AOTBOR
TITLE

IHSTITOTIOH
POB DATE
mots
AVAILABLB FROB

BDRS PRICE
DESCRIPTORS

IDEMTIFIBS

Marco, Gary 1.
The Performance of Pirst-lear Graduate Students on the Graduate Record Braninations. Gradurite Becord Bxauinations Special Eefcrt minber 68-2.
Bducationsi Testing Servić; Princetca; f. J. Jul 68
30p.; llot available $\pm \mathrm{n}$ bard cops dee to print gunlifty
-Graduate Record graminationé, Bducational Testing Service, princetod. Ify Jerscy 08541 (free while supplies last)
 - Achievenent Tests; Aftitvie iests: college Entrance Eraninations; College gajore; © Eraduate Students; Graduate Study; Eigher Education: Horas; -Scores; Standard Brror of Heasurenent; Statistical Data: Test Interpretaticn; 略ighted Scares *Graduate acord freainations

ABSTRACT
Woraative data vere ottaiacd on the ferfcrance of first-year graduate students on the frtitede Test and divanced Tests of the craduate Record Examinations. Tbs population ccpsisted of students encolled as full-time graduatc gtudfats for the first tiae if the fall of 1964 sin a collega or veiveraity belonging to the Council of tractate Schools (CGS). Sutpopulaticys wthin exch institution vere conposed of stuatents enfelici in grsduatf study in 18 najorfields-bioloyy, business, chienstry, ccorncaics; aducation. -atreatics, philosophy, phyaical ducaticn, phyics. psychology. sociology. and Spanish. Hajors in thesa fields.tcck both the aptitude Test and the appropriste Adyinced Test. Mcren ire providud for the Verbal and quantitative sectionslof the iptitede test by ajor field of study and for perioriance on the 15 Adrance'd Testa.. Accuracy of the data is discussed and the appendicen contain the naies of the institutions participating, the plan of data collincticn, seights used in prepardag the freguency distributicne, and sfecific fields classified under particular fields for mafiling purposer. (EOF)


The Performance of First.Year Graduate Students $\because$ on the
Graduate Record Examinations

Cary L. Marco

July 1968
Graduate Record Examinationt Special Report * ${ }^{2}$ Nomber 68.2.

ERIC

## GRE

THE PERFORMANCE OF FIRST-YEAR GRADJATE STUDENTS
ON THE GRADUATE RECORD EXAMINATIONS

Cary L. Marco

> The National Program for Oraduate School Selection is under the policy direction of the Craduate Record Examinations Board, an affiliate of the Association of Graduate Schools, the "Council of Cradtrate Schools, and Educational Testing Service.

Copyright © 1968 by Educational Testing Service. All rights reserved.

ACKNOWTEDGMENTS

The author gratefully acknowledges the assistance of Dr. Philip R. Harvey, who assumed the administrative responsibility for collecting the data for this study and assisted in writing the introductory section. He is also-grateful for the assistance of Miss Barbara Weeks, who wrote the computer programs. The author thanks particularly Dr. William Schrader for his help in planning the study and for his continued interest in and support of the study while it was being inplemented.
1.2
CONTEVTS
$*$ Page
PURPOSE AND BAEKCROUND ..... 5
POPULATION AND. SAMPIE. ..... 6
PERFORMANCE ON THE GRE TESTS ..... 9
ACCURACY OF THE DATA ..... 14
REFERENCES ..... 20
APPENDICES
A. Colleges and Universities which Participated in the CRE Study of First-Year Graduate Students. ..... 23
B. Plan for Data Collection ..... 24
C. Weights Used in Preparing the Frequency Distribution. ..... 27
D. Specific Fields Classified Under Particular Major Fields for Sampling Purposes. ..... 29
$*$

## TABLES

Page, 1:
Percent of Graduate Students Scoring Lower than Selected Scaled Scores on the Verbal Section of the Aptitude Test.. ..... 11
Percent of Graduate Students Scoring Lower than Selected Scaled Scores on the Quantitative Section of the Aptitude Test. ..... 12
Percent of Graduate Students Scoring Lower than
Selected Scaled Scores on the Advanced Tests. Percent of Graduate Students Scoring Lower than
Selected Scaled Scores on the Advanced Tests. ..... 13 ..... 13Number and Percentage of Students in a ParticularSample Who Had Advanced Test or Aptitude Test
Scores ..... 16

Until the 1967-68 academic year, the primary interpretive materials for the Graduate Record Examinations (GRE) were the scores for basic reference groups of college seniors, who for the most. part took the tests in 1952 (1). At its meeting in 1964, the Committee on Testing of the Association of Graduate Schools recomended that new data--data representative of the performance of enrolled first-year graduate students...be collected and published. This rec̆omendation resulted in the initiation. of the current study. The purpose of this study was to obtain data on 4 the performance of firstiyear graduate students on the Aptifude Test and the Advanced Tests of the Craduate Record Examinations. "R

Today the primary interpretive data for GRE tests are norms for National Program cqndidates tested from May 1964 through April $1967^{\circ}$ (1). However, since not all National Program candidates 免ttend graduate school, the data on first-year graduate students described here should be of further use to graduate schools and fellowship sponsors in interpreting GRE scores.

## POPULATION AND SAMPLE

The population of the study consisted of those students who enrolled as full-time graduate students for the first time in the fall of 1964 in a college or university belonging to the 228 -member Council of Graduate Schools (CGS). Students from countries where the primary language is not English and students who had physiçal handicaps (such as blindness) which might have affected their GRE scores were excluded from the population.

Subpopulations were composed of students enrolled in graduate study in 19 major field categories --18 major fields for which GRE Advanced Tests were appropriate and a nineteenth category called "other."

The Advanced Tests for the 18 major fields were dis follows:


A two-stage sampling plan was used to select the samples for each of the various subpopulations. The first stage involved the selection of universities and colleges from the CGS members. CGS member institutions were grouped into categories according to the number of graduate degrees awarded in 1961-62 as reported by the U.S. Office of Education (3). The
$\qquad$
*The name of this test will be changed to Political Science in the fall of 1968.
thirty institutions granting the most graduate degrees in 1961-62 were included in the sample. Also included were 15 institutions randomly selected from the 60 universities or colleges with the next highest. number of conferred grachate degrees, and 15 randomly selected from the remaining institutions. A greater proportion of the large universities was included in the sample, so that those fields in which only a small number of students were enrolled would be adequately represented.

Of the 60 institutions in the sample invited to participate in the study, 52 agreed to furnish rosters of students and 35 agreed to test each of those students who had not taken the Aptitude Test or an Advanced Test in his major field. Eight schools did not participate. (A list of the participating colleges and universities in each of the three size categories is given in Appendix A.)

The second stage of sampling involved the selection of students within each institution. Some of the participating institutions selected their own samples following a plan'suggested by ETS (see the "Plan of Daṭa Collection" in Appendix B), while other institutions sent rosters to ETS for sampling. A spaced sample was chosen according to the ETS plan. The sampling ratios, which varied from field to field, were designed to yield a minimum of 300 studente in the group for each of the 19 major field classifications. Thus, smaller percentages of students were. sampled from the more popular major fields.

After the second stage of sampling was finished, the 1963 and 1964 ETS files were searched to determine which of the students in the sample had appropriate $G R E$ scores. Each student who attended an institution

POPULATION AND SAMPLE - Page 8
that agreed to conduct special testings was asked to take the Aptitude
Test if Aptitude scores were missing and the Advanced Test in his major field (provided his undergraduate and graduate major fields were the same) if the Advanced Test .score was missing. The special testings were conducted late in the first semester and early in the second semester of the 1964-65 academic year.

## PERFORMANCE ON THE GRE TESTS

How well did the first-year graduate students perform on the GRE tests? Score data for Verbal Ability, Quantitative Ability, and the Advanced Tests are given in Tables 1-3. These tables show the percent of graduate students scoring lower than selected scaled scores. It is important to note that, because of missing scores, the Advanced Test distribution for a particular major field was not based on the same students as the Aptitude Test distribution for that field.

The data in Tables $1-3$ are based on weighted frequency distributions. Weighting was necessary to estimate the population frequency distributions from the sample data. The weights were determined primarily from the number of full-time first-year graduate students reported by the U.S. Office of Education for each of the participating institutions (4).

The score frequencies for a particular field and institution were adjusted for (a) the size of the institution's class of full-time firstyear graduate students, (b) the number of full-time first-year graduate .students, in the size category (large, medium-sized, small) in which the institution was classiffed, and (c) the number of CGS schools in that particular size category. (The actual formula used in computing the weights is given in Appendix C.)

Since the score data for the fields of business, education, and physical education were probably biased (see next section), neither

PERFORMANCE - Page



 overall Aptitude Test distributions.

$$
12
$$

挔




AOCURAEY OF THE DATA

How accurate are the data reported in Tables 1-3? One way to answer this question is to determine confidence intervals for the various means of the score distributions However, such a procedure can be used legitimately only if the sampling is random. Since there is reason to, believe tha $\ddot{t}$ the final samples were fottrandom, this procedure was not considered appropriate. Another way to approach the question of accuracy is to examine (a) the representativeness of the original sample and (b) the muber of students in the original sample with appropriate (RE scores.

The samples In three fields were considered unrepresentative because there were insufficient mumbers of students in the samples from particular graduate schools. For example, because of a communication problem, no students from the University of Michigan were represented in the sample for education, even though a lange number of firgt-year grachuate students were enrolled in this field at the University of Michigan in the fall of 1964. Thas, the sample for education could not be considered representative. The other flelds which were judged to have biased data because of inadequate representation from some colleges or universities were tustness, physical education; and "other".

The samples in some fields werer also considered unrepresentative because of ambiguity in the definition of subpopulations. Since gradusts

DATA - - Page 15
school personnel should be more able than anyone else to determine the subject-matter areas for which a given Advanced Test is appropriate, the colleges and universities that selected their own samples were asked to define the subpopulations for the various fields. Educational Testing Service defined the subpopulations when it selected the samples. (Subjectmatter areas included under a particular major field heading when ETS selected the samples are indicated in Appendix D.) As a result, certain ambiguities arose. The fields most seriously affected were probably biology, education, and "other".

Even if samples of students are representative, the data might still be biased if appropriate test scores are not available for all of these students. In this study, not all of the students in the original sample had scores. Students at if institutionst were not offered special testings, and only about $30 \%$ of the students asked to participate in the special testings offered by the other 35 institutions actually were tested. The number and the percent of students in the original sample for this study who had Advanced Test or Aptitude Test scores are reported in Table $\mathrm{L}^{2}$. The percent of students who had appropriate test scores ranged from 9\% (in business) to $66 \%$ (in physics) for the Advanced Tests, and from $21 \%$ (in business) to 73\% (in psychology) for the Aptitude Test. Because the percentages of students with $\operatorname{CRE}$ scores were so low, the Advanced Test distributions in'the fields of business, dducation; and physical education and the Aptitude Iest distributions in the field of business were probably biased.

Another way to approach the question of accuracy is to compare the distributions with data for other groups. Two such groups are the

DATA - Page $16^{\circ}$

Table h

NUMBER AND PERCENTAGE OF STUDENTS IN A PARTICULAR SAMPLE WHD HAD ADVANCED TEST OR APTITUDE TEST SCORES

| Major Field | Nuniber in Sample | Students with Advanced Test Scores |  | Students with Aptitude Test Scores |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | Number | 8 | Nhamer | 9 |
| Brology | 243 | 129 | 53 | 163 | 67 |
| Business | 332 | 31 | 9) | 69 | 21 |
| Chemistry | 458 | 268 | 59 | $280^{\circ}$ | 61 |
| Economics | 445 | 234 | 53 | 260 | 58 |
| Education | 353 | 69 | 20 | 116 | 33 |
| Engineering | 348 | 159 | 46 | 175 | 50 |
| French | 397 | . 152 | 38 | 146 | 37 |
| Ceology | 202 | 99 | 49 | 126 | 62 |
| Oovernment | 470 | 164 | 35 | 212 | 45 |
| Hisistory | 505 | 232 . | 46 | 259 | 51 |
| Literature in English | 469 ' | 232 | 49 | 237 | 51 |
| Máthematics | 506 | 265 | - 52 | 289 | 57 |
| Philosophy | 436 | 188 | 43 | 228 | 52 |
| Physical Education | 149 | 36 | 24 | 55 | 37 |
| Physics | 346 | 227 | 66 | 233 | 67 |
| Psychology | 355 | ${ }^{2} 22$ | 63 | 258 | 73 |
| Sociology | 555 | 187 | $34^{\circ}$ | 251 | 45 |
| Spanish | 1 251 | 82 | 33 | 77 | 31 |
| Other | 1044 | -- | -- | 378 | 36 |

DATA - Page 17

GRE basic reference groups of college seniors (1) and 1963-64 CRE National. Program Candidates (2). In general, the means for- the Aptitude Test and the Advanced Tests from the current study were somewhat larger than those reported in the 1963-64 study and much larger than the means for the college seniors. The Verbal Ability and Quantitative Ability means of first-year graduate students were approximately 15 points higher than the corresponding means of the 1963-64 National Program candidates. The Verbal Ability mean for first-year graduate students was, on the average, 55 points higher than; the corresponding mean of eollege seniors, while the Quantitative Ability . mean was about 75 points higher: The Advanced Test scores of first-year . graduate students averaged about $25^{\prime}$ points higher than the scores of National Program candidates and about 100 points higher than the scores of college seniors. However; the means for first-year graduate students were slightly lower than the means for National Program candidates in biology (Verbal Ability mean), goverment (Quantitative Ability mean) and mathematics (Advanced Test mean)'.

A third comparison group is the 1964-67 norms group, who also were National Program candidates (see the Handbook for the Interpretation of CRE Scores, 1967-68). No data by major, field are available on the 1964-67 group for the Aptitude Test. However, the overall Verbal Ability and Quantitative Ability means for first-year, graduate students were approximately 25 points higher than the means for the 1964-67 norms group. The Advanced Test means for first-year graduate students in all fields but mathematics were also higher than the means for the 1964-67 group. In mathematics the mean for the norms group slightly exceeded the mean for first-year graduate students. In the other fields, the means of first-year

DATA. - Page 18
graduate students ranged from 23 points higher in biology to 58 points higher in geology.

Generally; these findings seem reasonable. Most of the persons who took the $G R E$ tests in the National Program intended to go to graduate school or in some cases were already enrolled. However, a number of these persons were probably not admitted to CGS member schools, and others failed to attend any graduate school. Thus, one would expect the scores of the National Program candidates to be lower on the average than the scores of enrolled first-year graduate students. Likewise one would expect the scores of enrolled first-year graduate students to be substantially higher on the average than the scores of complete classes of college seniors, not all of whom actually attend graduate school.

That the means for first-year graduate students were lower than the means for National Program candidates in biology and government on the Aptitude Test and in mathematics on the Advanced Tests is somewhat disturbing. It might be that the means for first-year graduate students. in these fields were low. However, a large percentage of the National Program candidates were applicants for National Science Foundation (NSF) fellowships. Each applicant was required to take the Aptitude Test and an Advanced Test if one was offered in his field. Advanced Tests were offered in biology, chemistry, economics, engineering, geology, government, mathematics, physics, psychology, and sociology. The percent of applicants for NSF fellowships in the 1964-67 norms group was particularly high in chemistry ( $26 \%$ ), mathematics (25\%), physics (23\%), and geology ( $20 \%$ ), respectively. Thus, the Advanced Test means for these four fields especially and the Aptitude Test means for the $1964-67$ norms

DaTA - Page 19
group were somewhat inflated. If one removed the effect associated with the percentage and caliber of NSF fellowship applicants, the means for first-year graduate stucents would probably haye beeni consistently higher than the means for National Program candidates.
. In sumbary, then; the Aptitude and Advanced Test data for the fields of business, education, and physical education and the Aptitude Test data for "other" fields were probably biased and are not reported. However, the data for the other fields appear to be reasonably consistent . with comparative data from other sources, particularly when the percent5
ages and caliber of NSF fellowship applicants are considered. The reported data are probably fairly accurate estimates of the performance of full-time first-year graduate students who attended colleges or universities belonging to CGS in the fall of 1964.

1. Handbook for the Interpretation of GRE Sepres, 1967-1968. Educational Testing Service, Princeton, New Jersey, 1967.
2. Harvey, Philip R:, and Marco, Gary L., Aptitude and Advanced Test Scores of 1963-1964 National Program Candidates by Indergraduate Major Flelds. Craduate Record Ecaminations Special Report No. 65-3, Efucational Testing Service, Princeton, New Jersey, 1965.
3. Tolliver, Wayne E., Earned Dagrees Conferred, 1961-1962,

Bachelor's and Higher Dejeees. U.S. Office of Education Report No. OE-5LO13, Washington, D. C., 1963.
4. U:S. Office of Eduration, Mational Center for Educational Statistics, Enrollmel t for Master's and Higher Degrees, Fall 1964. U.S. Office of Education Report No. OE-54019-64, Washington, D. C., 1966.

22

APPENDICES

0


Appendix B
plan óf data collection
I.- Identification of students to be included in the sample

Major fields to be sanpled. In the oratuate Record Examinations National Program for Craduate School Selection, the Aptitude Test and eighteen Advanced Tests are administered. It is planned to obtain Advanced Test norms on first-year graduate students enrolled in the corresponding major fields. Aptitude Test norms will be obtained on the basis of a cross section of all firstayear graduate students. In order to provide Aptitude Test norms it is necessary to obtain score data on the Aptitude Test for students enrolled in other fields as well as for students enrolled in the eighteen major fields. In addition, it is planned to provide Aptitude Test norms on students enrolled in each major field. The major fields to be sampled are listed in the cable in Section I-C.
B. Sampling ratios for the major fields. The number of cases. needed in the sample to represent. the population of each major field adequately is approximately 300. Since more students are enrolled in some fields than in others, it is desirable to-use different sampling ratios for the different major fields to obtain adequate samples. Thus for a field with a large enrollment, such as engineering, a smaller proportion of the total enrollment would need to be selected than for a field with a smaller enrollment. In the case of a field with a very small enrollment every student may have to be selected., To adjust for differences in enrollment in the major fields the sampling ratios given in the table in Section I-C have been chosen.
C. Procedure for 'the selection of an unbiased sample. It is essential that the sampling within each major ríield be as nearly random as possible in order not to bias the norms. The following procedures are designed to produce an unbiased sample.

1. The popilation to be sampled consists of all students entering graduate school in the fall of 1964. Since sampling ratios differ from field to field (see table on following page), the total population of students must be identified according to the eighteen major fields designated in the table. All of the names of the students in major fields not specifically designated in the table should be listed on a combined roster and identified by
major field. Although a single roster is to be used for major fields not listed in the table, the major field with which each student is affiliated should still be indicated.
2. Within each of the eighteen flelds named in the table and within the combined category, each student should be assigred a number from 1 to $N$ (where $N$ equals the total number of students enrolled in the field or in the combined category).
3. In the actual selection of the students for a major field sample, the first student to be selected will be the one whose number corresponds to the initial selection number given in the table. Then the following cases to be selected will be detemined by the sampling ratio for that major field. Of course in the case of a sampling ratio of $1 / 1$, all cases would be included in the sample. For example, if at institution A there were 100 graduate students enrolled in the History on a roster and assigned consecutive numbers. In accordance with the initial selection muber in the table, the first student to be selected would be student No. 2; and in accordance with the sampling ratio for History (1/3), every third student, thereafter, i.e., students numbered 5, 8, 11, etc., would be selected until the entire list of students in the field was exhausted.

Sampling Ratios and Initial Selection Numbers for Major Fields
Major
Field

Biology
Business
Chemistiry
Eoonomics
Education
Engineering
French
Geology
Government
History

| Sampling | Initial |
| :--- | :--- |
| Ratio | Major | $\begin{array}{lll}1 / 5 & 1 & \bullet \\ 1 / 10 & 1 & \\ 1 / 3 & 2 & \\ 1 / 2 & 1 & \\ 1 / 10 & & \\ 1 / 10 & - & \\ 1 / 1 & 1 & \\ 1 / 1 & -1 & \\ 1 / 2 & 1 & \\ 1 / 3 & 2 & \end{array}$


| Literature - | 1/5 | 1 |
| :---: | :---: | :---: |
| Mathematics | 1/3 | 2 |
| Philosophy | 1/1 | 1 |
| Physical |  |  |
| Education | 1/2 | 1 |
| Physics | 1/3 | 2 |
| Psychology | 1/3 | 2 |
| Sociology | 1/1 | 1 |
| Spanish | 1/2 | 1 |
| Major Fields | 1 | , |

D. Preparation and mailing of the rosters. On the enclosed rosters the names of the students selectec for the samples should be listed according to major fields. If $C R E$ scores are readily available on the students, they should be recorded in Columns. $4-6$ on the rosters.

For a student who has one or more missing CRE scores, the name of the undergraduate college he last attended should be given; and his date of birth should be provided if it is readily available. When the rosters are completed, they are to be sent to:

Dr. Philtp R. Harvey<br>Assistant Program Director<br>Graduate Record Examinations<br>Educational Testing Service<br>Princeton, New Jersey

II. Spocial test administration for securing missing data on students in the sample

The studenis in the sample for whom $\operatorname{GRE}$ scores are available will not need to be tested. ETS will search the files for the scores of those students whose scores are not furnished by the cooperating institutions. Students for whom the Advanced Test score and/or the Aptitude Test scores are missing will be requested to take the test(s) for which scores-are not available. ETS will report to each cooperating institution the names of students to be tested, and a test date will be arranged for the institioion. Examination materials and instructions for testing will be provided by ETS.

[^0]Appendix C

WEIGHTS USED IN PREPARING THE FREQUENCY DISTRIBUTIONS

The weight for a college or university i in a given field was corquted as follows:

$$
\begin{equation*}
H_{i}=\frac{N_{i}}{F_{i}} \times \frac{N}{N^{\prime}} \times R \tag{i}
\end{equation*}
$$

where $N_{i}=$ The mumber of full-time first year graduate students at college o: university 1 in the sample;
$F_{i}=$ the mumber in the sample who had the appropriate CRE score;
$N=\Sigma N_{i}$ for all colleges or universities in the same size category as college or university i;
$N^{\prime}=\sum N_{i}\left(F_{i} \neq 0\right)$ for all colleges or universities in the same size category as college or university i;

R a the weight based on the sampling ratio in the appropriate size category ( 1 for large colieges or universivies, 4 for mediumsized colleges or universitiee, and 138/15 for smail colleges or undersities).

To see how this weighting scheme-worked, consider the following example for the field of Biology:

College or University
Size Category ID No.

Large
( $\mathrm{R}=1$ )
21
48
24
Medium-Sized ( $\mathrm{R}=\mathrm{L}$ )

Small
(R-738/15)

31
32
45
ID No.
1 45 48 49 60

| $N$ | $F$ | $H$ |
| ---: | :---: | :---: |
| 21 | 0 |  |
| 48 | 4 | -2 |
| 24 | 2 | 15.5 |
| 21 | 3 | 15.5 |
| 0 | 0 | 28.0 |
| 19 | 1 | 76.0 |
| 18 | 4 | 41.24 |
| 8 | 6 | 12.27 |
| 9 | 1 | 82.8 |

6) 

## APPENDICES - Page 28

In this example the weights for colleges or universities 3,31 , and 49 in Biology were computed as follows:

$$
\begin{aligned}
& \text { No. } 3: \frac{48}{4} \times \frac{93}{72} \times 1=15.5 \\
& \text { No. } 31: \frac{21}{3} \times \frac{40}{40} \times 4=28.0 \\
& \text { No. } 49: \frac{8}{6} \times \frac{35}{35} \times \frac{138}{15}=12.27
\end{aligned}
$$

For both the sptitude Test and the Advanced Tests, the irgquencies were weighted according to the weights caiculated by the formula given previouisly. Howover, for the Aptitude the froquencies were combined across fields to give ovorali distributions ion Verbal sbility and guanititative Ability as well as distributions for each of the major Eieldis.

SPECIFIC FIELDS CLASSIFIED UNDER PARTICULAR MAJOR FIELDS FOR SAMPLING PURPOSES



[^0]:    か.

