A STUDY OF THE EFFECTS OF SPECIAL PREPARATION ON GRE ANALYTICAL SCORES AND ITEM TYPES

Spencer S. Swinton
and
Donald E. Powers

GRE Board Research Report GREB No. 78-2R
ETS Research Report 82-1
January 1982

This report presents the findings of a research project funded by and carried out under the auspices of the Graduate Record Examinations Board.

EDUCATIONAL TESTING SERVICE, PRINCETON, NJ

BAIRD, L. L. An Inventory of Documented Accomplishments. GREB No. 77-3R, June 1979.


BAIRD, L. L. The Relationship Between Ratings of Graduate Departments and Faculty Publication Rates. GREB No. 77-2aR, November 1980.

BAIRD, L. L. and KNAPP, J. E. The Inventory of Documented Accomplishments for Graduate Admissions: Results of a Field Trial Study of Its Reliability, Short-Term Correlates, and Evaluation. GREB No. 78-3R, August 1981.


CENTRA, J. A. How Universities Evaluate Faculty Performance: A Survey of Department Heads. GREB No. 75-5bR, July 1977. ($1.50 each)

CENTRA, J. A. Women, Men and the Doctorate. GREB No. 71-10R, September 1974. ($3.50 each)


CLARK, M. J. Program Review Practices of University Departments. GREB No. 75-5aR, July 1977. ($1.00 each)


HARTNETT, R. T. Sex Differences in the Environments of Graduate Students and Faculty. GREB No. 77-2bR, March 1981.

HARTNETT, R. T. The Information Needs of Prospective Graduate Students. GREB No. 77-6R, October 1979.


LANNHOLM, G. V. and PARRY, M. E. Programs for Disadvantaged Students in Graduate Schools. GREB No. 69-1R, January 1970.


ROCK, D.; WERTS, C. An Analysis of Time Related Score Increments and/or Decrements for GRE Repeaters across Ability and Sex Groups. GREB No. 77-9R, April 1979.


SCHRADER, W. B. GRE Scores as Predictors of Career Achievement in History. GREB No. 76-1bR, November 1980.

SCHRADER, W. B. Admissions Test Scores as Predictors of Career Achievement in Psychology. GREB No. 76-1aR, September 1978.

WILD, C. L. Summary of Research on Restructuring the Graduate Record Examinations Aptitude Test. February 1979.


WILTSIE, R. G. Doctoral Use of Foreign Languages: A Survey. GREB No. 70-14R, 1972. (Highlights $1.00, Part I $2.00, Part II $1.50).

A Study of the Effects of Special Preparation on GRE Analytical Scores and Item Types

Spencer S. Swinton

and

Donald E. Powers

GRE Board Research Report GREB No. 78-2R

Copyright ©1982 by Educational Testing Service. All rights reserved.
ABSTRACT

A brief curriculum of special preparation for the analytical section of the GRE Aptitude Test was developed and administered to a sample of self-selected GRE candidates at a large eastern university. Students' reactions to the materials and procedures used were obtained. The effects of the program on analytical scores were determined by comparing the scores subsequently obtained by the candidates who undertook special preparation with the scores of all the other candidates who took the test on the same date at the university. Analyses revealed an effect on analytical scores that was both practically and statistically significant. The effect stemmed from improved performance on two of the three analytical item types included in the 1977-1981 analytical section. Hypotheses concerning the components of the special preparation that were responsible for the improvement are discussed.
A Study of the Effects of Special Preparation on GRE Analytical Scores and Item Types

The effects of special preparation on test scores are of interest to test specialists in several respects. First, a test's susceptibility to score changes that result from various interventions may reflect directly on the validity with which the intended construct is measured. For example, tests designed to reflect student achievement are generally more sensitive to certain interventions, even those that are short-term in nature, than are aptitude tests, which are intended to measure abilities or skills that are developed over a longer period of time. For aptitude tests, sensitivity to short-term special preparation of the kind studied here may indicate that a substantial proportion of the variation in test scores is due to differences in examinees' facility with the methods of assessment or familiarity with the format of items, in contrast to differences in the presumably more stable trait that is being measured. Second, if special preparation increases the test score variance due strictly to differential facility in taking the test, then the predictive power of the test may be impaired since it is unlikely that the same method-related variance would underlie any important criterion. And finally, because all candidates may not have equal access to special preparation, a potential for inequity exists, especially for admissions tests, since any benefits may accrue disproportionately to those examinees who have the time and the financial means to obtain special preparation.

The Restructured GRE Aptitude Test

The analytical section of the GRE Aptitude Test was introduced in 1977 as a result of a major restructuring of the test (Miller & Wild, 1979). The analytical portion of the test is intended to measure analytical reasoning abilities that, like the verbal and quantitative skills measured by the test, are assumed to develop over a relatively long period of time.

The item types originally used in the analytical section of the test are described briefly as follows:

1. Analysis of Explanations: A fact situation paragraph and a surprising result, given the situation, are presented. Approximately 10 numbered statements follow. Each statement is to be evaluated separately from the other statements as to whether it:

   (A) contradicts something in the fact situation, the result, or both taken together;

   (B) presents a possible adequate explanation of the result;
(C) is deducible from the fact situation, the result, or both;

(D) is relevant to a possible explanation of the result; or

(E) is irrelevant to an explanation of the result.

The choices are to be considered in order, and the first choice that cannot be eliminated is correct. For example, if a statement is both deducible from the fact situation and irrelevant to an explanation, the statement should be marked "C," rather than "E." Common sense, rather than formal logic, is presupposed in deciding whether the explanations are adequate or whether statements are inconsistent or deducible.

2. Logical Diagrams: Five diagrams, each consisting of three circles with differing patterns of inclusion, exclusion, and overlap are presented. Each question consists of the names of three classes (e.g., mammals, cows, birds). The diagram that best illustrates the logical relationship among the three classes is to be chosen. The size of the circles does not indicate the relative size of the classes.

3. Analytical Reasoning I: A passage or a set of statements giving a rule set or describing a logical structure is presented. Several multiple-choice questions requiring deductions from the rule set or structure are asked. These items are logical puzzles, and it is frequently helpful to draw a diagram or table in answering them.

4. Analytical Reasoning II: In this subtype of analytical reasoning items, an argument is presented and the examinee is asked to identify assumptions or additional points that would strengthen or weaken the argument or to detect errors in reasoning.

Purpose

Anastasi (1976) has suggested that the effects of special preparation may depend on the ability and previous experience of examinees, the nature of the test and the items it contains, and the amount and type of special preparation provided.

The present study was undertaken because the results of the many previous studies of the effects of special preparation on the Scholastic Aptitude Test (e.g., Coffman & Parry, 1967; Dear, 1958; Dyer, 1959; Evans & Pike, 1973; Frankel, 1960; French & Dear, 1959; Pallone, 1961; Roberts & Oppenheim, 1966; Whitla, 1962) were not, because of differences in the characteristics of examinees and the types of test items involved, readily generalizable to the new
analytical section of the GRE Aptitude Test.\textsuperscript{1} Evans (1977) investigated the susceptibility of the quantitative section of the GRE Aptitude Test to special preparation, but severe sample attrition rendered the results inconclusive. Also, each of the three new analytical item types (analysis of explanations, logical diagrams, and analytical reasoning) involve relatively complex formats and two use fixed response options, which have been suggested to be more susceptible than simpler formats to coaching and practice (Vernon, 1954; Loret, 1960). Because these item types have not been used before in national testing programs, they are also much less familiar to candidates than are various other question types.

The major objective of the study was to develop and pilot test a brief curriculum of special preparation for the analytical portion of the GRE Aptitude Test in order to assess the susceptibility of three analytical item types to short-term intervention designed to improve performance on the item types. The results were intended to aid decisions about the composition of the analytical section of the test when it becomes fully operational.\textsuperscript{2} Originally, the study reported here was intended only as a pilot study to determine student reactions to the special preparation curriculum, which then would be revised and administered to larger samples of GRE candidates under more rigorous experimental conditions. As we proceeded, however, the opportunity also arose to assess the effects of the special preparation that was administered.

Methods

Development of Materials

A variety of techniques were used to develop the special analytical preparation materials, including (a) the identification

\textsuperscript{1}Other studies of special preparation that were conducted concurrently with our study and/or completed after we began (e.g., Alderman & Powers, 1980; Federal Trade Commission, 1979) also involved the SAT and, hence, were not directly generalizable to the GRE Aptitude Test.

\textsuperscript{2}Since the analytical section was introduced, the GRE guide to using test scores and the score report form have emphasized the introductory nature of the new measure and has cautioned schools against using scores from this section until the relationship between the scores and academic performance in graduate school is established.
and review of existing preparation resources, such as the commercially available books on test preparation and the materials used in courses of special preparation for the GRE Aptitude Test, and (b) interviews with students—both prospective GRE candidates and students who had previously taken the test.

These techniques suggested that any special preparation materials should include (a) hints or strategies for approaching each of the three item types, (b) opportunity for extra practice on each item type, beyond that provided by the test sponsor, and (c) feedback or knowledge of results in the form of explanations of the extra practice questions. The hints included, for example, suggestions to learn the instructions and response formats ahead of time and to be aware of the particular demands of each item type (e.g., the need to consider choices in a sequential manner for analysis of explanations items and the desirability of drawing an appropriate diagram for certain analytical reasoning items). The practice tests were two full-length analytical sections that directly paralleled the analytical section of the Aptitude Test. One is included in the GRE Information Bulletin that is provided to all test registrants; the other was developed especially for the study. A set of explanations, giving for each item the rationale underlying the correct choice and a procedure for arriving at the correct answer, was developed for each of the practice tests.

Sample Selection

Arrangements were made with the staff of the study improvement center at a large eastern university to offer a free course to prepare for the GRE Aptitude Test. Students were recruited with an advertisement in the student newspaper and posters placed at strategic points on the campus, as well as with help from the staff of the study improvement center. The course was billed as preparation for the GRE Aptitude Test, with emphasis on the new analytical section of the test. Initially the course attracted more than 40 students, most of whom anticipated taking the test in October or December or in both months.

Course Description

The course was structured in such a way that all of the six sessions (a total of seven contact hours) that were devoted to the analytical section of the test preceded the October administration. The average attendance for these six sessions was 31 students. The remaining sessions, devoted to practice on various aspects of the verbal and quantitative sections of the test, were held primarily as a courtesy to students for helping us learn more about the analytical test. These sessions were held after the October
administration. The seven hours of analytical preparation included administration of the analytical part of the sample test in the GRE Information Bulletin, discussion of each of the three analytical item types and the relevant hints and strategies for each type, making available additional practice analytical items for homework, explanation of correct answers to practice items, and, finally, administration and discussion of the second full-length sample analytical test.

Thus, the special preparation considered in this study was relatively brief, focusing on strategies and techniques specific to the analytical portion of the GRE Aptitude Test and to its specific item formats, rather than on development of the cognitive abilities the test is designed to measure. This definition is consistent with Gulliksen's (1950) position that an intervention that raises the criterion to the same degree that it raises the prediction test score is not considered to be special test preparation, but instruction in skills and/or knowledge measured by the test and represented in the criterion.

Data Collection

At the end of the course, students were asked to complete a brief evaluation form giving their reactions to the appropriateness and the helpfulness of the special preparation. The GRE Aptitude Test scores of course participants and the scores of all other candidates who took the test in October at the university were retrieved from test score files. Background data, which are supplied by all GRE test registrants, were also retrieved from the files for these two groups of candidates.

Analysis

The major dependent variables investigated were (1) total GRE analytical score expressed on the 200–800 point GRE converted score scale and (2) raw scores on each of the three analytical item types. Two separate analyses were also computed with either GRE-verbal or GRE-quantitative scores as the dependent variables. Using GRE-V and GRE-Q as dependent variables served primarily to assess any unanticipated effects of analytical preparation on these scores.

Independent variables included sex, undergraduate major field, and self-reported undergraduate grade-point average (UGPA). GRE-V and GRE-Q scores were also entered as independent variables in the analyses of analytical scores.
A dichotomous variable representing treatment condition was entered in the multiple regression analysis with the other independent variables. Each of the four analytical dependent variables (total score and item type subscores) was regressed on the independent variables and the dichotomous treatment variable. Interactions between major field and UGPA were also allowed to enter into the regression analysis if these interactions had significant regression weights. These four analyses were repeated with GRE-V and GRE-Q as additional covariates.

In summary, the tryout of the analytical practice material was a quasi-experimental study in which students volunteered to participate. They were not randomly assigned to treatment and control groups. Therefore, we would expect the treatment group to differ from other students who took the test at this university even before the treatment took place, and to continue to differ after the treatment even on scores that were not the subject of the special preparation. That is, a number of personal characteristics (e.g., motivation or anxiety) that may have been related to students' decisions to seek special preparation for the test may be reflected in their subsequent performance on each section of the Aptitude Test, thus giving the appearance of effects from the preparation they received. Because the special preparation was restricted to the analytical portion of the test, the analytical scores reflect not only the personal characteristics related to student self-selection but also the effects of the special preparation course. By using multiple regression procedures and several independent variables, some of the effect of self-selection that is related to these variables could be reduced, thus giving less biased estimates of the effects of special preparation on analytical scores. The major assumption underlying the use of this analytic strategy is that the personal characteristics related to self-selection are reflected in the scores on each section of the test, but that the treatment has an effect on only the analytical portion of the test. The likely consequences of failing to meet these assumptions are discussed below.

Results

Student Reactions

Generally, students' comments suggested that they were pleased with the course and with the preparation techniques that were used. An analysis of student evaluation forms revealed an unanimous agreement that the course would improve performance on the analytical section of the test. Receiving hints about specific item types and
studying additional practice items were rated as the most useful preparation activities.

Description of the Sample

Table 1 gives descriptive statistics for the treatment and control groups. As can be seen, the groups differed on several background variables. Students who undertook special preparation reported significantly higher undergraduate grade averages than those who did not, and were more frequently enrolled in a humanities major than their counterparts at the same university. Students in the special preparation program obtained slightly, but not significantly, lower GRE-V and GRE-Q scores and significantly higher GRE-A scores. Their higher analytical scores stemmed from significantly better performance on both analysis of explanations items and logical diagrams items. The pattern of higher undergraduate grades and somewhat lower GRE-V and GRE-Q scores suggests that the students who volunteered for coaching may have been those who tend to obtain lower test scores than would be expected from their academic performance. The higher GRE analytical scores suggest that the special preparation may have been effective in raising scores on this section of the test.

Effects on Total Scores and Item Type Scores

Regression analyses of total scores. Table 2 gives regression results for quantitative and verbal scores, taking sex, undergraduate major, and grade-point average into account to estimate the number of points by which the self-selected treatment group (25 of whom received GRE scores) differed from the control group (the 415 students who took the test at the same center, but who did not choose to attend these preparation sessions). Although the treatment was not designed to influence verbal and quantitative scores, there exists the possibility that it may have had some generalized positive effect on testwiseness or confidence and therefore on verbal and quantitative scores. To the extent that the treatment may have, unexpectedly, improved verbal and quantitative scores, using GRE-V and GRE-Q scores as covariates with the treatment variable in the regression analyses would tend to underestimate the effects of the treatment on analytical scores.

Table 2 shows a 24.6 point deficit in quantitative scores for the treatment group, taking the effects of background variables into account. As is usual in analyses of quantitative scores, the regression weights indicate that females scored lower than males, and physical science majors higher than both social science and humanities majors, with biological science majors about midway between. Undergraduate grade-point average made the most significant contribution to the prediction.
Table 1
Characteristics of Treatment and Control Groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Treatment (N=25)</th>
<th>Control (N=415)</th>
<th>Significance of Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>S.D.</td>
<td>Mean</td>
</tr>
<tr>
<td>GRE-V</td>
<td>495.5</td>
<td>96.8</td>
<td>498.3</td>
</tr>
<tr>
<td>GRE-Q</td>
<td>513.2</td>
<td>134.9</td>
<td>535.5</td>
</tr>
<tr>
<td>GRE-A</td>
<td>591.8</td>
<td>70.8</td>
<td>530.7</td>
</tr>
<tr>
<td>Analytical Item Types</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis of Explanations (40 items)</td>
<td>Mean 28.6</td>
<td>S.D. 4.3</td>
<td>Mean 24.2</td>
</tr>
<tr>
<td>Logical Diagrams (15 items)</td>
<td>Mean 12.1</td>
<td>S.D. 2.1</td>
<td>Mean 10.7</td>
</tr>
<tr>
<td>Analytical Reasoning (15 items)</td>
<td>Mean 7.5</td>
<td>S.D. 2.8</td>
<td>Mean 7.2</td>
</tr>
<tr>
<td>UGPA&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Mean 6.05</td>
<td>S.D. .79</td>
<td>Mean 5.55</td>
</tr>
<tr>
<td>Female (%)</td>
<td>68.2</td>
<td>54.0</td>
<td>n.s.</td>
</tr>
<tr>
<td>Social Science Majors (%)</td>
<td>27.3</td>
<td>39.0</td>
<td>n.s.</td>
</tr>
<tr>
<td>Physical Science Majors (%)</td>
<td>18.2</td>
<td>17.9</td>
<td>n.s.</td>
</tr>
<tr>
<td>Humanities Majors (%)</td>
<td>27.3</td>
<td>12.6</td>
<td>p &lt; .05</td>
</tr>
<tr>
<td>Biological Science Majors (%)</td>
<td>27.2</td>
<td>30.5</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

<sup>a</sup>UGPA is expressed on a scale of 1-7 with D- or lower = 1, C- = 2, C = 3, B- = 4, B = 5, A- = 6, A = 7.
Table 2
Effects of Special Preparation on GRE Scores

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Multiple Correlation</th>
<th>Estimated Treatment Effect</th>
<th>Standard Error of Effect Estimate</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRE-Verbal</td>
<td>.29&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-14.3</td>
<td>22.5</td>
<td>n.s.</td>
</tr>
<tr>
<td>GRE-Quantitative</td>
<td>.52&lt;sup&gt;a&lt;/sup&gt;</td>
<td>-24.6</td>
<td>22.2</td>
<td>n.s.</td>
</tr>
<tr>
<td>GRE-Analytical</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis I</td>
<td>.37&lt;sup&gt;a&lt;/sup&gt;</td>
<td>48.7</td>
<td>22.9</td>
<td>p&lt;.05</td>
</tr>
<tr>
<td>Analysis II</td>
<td>.80&lt;sup&gt;b&lt;/sup&gt;</td>
<td>65.8</td>
<td>15.0</td>
<td>p&lt;.01</td>
</tr>
</tbody>
</table>

Note. Results are based on 25 students who participated in the special preparation course and 415 students who did not. Analysis I and Analysis II are identical except for the addition, in Analysis II, of two more covariates—GRE-V and GRE-Q.

<sup>a</sup>Covariates were sex, undergraduate major field, and self-reported undergraduate grade-point average.

<sup>b</sup>Covariates were sex, undergraduate major field, self-reported undergraduate grade-point average, GRE-V score, and GRE-Q score.

<sup>c</sup>Effect is expressed in terms of scaled score points on the 200-800 GRE Aptitude Test scale.
Table 2 also shows a 14.3 point estimated deficit for the treatment group on the verbal section of the test. As with the quantitative score, this difference was not statistically significant. Although females also scored lower on the verbal section when major field and grade-point average were taken into account, the discrepancy was smaller and not statistically significant.

Table 2 further gives the results of a similar regression on analytical scores (Analysis I). It should be reiterated that, because of the quasi-experimental nature of the study, the estimated treatment effect represents the combination of preexisting differences between the treatment and comparison groups as well as any effect attributable to the treatment. The estimated treatment effect, 48.7 points in favor of the treatment group, suggests either that students were capable of highly specialized self-selection (e.g., on the basis of both low verbal and quantitative ability and high analytical ability) or that the treatment had a positive effect. Assuming that, without special preparation, the treatment group would have had analytical scores comparable to their V and Q scores, we would expect an analytical score deficit of nearly 20 points for the treatment group (i.e., the average of the V [14.3] and Q [24.6] deficits, since GRE-A scores are scaled to be midway between V and Q scores) in contrast to the positive 48.7 point effect.

Table 2 also contains the results of adding verbal and quantitative scores to the regression analysis (Analysis II). Because these concurrent scores could themselves have been raised by the treatment and, therefore, are not technically appropriate covariates, their use may give a conservative estimate to the question, How much higher were the treatment group's analytical scores than would be expected from their verbal and quantitative scores and other background information? The answer is a practically and statistically significant 66 points. With a multiple correlation coefficient of .80, it would appear that any remaining bias due to unmeasured correlates of self-selection could not be sufficient to reduce this effect to nonsignificance.

In summary, it appears that scores on the analytical section of the GRE Aptitude Test, as currently constituted, may be improved under at least some conditions by relatively short-term interventions that focus primarily on practice and familiarization. The question addressed below is the extent to which this score improvement is associated with gains on each of the three analytical item types.

Regression analyses of item scores. As shown in Table 1, the control group exhibited a lower mean and a higher standard deviation on each analytical item type, reversing their slight superiority on verbal scores and quantitative scores.
**Table 3**
Effects of Special Preparation on Three Analytical Item Type Scores

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Multiple Correlation</th>
<th>Estimated Treatment Effect</th>
<th>Standard Error of Effect Estimate</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Analysis of Explanations (40 items)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis I</td>
<td>.32&lt;sup&gt;a&lt;/sup&gt;</td>
<td>3.57</td>
<td>1.55</td>
<td>p&lt;.05</td>
</tr>
<tr>
<td>Analysis II</td>
<td>.71&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.54</td>
<td>1.16</td>
<td>p&lt;.01</td>
</tr>
<tr>
<td><strong>Logical Diagrams (15 items)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis I</td>
<td>.31&lt;sup&gt;a&lt;/sup&gt;</td>
<td>1.17</td>
<td>.70</td>
<td>n.s.</td>
</tr>
<tr>
<td>Analysis II</td>
<td>.62&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.54</td>
<td>.58</td>
<td>p&lt;.01</td>
</tr>
<tr>
<td><strong>Analytical Reasoning (15 items)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analysis I</td>
<td>.31&lt;sup&gt;a&lt;/sup&gt;</td>
<td>.05</td>
<td>.64</td>
<td>n.s.</td>
</tr>
<tr>
<td>Analysis II</td>
<td>.61&lt;sup&gt;b&lt;/sup&gt;</td>
<td>.46</td>
<td>.54</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

*Note. Results are based on 25 students who participated in the special preparation course and 415 students who did not. Analysis I and Analysis II are identical except for the addition, in Analysis II, of two more covariates—GRE-V and GRE-Q.*

<sup>a</sup>Covariates were sex, undergraduate major field, and self-reported undergraduate grade-point average.

<sup>b</sup>Covariates were sex, undergraduate major field, self-reported undergraduate grade-point average, GRE-V score, and GRE-Q score.

<sup>c</sup>Effect is expressed in terms of number of items correctly answered for each item type subscore.
Table 3 gives the regression results for each item type, both when concurrent GRE-V and GRE-Q scores were excluded from the analysis (Analysis I), and when they were included (Analysis II). As can be seen, a significant effect was noted for each analysis. The 4.54 raw score point effect (p < .01) estimate when V and Q scores were included in the analysis corresponds to .61 of the control group's standard deviation on this item type. Inspection of the item analysis data showed that the students with special preparation reached, on the average, about half an item more on the timed section containing analysis of explanations than did the control group, suggesting that increased efficiency in responding cannot by itself account for the sizable effect.

An effect on logical diagrams of 1.54 raw score points (p < .01), about .45 control-group standard deviations, was computed when GRE-V and GRE-Q scores were included. Because of the greater variability in scores on this item type, the estimate of 1.17 when GRE-V and GRE-Q scores were excluded was not significant. Because certain common logical fallacies (e.g., stereotypic classification) may have been eradicated, at least temporarily, during the special preparation, we are reluctant to ascribe the effect on logical diagrams totally to "coaching" as usually defined. If this training is retained, it may improve performance in graduate school, as well as performance on this item type.

There was no significant treatment effect on analytical reasoning items. The fact that analytical reasoning scores were not susceptible to our special preparation does not assure us that they would not respond to some other form of instruction. However, the fact that each item set presents a novel puzzle with different response choices means that familiarity with sample problems is less likely to transfer than it is for logical diagrams, and that performance cannot be improved solely by gaining familiarity with a fixed response format as with analysis of explanations.

Discussion

Responsiveness to instruction does not necessarily disqualify an item type designed to measure developed academic ability. An item type totally immune to experience would have to measure some innate or unconscious quality, which would probably be less relevant to academic success than developed ability is. However, over-responsiveness of an item type to intensive study does raise questions of equity.
The problem in this case, however, is not so much one of money as one of time. The special preparation given here contained no "inside information" or tricks, and any candidate who took the sample test under timed conditions, learned instructions for the item formats completely, thought carefully about the rationale behind the correct answers, and applied pacing and guessing strategies as given in the GRE Information Bulletin would probably have done as well on the analytical section as did the experimental group. However, it appears that the penalty for lacking the motivation or time to prepare may be excessive for the analysis of explanations item type as it is now constituted, and possibly for the logical diagrams item type as well. There are, to be sure, a variety of other criteria that need to be considered in making choices among item types.

One recommendation based on both a logical analysis of the test format and the results of the study reported here, is that, were the items to continue in the test, and if it were operationally feasible, examinees be allowed to read the analysis of explanations instructions before they begin to work on the test. Moving these complex instructions outside the timed section and providing a realistic sample problem with feedback before testing might bring the item type's susceptibility to practice closer to the levels appropriate for testing developed ability.
References


Boldt, R. R. Comparison of a Bayesian and a Least Squares Method of Educational Prediction. GREB No. 70-3P, June 1975.


Campbell, J. T. and Donlon, T. F. Relationship of the Figure Location Test to Choice of Graduate Major. GREB No. 75-7P, November 1980.


Donlon, T. F.; Reilly, R. R.; and McKee, J. D. Development of a Test of Global vs. Articulated Thinking: The Figure Location Test. GREB No. 74-9P, June 1978.


Rock, D. A. The "Test Chooser": A Different Approach to a Prediction Weighting Scheme. GREB No. 70-2P, November 1974.

Sharon, A. T. Test of English as a Foreign Language as a Moderator of Graduate Record Examinations Scores in the Prediction of Foreign Students' Grades in Graduate School. GREB No. 70-1P, June 1974.


