



UNITED STATES GENERAL ACCOUNTING OFFICE

WASHINGTON, D.C. 20548

PROGRAM EVALUATION
AND
METHODOLOGY DIVISION

SEPTEMBER 10, 1985

B-220175

The Honorable James J. Howard
House of Representatives

Dear Mr. Howard:

Subject: Review of Two Studies on College Athlete
Graduation Rates (GAO/PEMD-85-9)

This report responds to your June 26, 1985, request for a GAO review of two studies on college athlete graduation rates. You asked us to "determine whether or not the methodology used was accurate and whether or not these studies are relevant" in discussing your proposed legislation H.R. 2620. Your bill would require colleges to graduate at least 75 percent of their scholarship athletes (defined as students who receive athletic scholarships for at least 3 academic years) within 5 years in order for contributions to their athletic departments to qualify as tax deductible.

The two studies we reviewed are (1) "National Collegiate Athletic Association Survey of Graduation Rates After Five Years for Males First Entering College in Fall 1975," prepared by the American College Testing Program in April 1981 (we refer to this study as "ACT"), and (2) "Study of Freshman Eligibility Standards Technical Report," prepared by Advanced Technology, Inc., in August 1984 (we refer to this study as "AD TECH"). ACT concluded that male athletes graduated at a rate equal to or higher than that of male nonathletes. AD TECH concluded that student athletes graduated at about the same rate as students in general. Our review of these two studies was based on professionally accepted methodological standards for behavioral research and, in the interest of timeliness, limited to information provided in the two written reports.

We found that both studies have limited relevance for your legislation. H.R. 2620 is directed at athletes who remain at the same institution and receive athletic scholarships for 3 years. Both studies reported graduation rates for athletes and nonathletes only for freshman college entrants rather than for those who had been enrolled for 3 years. In addition, ACT did not report analyses on athletes by scholarship status. The AD TECH researchers conducted analyses by grant-in-aid (scholarship)

(973599)

033184

status although they included some athletes at institutions that do not offer athletic scholarships.

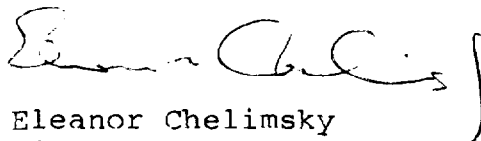
We found serious methodological problems in each study. First, generalizability: the characteristics of the samples signify that generalizations can be drawn from each study only to the schools that participated in that study. Second, data quality: the poor quality of the data in ACT limits the confidence that can be placed in the study's results and the precision with which the results can be interpreted. Third, data analysis: inappropriate data analyses in both studies mean that the reported findings cannot be used to compare the graduation rates of athletes and nonathletes. More appropriate and meaningful analyses could be performed on the data collected for both ACT and AD TECH, provided that these data have been retained. We believe the AD TECH data are of sufficiently high quality to allow confidence in the results of such analyses.

More specific details are provided in the enclosures to this report. Our review of ACT is in enclosure I, of AD TECH in enclosure II. Enclosure III contains a hypothetical illustration of what we mean by "appropriate" and "inappropriate" data analyses.

The problems we found are such that the two studies are inconclusive for the purpose of answering your questions. AD TECH was designed primarily to assess the effects of instituting eligibility standards for participation in athletics in the freshman year, and we did not review its relevance or quality for this purpose. As the two studies stand, however, they do not provide a basis for concluding that the graduation rates of college athletes are higher or lower than, or the same as, those of all other college students.

As requested by your staff, we did not send a draft of this report for advance review to the American College Testing Program, Advanced Technology, Inc., and the National Collegiate Athletic Association, in the interest of timeliness. No further distribution of this report will be made for 2 days from the date of this letter. At that time, we will send copies to these three organizations and will make additional copies available to others who are interested.

Sincerely yours,



Eleanor Chelimsky
Director

REVIEW OF ACT

1. The characteristics of the sample limit the conclusions that can be drawn from this study to men at 46 member institutions of the National Collegiate Athletic Association (NCAA).
 - * ACT appropriately chose to take a random sample of athletes and nonathletes to allow generalization of the results to all NCAA institutions. The researchers sent questionnaires on student graduation rates to a random sample of 200 member institutions of NCAA. However, only 46 institutions responded. This response rate, 23 percent, is low by professional standards.
 - * ACT asked responding institutions to report graduation rates for men only.
 - * ACT did not establish that the 46 responding institutions were typical of NCAA member institutions in general. The low response rate upset the initial design, which called for a random sample for drawing general conclusions about NCAA member institutions. ACT did not report standard analyses to demonstrate that participation in the study was unbiased. Thus, the institutions that participated may differ from other institutions on factors such as institutional size, prestige, or NCAA division, which could affect reported outcomes, or they may differ from nonparticipants directly in graduating a higher or lower percentage of athletes.
2. The quality of the data in this study is low, by professional standards, for placing confidence in comparisons of athletes and nonathletes.
 - * The data may contain reporting bias. The respondents for each institution surveyed were largely directors of athletic departments who, it could be argued, may have been predisposed to report more favorable information for athletes than nonathletes. Although there is no proof of bias, ACT did not report any efforts to minimize its likelihood, such as independent verification of reported data. Thus, differences between athlete and nonathlete graduation rates could in fact be less favorable to athletes than those reported.
 - * Some of the students classified as nonathletes may participate in sports as frequently as some of the students classified as athletes. The survey instructions did not clarify how status as an athlete was to be defined. The definition was left to the discretion of the respondents, who may have used different criteria for classifying athletes and nonathletes. In addition, classification was based on status upon college entrance.

Some students may have changed status during the first year or may have been misclassified at the outset of the study. ACT did not report information on the reliability or stability of classifications. Therefore, confidence in the reported contrasts between athletes and nonathletes is less than it might have been.

3. Inappropriate data analyses in this study limit conclusions.

- * The overall rates of graduation of athletes and nonathletes reported in ACT do not reflect the typical pattern in the 46 institutions surveyed. The overall rates reported in this study are biased toward the rates of graduation at the largest institutions in the sample. ACT did not provide information on the range or distribution of rates across institutions and did not report analyses to explain the large differences between overall and median rates.

The data suggest that a few of the larger institutions in the survey had relatively high rates of graduation for all students (athletes and nonathletes) compared to the other institutions surveyed and that the rates for athletes were exceptionally high at a few institutions. Therefore, the overall rates may seem higher than actual rates in most schools for both athletes and nonathletes.

- * The median rates of graduation for athletes and nonathletes reported in ACT do not reflect the typical pattern of differences in the 46 institutions surveyed. (The median is the point at which half the schools graduated students at higher rates and half at lower rates.)

There is no problem with using median rates to describe the central tendency of a single group or to compare rates in two groups when the groups are not matched on a relevant characteristic. There is a problem with using two separate group medians to compare athletes and nonathletes matched by institution of attendance. In ACT, the median rates indicate only that 23 of the institutions graduated nonathletes at rates lower than 33.8 percent and 23 graduated athletes at rates lower than 36.9 percent. These are not necessarily the same 23 institutions. An appropriate procedure is first to calculate the difference in the rates of graduation at each institution and then to find the median of this difference.

- * ACT did not report appropriate analyses for establishing the typical pattern of difference between athlete and non-athlete graduation rates within the 46 institutions surveyed. The average difference in graduation rates for athletes and nonathletes in these 46 institutions may be zero or it may favor athletes or nonathletes. There is no way to determine this from the analyses that were reported.

REVIEW OF AD TECH

1. The characteristics of the sample limit the conclusions that can be drawn from this study to black and white athletes at 206 Division I member institutions of NCAA.
 - * AD TECH decided to survey all 276 NCAA Division I member institutions, which is one appropriate way to draw adequate conclusions about differences in graduation rates for this type of institution. The response rate, 206 of 276, or 75 percent, is reasonable by professional standards.
 - * However, the participating institutions are not fully representative of NCAA Division I member institutions. The participating institutions had larger fall 1982 enrollments, more were public than private, and more offered a doctorate or first professional degree than Division I institutions as a whole. Rates of participation by NCAA subdivision and NCAA district varied. Differences between participating and nonparticipating institutions on any of these factors could imply differences in graduation rates as well.
 - * AD TECH reported results only for black and white athletes.
 - * The intended sample of grant-in-aid athletes included an unknown number of athletes who did not receive grants-in-aid. An unreported number of participating institutions did not provide athletic grants-in-aid. These institutions provided information on a sample of their athletes. AD TECH included these data in the overall analyses of the performance of grant-in-aid athletes. Therefore, confidence in drawing conclusions from the study about grant-in-aid athletes is less than it might have been.
2. The quality of the reported data in this study appears to be high.
 - * Data collection was coordinated by institutional liaisons designated by each institution's chief executive officer, and in some institutions as many as five departments participated in data-reporting. Answering the survey required access to formal student records for information on athletes.
 - * The definition of "athlete" for most institutions was a student receiving an athletic grant-in-aid. For military academies and other institutions that did not award athletic grants-in-aid, "athlete" was defined as a student participating in NCAA-recognized sports.
 - * Quality-control procedures were used to minimize errors in preparing the data for analysis.

3. No appropriate analysis was reported to support a contrast between athletes and students in general on graduation rates.

- * How the graduation rates of athletes compared to those of nonathletes in the 206 participating institutions is unknown. AD TECH reported overall graduation rates for athletes and median institutional graduation rates for students in general. These two statistics cannot be meaningfully compared.

HOW STATISTICS CAN BE MISLEADING

We constructed hypothetical data to illustrate how the use of different summary statistics can lead to different conclusions. In this example, four hypothetical schools A-D are equal in size and School E has a much larger student body and a larger number of athletes. School E is the only school that graduates athletes at rates greater than it graduates nonathletes. Tables 1 and 2 show these hypothetical data. Table 3 shows appropriate and inappropriate analyses and explains why the ACT and AD TECH analyses were inappropriate.

Table 1Data on Graduation Rates of Individual Students
at Five Hypothetical Schools

<u>School</u>	<u>No. of nonathletes</u>		<u>No. of athletes</u>	
	<u>Students</u>	<u>Graduates</u>	<u>Students</u>	<u>Graduates</u>
A	500	400	50	20
B	500	400	50	20
C	500	150	50	0
D	500	150	50	10
E	<u>2,000</u>	<u>500</u>	<u>200</u>	<u>150</u>
Total	4,000	1,600	400	200
% graduating (overall graduation rate)	1,600/4,000 = 40%		200/400 = 50%	

Group with higher rate: Athletes

Table 2Data on Graduation Rates of Nonathletes and Athletes
in Five Hypothetical Schools

<u>School</u>	<u>Graduation rate</u>		<u>Difference in rate</u>	<u>Group with higher rate</u>
	<u>Nonathletes</u>	<u>Athletes</u>		
A	80%	40%	40	Nonathletes
B	80	40	40	Nonathletes
C	30	0	30	Nonathletes
D	30	20	10	Nonathletes
E	25	75	-50	Athletes
Number of schools	5	5	5	
Graduation rate				
Mean	49%	35%	14%	
Median	30	40	30 ^a	

^aThe median difference is computed vertically, not horizontally.

Table 3Possible Comparisons from the Hypothetical Data

<u>Statistics for 2 groups from the same 5 schools</u>	<u>Athletes</u>	<u>Nonathletes</u>	<u>Group with higher rate</u>
Graduation rate			
Median	40%	30%	Athletes
Overall	50	40	Athletes
Mean	35	49	Nonathletes
Difference in graduation rate			
Mean		14%	Nonathletes
Median		30	Nonathletes
Percentage of schools graduating students at higher rates	20%	80%	Nonathletes

Median rates are appropriate summary statistics for one group or for two independent groups. In this example, the two groups are not independent because the samples of athletes and nonathletes were drawn from the same schools. In this kind of situation, it is inappropriate to compare two medians. Such a comparison provides misleading information about the typical case.

Overall rates are appropriate for answering questions about aggregate results (all athletes and nonathletes). They are inappropriate for answering questions about the typical case, because they ignore the information that students come from different schools. Thus, overall rates are biased toward the rates at institutions with the largest number of students. In this example, the difference between overall rates is largely a consequence of the unique pattern at School E.

Mean rates are appropriate for answering questions about the typical case. They can be misleading when interpreted alone, because they are sensitive to extreme values. Thus, it is inappropriate to draw inferences from means in the absence of other information.

The mean difference in graduation rate is algebraically equivalent to the difference in mean rates and is therefore interpreted in the manner described above for mean rates.

The median difference in rates is always appropriate for answering questions about the typical case. It is not sensitive to extreme values or institutional size. The median difference is not equivalent to the difference in medians.

The percentage of schools showing a particular outcome is also an appropriate statistic. It provides information not on the size of the typical difference between two groups but only on the pattern of difference across schools.

Finally, it is always inappropriate to compare two different statistics to draw inferences about group differences. Each statistic is calculated in a different way and each has a different meaning.

AD TECH compared two different statistics--median institutional graduation rates of students in general and overall graduation rates of athletes.

ACT compared median institutional rates for athletes and nonathletes as well as overall rates for both groups.