DOCUMENT RESUME

ED 439 646 HE 032 720

AUTHOR	Wood, Anda L.; Ridley, Dennis R.; Summerville, Richard M.
TITLE	University Grade Inflation through Twenty Years: An Analytical and Empirical Investigation. ASHE Annual Meeting Paper.
PUB DATE	1999-11-18
NOTE	54p.; Paper presented at the Annual Meeting of the Association for the Study of Higher Education (24th, San Antonio, TX, November 28-21, 1999).
AVAILABLE FROM	Director of Institutional Research and Planning, Virginia Wesleyan College, 1584 Wesleyan Drive, Norfolk/Virginia Beach, VA 23502-5599. E-mail: dridley@cnu.edu.
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Academic Achievement; *Academic Standards; Evaluation Methods; Evaluation Research; *Grade Inflation; Grade Point Average; *Grades (Scholastic); Grading; Higher Education; Scores; *Student Evaluation
IDENTIFIERS	*ASHE Annual Meeting

ABSTRACT

This paper presents a new method to test for grade inflation. Defining grade inflation as the decreasing value of grades in the coin of student achievement, the study avoids the assumption that a rise in mean grade is a necessary condition of grade inflation. The study, which was conducted at a single university, also sought to determine whether grade inflation could be demonstrated if mean grades have leveled off. To measure general academic abilities, Graduate Record Examination (GRE) data was used. Data was grouped into three time spans (1976-80, 1984-88, and 1992-96). To examine the influence of students' majors, data were also separated into two broad groups, "humanistic" and "scientific," because literature had shown roughly that the former revealed more grade inflation than the latter. Findings suggested grade inflation relative to GRE total and verbal scores, but no evidence of grade inflation when GRE quantitative or GRE analytical scores were used, and it is postulated that this is a result of declining academic standards in the sense that verbal reasoning ability associated with grade point average is lower than it used to be. Differences were also found between disciplines, with some evidence of grade inflation found for science majors. (Contains 17 references.) (RH)

University Grade Inflation Through Twenty Years: An Analytical and Empirical Investigation[1]

Anda L. Wood

JSI FundRaising Systems, Inc.

Dennis R. Ridley

Virginia Wesleyan College

Richard M. Summerville

Christopher Newport University

Anda L. Wood is Conversion Programmer with JSI FundRaising Systems, Inc., of Williamsburg, Virginia. She is a 1997 graduate Christopher Newport University, in Newport News, Virginia with a B.S. in Mathematics.

Dennis R. Ridley is the Director of Institutional Research and Planning, Virginia Wesleyan College, Norfolk/Virginia Beach, Virginia.

Richard M. Summerville is a Professor in the Department of Mathematics at Christopher Newport University.

Paper presented at the 1999 Annual Meeting of the Association for the Study of Higher Education, November 18, 1999.

ASHE *Association for the Study of Higher Education*

Headquartered at the University of Missouri-Columbia • College of Education • Department of Educational Leadership and Policy Analysis • 211 Hill Hall, Columbia, MO 65211 • 573-882-9645 • fax 573-884-5714

This paper was presented at the annual meeting of the Association for the Study of Higher Education held in San Antonio, Texas, November 18-21, 1999. This paper was reviewed by ASHE and was judged to be of high quality and of interest to others concerned with higher education. It has therefore been selected to be included in the ERIC collection of ASHE conference papers.

ABSTRACT

This article presents a new method to test for grade inflation. Defining grade inflation in economic terms, i.e., "decreasing value of grades in the coin of student achievement," the study develops a method for analysis of relevant data spanning twenty years. As the definition avoids assuming that mean grade rise is a necessary condition of grade inflation, so also the study addresses whether grade inflation can be demonstrated when mean grades had leveled off. This approach required a valid measure of significant general academic abilities, providing covariates to assess important changes in the value of grades. The available Graduate Record Examinations (GRE) data were used for three time spans, 1976-80, 1984-88, and 1992-96, for students who took the GREs as seniors. The exercise produced a wealth of findings that are highly relevant to the problem of identifying grade inflation and declining academic standards.

INTRODUCTION

Concerns over grade inflation in higher education go to the heart of the issue of academic standards and whether they are on the rise or falling. When analyzed, the concept of grade inflation reveals a strong connection with this issue of standards. Bejar and Blew (1981) provided a useful definition of grade inflation as, "...an increase in grade point average without a concomitant increase in achievement." In a moment we address whether this definition can be improved upon and what difference such a refinement might make. However, Bejar and Blew's definition seems correct, at least, in making a connection with academic standards and concerns over those standards. If average grades are on the rise at the same time that the correlated measures of students' achievement apart from grades—i.e., data revealing what students know and can do—are not keeping pace, that situation signals declining academic standards. In the Doonesbury comic strip, incompetent students have been portrayed assailing their professors for awarding them any grade lower than an "A." Hu and Kuh (1998) cite evidence that many professors experience these pressures; others point out that such pressures are brought home through the power exerted by students via student evaluations (Greenwald, 1996, 1997; Trout, 1997). As portrayed in Doonesbury, when they are pressed by irate students on one side and compromising administrators on the other, these thoroughly cowed professors give in and award higher grades than the students deserve, thereby lowering the average performance expected to receive that high grade in the future. After traveling that route, a university's academic standards will have disappeared, even while the traditional token of academic achievement, the aggregate GPA, increases.

Because such a scenario has been known or suspected in many quarters for a long time,

the appearance or reappearance of an upward trend in average GPAs amounts, for many academics, to an automatic call to action. Such a response is particularly expected if university administration (as in the Doonesbury episodes) reinforces lowering academic standards (Rotfeld, 1997, 1998). However, there is at least one serious difficulty with using average grades as an automatic barometer of declining academic standards. As this article attempts to show, the difficulty can arise from failure to take fully into account all of the essential parts of grade inflation. Bejar and Blew's definition, "...an increase in grade point average without a concomitant increase in achievement," reveals grade inflation to be an economic concept, fundamentally concerned with the changing value of one measure, grades, in terms of something else, student achievement. As an economic concept, it does not permit one to infer a change in the value of grades when the behavior of only one key variable (grades) is known and the second variable (student achievement), essential to defining the value of the first, is not known. If student achievement rises concomitantly with grades, the value of grades is unchanged. Therefore, from rising grades alone, although one may suspect, one can not infer grade inflation or the undermining of academic standards.

Further, the problem of identifying whether true grade inflation exists is even more fundamental than finding out whether a rise in grades accompanies a decline in independently measured achievement. This point arises as soon as one asks whether grade inflation can exist even when average grade trends have leveled off. Such a leveling off in recent years described the situation at the institution where the current study took place. While fewer concerns might be voiced, the possibility of grade inflation can not be dismissed. Embedded in an economic concept of grade inflation is the implication that rising grade trends are *not* a necessary condition for grade

inflation. We have already concluded that such a trend is not a *sufficient* condition. We now assert that Bejar and Blew's definition is not completely correct in that a rising aggregate GPA is not even a *necessary* condition. A corrected definition might state: "decreasing value of grades in the coin of students achievement." From this correction, it is clear that grade inflation is theoretically consistent with a leveling or declining state in average grades.

To make this fundamental point, however, is to raise an equally fundamental practical consideration. If we suspect that grade inflation, and slippage of academic standards, may be a problem, how can we test our suspicions in a particular institutional context? This article reviews and critiques an earlier attempt to accomplish that goal, and finally proposes a new methodology that is particularly suited to situations in which the aggregate GPA is not rising. The article then goes on to apply this method to an institutional context, and examines the findings that emerge from that application.

Our article further develops and illustrates the consequences of such a definition for the analysis of one institution's grade data. Before presenting the immediate background of this study, however, it is instructive to contrast the above points with another significant recent approach to the problem of grade inflation (Hu & Kuh, 1998). After reviewing evidence of rising grades coupled with diminished academic effort, Hu and Kuh comment: "...grade inflation may be at work, a discriminatory practice whereby earlier cohorts of students received lower grades for academic effort comparable to their counterparts in the 1990s. That is, grades are inflated due to an artificial increase independent of academic effort or student characteristics (e.g., ability, motivation)" (Hu & Kuh, 1998). This definition and the study based upon it recognized a central role of student effort, in particular as measured by various scales of the College Student

Experiences Questionnaire (Pace, 1998), the data source for their large sample of national data. The article also alludes to the role of measured academic ability in grade inflation and acknowledges the non-inclusion of such data as a limitation of the approach. Despite this limitation, Hu and Kuh's article is helpful in analyzing the upward trend in college grades and in suggesting a limited role of grade inflation, as they define it, in such a trend. Their thoughtful discussion lies outside the scope of this article except as regards their claim to have isolated grade inflation. One problem, which they acknowledge in their discussion of limitations, is that the variables they focus upon, including student academic effort and student background characteristics, typically account for only a small amount (10-15%) of the variance in students' grades. This problem is exacerbated by their reliance upon students' self-estimates of their GPAs, a criterion which itself fails to capture approximately 30% of the variance of actual grades. However, in our perspective the non-inclusion of developed academic ability measurement is the most serious drawback as regards grade inflation. While student academic effort is an important consideration, our approach stays with the Bejar and Blew (1981) definition (as modified above) in emphasizing students' measurable academic competence as even more critical in grade inflation. The connection between grade inflation and academic standards, in particular, turns upon that emphasis. Educators are most animated by the concern that students lack academic competence; student effort, while important, does not guarantee competence as an outcome. Finally, Hu and Kuh appear to accept the common assumption called into question above that a rising grade trend forms one necessary condition of grade inflation. To return to the fundamental point that grade inflation does not depend on an average increase in grades, this insight raises an equally fundamental practical consideration regarding how to test for grade inflation.

The immediate background for the study reported here was a study of average grade data at one institution and comparative nation-wide data (Summerville, et al, 1990). This study identified stable disciplinary differences, suggesting grade inflation within some disciplines. Average grades earned by students within these disciplines were systematically higher than the grades they earned in courses within other disciplines taken at the same time. This result suggested grade inflation dividing along disciplinary lines, helping to identify average grade rise in at least some sectors of the curriculum. Additional research at this institution found evidence revealing the effects of different grading standards in different disciplines (Ridley, Quanty & Sciabica, 1998). However, the limitations of the Summerville *et al.* study have become more clear in retrospect. Again, grades alone can not establish a change through time in the relationship between grades and academic achievement. While grades themselves are indicators of achievement, it is important to have valid and credible measurements of academic achievement, which are independent of the grading process. Such measurement seems required by the concept of grade inflation as defined above. Further, such an approach is required to account for ways in which the role of disciplines may have changed across time. Thus, a new approach is required to test for grade inflation in the situation where the rate of increase in average grades has finally leveled off.

Again, to conduct such a test required substantial data from a widely recognized valid measure of developed academic ability or achievement. For this study the Graduate Record Examination (GRE) was chosen as the operational measure. Roughly a more difficult version of the SAT, this measure provides a generalized measure of developed academic abilities with acceptable reliability and validity coefficients (Anastasi, 1976). Like the SAT, this test provides

separate verbal and quantitative scores along with a score on analytical ability. Also like the SAT, the meaning of scores may have changed through time with population changes and evolution in the relevant disciplines. However, as a source of relevant and valid long-term information, there is nothing comparable to this test. Other considerations influenced this choice. First, the database of GRE scores reported to the institution on graduates and current students considering graduate school, was fairly extensive going back to the 1970s. Of course, these data were limited to those who aspired to graduate school and therefore comprised a sample that was not necessarily representative of the institution's students. However, by observation the ranges of achievement and grade data did not appear restricted and the dispersion was similar to that of unselected seniors; further, the mean GPA of the GRE examinees was only very slightly higher than those of other seniors. Thus, exercising due caution, the investigators opted to study the GRE database together with examinees' grades. The GRE data offered a potentially enlightening source to test for local grade inflation, where grade inflation depended on changes over time in grades as judged commensurately with the GREs. At the least, the results could reveal changes through time in patterns of "over-aspiration" or "under-aspiration" to graduate school, where "aspiration" (in the special sense used here) hinged on how well the several cohorts' average grades have kept pace with their average GRE scores.[2]

In addition, the influence of students' majors was open to study. The influence of majors or disciplines on grade inflation was revealed in the literature (Ekstrom & Villegas, 1992; McKenzie & Tullock, 1981; Ridley, Quanty & Sciabica, 1998; Summerville, Ridley, & Maris, 1990). Stone (1995) found evidence that fields showing the greatest grade rise from about 1965 to 1980 also showed the greatest declines in achievement measured by GRE scores. These fields included

political science, sociology, psychology, education, history, and English literature. On the other hand, student achievement was more consistent across time in biology, chemistry, physics, mathematics, and economics, despite lower average grades than in the first set of disciplines. Thus, the past literature suggested that students' majors would have a potent influence on the relationship between average grades and achievement over time. The literature showed roughly that "humanistic" as opposed to "scientific" majors would reveal more grade inflation as shown by declining achievement in relation to average grades across time. While the cited references relied upon rising mean grades as a key element in the evidence of grade inflation and falling standards, the current study tightens the argument by allowing grades to vary in either direction or remain level.

In all of this work, the authors have been mindful of the intrinsic complexity of the grading process and, in particular, of the diverse professional philosophies and disciplinary and institutional cultures which condition this process. Simultaneously, however, the authors are also mindful of the reality that out of this complex environment is precipitated information about student achievement (GPAs, etc.) which institutions and their constituents use freely and often with little regard for the complexity of its origins. The questions addressed in the present paper are therefore properly viewed as an effort to gain insight into just one dimension of a very large and complex phenomenon--and not in any sense as a denial of the complexity of the overall environment of which it is a part.

The authors considered two primary questions. (1) Will there be evidence of a decline in real grades after adjustment for GRE scores? [As used here, the term "real grades" refers to the correction made by the analysis; it is relative to the covariate used and is not an invariant

quantity.] Such a result might also suggest greater "over-aspiration" to graduate school among more recent graduates. (2) What are the discipline-related trends, as indicated by mean grades in relation to GREs?

METHOD

The body of electronically filed data went back to the 1970s. The investigators separated it into three groups with an equal number of years in the span and in the skipped periods, i.e., 1976 through 1980, 1984 through 1988, and 1992 through 1996. Each period yielded a randomly selected sample except the 1976-1980 group where the data were sparse. Each subject must have graduated around the same time when he or she took the GRE and have earned at least one hundred credit hours at the time of testing. Such degree progress placed him or her in the senior status. After obtaining each person's major from either the registration database or from the Graduate Record Examinations data, the researchers separated the subjects into two broad disciplinary groups labeled for convenience by the authors as "humanistic" and "scientific." Humanistic majors included education, psychology, history, sociology, philosophy and English; the science group comprised mathematics, physics, biology, chemistry, and economics.

The analysis plan relied upon the use of a multiple factor (two by three) analysis of covariance for two factors: discipline (two levels) and times (three levels) as defined above. Students' GPAs and several GRE scores provided the variate and covariates, respectively. Total scores of the GREs were used initially, with further identical analyses using the various sub-scales of the GRE as the covariate. This procedure allowed the investigators to explore in greater depth the meaning of the findings. Unequal cell frequencies required an unequal Ns analysis, for which

Winer (1971, 792-96) provided an unweighted means solution.

Additionally, several linear regression analyses provided more insight into how the variate and covariates were related within the various cells of the design. [Due to substantial unexplained variance, appropriate caution was exercised with regard to the linear regressions. Thus, the method was used only for confirming findings from the other analyses.] For example, grades were regressed upon total GRE scores for students within each cohort, 1976-80, 1984-88, and 1992-96. Separate regressions were done for examinees with majors classified as scientific and humanistic, respectively.

It was specifically predicted that there would be a significant effect of the second factor (time) in the analysis of covariance. This would suggest a decline through time in corrected or adjusted mean grades, i.e., grades adjusted for differences in GRE test scores. Secondly, an interaction effect was predicted, showing that the efficacy of time also depended upon the discipline variable.

RESULTS

Before presenting results, it is important to emphasize the general aim of the analyses. In the broadest terms, there was a two-fold purpose: first, to test whether there was evidence of grade inflation at one university across a 20-year time span; and second, to use this study to illustrate a method whereby the same questions might be investigated elsewhere. The two parts comprised the local and the general applications of the method. The second part, although more important in the long run, depended on the first. In that sense, the first part is chiefly illustrative; in any case, the findings from the first part, based upon one university, have limited

generalizability. Having made that obligatory caveat, however, the local study will prove to be an intriguing source of hypotheses and thus will fulfill a purpose beyond that of illustrating a method.

One distinguishing characteristic of these analyses must be kept in mind: that is, they were designed to permit us to apply a new definition of grade inflation: "decreasing value of grades in the coin of student achievement." This definition is free from the erroneous assumption that grade inflation must depend upon average grade rise. However, because the analyses attempt to measure change in grades relative to student achievement, a potentially confusing use of words can enter in, namely, the "change in real grades." By using this term, we do not mean to suggest that an invariant measure had been discovered; rather, we imply that we are measuring change in terms of a variable that has been corrected, or re-calibrated, by reference to its relationship to a measure of student achievement. This usage is similar to reference to change in "real dollars" after the value has been corrected for purchasing power.

The last point leads directly into the organization of the results into four parts. Since student achievement is complex and may be measured in various ways, real grades depend upon what measure is being used. The following analyses begin with the broadest context: grades and GRE total scores. The three subsequent parts seek further clarification by focusing on the several parts of the GRE total score—the quantitative, verbal, and analytical scales. These analyses remind us that, because student achievement is complex, the application of our definition of grade inflation must reflect that complexity.

Finally, cutting across all the analyses is the issue of whether the results depend upon the discipline or group of disciplines represented in the data—i.e., humanistic versus scientific disciplines. The authors find the outcome of this inquiry to be quite surprising and suggestive.

The first analysis examined grades and the GRE total score. Both main effects (discipline and time) were highly significant (Table 2). In addition, the interaction was a marginal result, with an F-ratio just under that required for significance at the conventional alpha level of .05. These results showed the scientific group of majors to have the highest corrected grades, despite their lower actual mean GPAs. In addition, the effect of time was toward a decreasing corrected GPA over time. The marginal interaction suggested that the advantage of science students in real grades diminished with time.

Insert Tables 1 and 2 about here

The first linear regression analysis regressed grades upon total GRE scores for each of the three cohorts, 1976-80, 1984-88, and 1992-96, and for each of the two discipline groups within these periods (see Figures 1.A., 1.B., and 1.C.). The results were consistent with those reported above. This analysis showed that the science group had a greater superiority to the humanities group in the lower GPA ranges. This difference tended to diminish notably in the later period. The science group in the latest period showed a slight score advantage at the lower GPAs and almost none at the higher GPAs.

Insert Figure 1 about here

The next analyses attempted to find out how the various parts of the GRE total score contributed to the general result (Table 2). First the researchers looked at the quantitative scale

of the GRE as the covariate. In that analysis there was a highly significant effect of discipline, but no significant effect due to time. Scientific students' adjusted grades were much higher than humanities students' grades when corrected for the effects of the quantitative scale, despite their lower mean GPAs. There was also a significant interaction effect. This showed that, while there was no general effect due to time, the discipline difference depended on time. That is, the discipline difference in real grades (favoring sciences when grades were adjusted for quantitative scores) has diminished.

The linear regression results for the quantitative GRE score followed a similar pattern (see Figures 2.A., 2.B., and 2.C.). The science group had a clear advantage on this score at all grade levels in the first period. This advantage diminished greatly by the middle period and virtually disappeared in the latest.

Insert Figure 2 about here

The verbal scale of the GRE was next entered as the covariate (Table 2). The second variable (time) again was significant, while the first (discipline) was just under the level required for significance at the .05 level. There was no significant interaction effect. Again, there was a real grade decline when the verbal scale was the covariate. Marginally, there was a difference in adjusted grades favoring sciences.

Again, the linear regression results for the verbal GRE score followed a similar pattern (see Figures 3.A., 3.B., and 3.C.). There was a general decline in the height of the regression line, showing generally higher verbal scores at all grade levels in the first period. The first period result

suggested some superiority of the science group in scores for the lower grade levels. The second and third periods show greater similarity between the two discipline groups.

Insert Figure 3 about here

Finally, the analytical scale of the GRE entered as the covariate (Table 2). This analysis found a highly significant effect due to discipline. There were no other significant effects. Science students' corrected GPAs were higher after adjustment for differences in the analytical scale.

The linear regression results for grades regressed upon the analytical scores gave similar results. There was a difference due to discipline, favoring the sciences, clearly present in the first period and smaller in the second. This difference disappeared in the latest period at the higher grades.

Insert Figure 4 about here

DISCUSSION

This study has illustrated a new methodology to test for the existence of grade inflation, defined as "decreasing value of grades in the coin of student achievement." We have started from the premise that grade inflation does not necessarily depend upon an upward trend in average grades. Based upon that insight we have adopted methods to fit the new definition, and we have

applied them to one institutional context. This application can not yet address the larger issue of grade inflation in U.S. higher education (Hu & Kuh, 1998); however, it is conceivable that our methodology could be applied directly to that problem and thus put the issue upon a sounder theoretical basis than we believe has been available heretofore. Before recommending the adoption of these methods on a larger scale, however, we need to review the success of the local application.

The first question that we will discuss is whether the study provided sound, credible evidence of grade inflation at one university. We suggest that the available evidence is consistent with that interpretation, although there may be alternative explanations. We return to the point made earlier that changes in *real* grades depend upon what measure is being used to define achievement. The application of our methods to grade inflation is complex because student achievement is complex and may be measured in different ways. Accordingly, we found evidence consistent with grade inflation over a 20-year period relative to the GRE Total Score and GRE-Verbal, but we failed to find evidence of grade inflation over all students when the GRE-Quantitative or GRE-Analytical were used to define achievement. Since the Total Score is essentially a composite, it is reasonable to assume that the decline in grades relative to the GRE Total Score was largely due to the Verbal component of the Total Score. Therefore, the important contrast is that between GRE-Verbal, on one side, and both the GRE-Quantitative and GRE-Analytical on the other. The major conclusion appears to be that to the degree grades reflect and depend upon performance measuring verbal ability (briefly, "the ability to reason with words in solving problems"—Graduate Record Examinations Board, 1997), the evidence suggests that *to that degree* grade inflation has occurred.

This conclusion, of course, is consistent with other literature (Stone 1995; Wingspread Group 1993). However, our study provides a reasonable attempt to make that point without relying upon average grade increase as a key element of the argument. In sum, our findings suggest declining academic standards in the sense that the verbal reasoning ability recently associated with a certain GPA is lower than it used to be.

As stated before, other interpretations are possible. One such plausible alternative interpretation is that aspirations to attend graduate school have risen relative to *real* grades, particularly when the latter are evaluated in terms of verbal reasoning ability. The extraordinary expansion of graduate school opportunities lends credence to that view. As these opportunities have increased, aspirations to pursue further education have risen among students in general, encouraging greater numbers of those who tend to score low on a challenging test of verbal reasoning ability to participate in the process by sitting for the GRE. Thus, the results suggesting possible grade inflation may reflect a decreasing incidence of self-selection out of that testing situation. However, it is also possible that both interpretations bear part of the truth. That is, during the same time span, academic standards have lowered and aspirations have increased.

A second issue concerns the broad disciplinary clusters of science and humanities. What differences emerged between these two clusters relative to the GRE and GPA analyses? How did such differences contribute to the previously discussed indications of grade inflation? The categorization (following Stone) into scientific and humanistic majors or disciplines, although rough, was fruitful of several provocative results that bear upon these questions. It is noteworthy that students in science-related majors, while having generally lower GPAs, had higher grades when evaluated in GRE terms. However, the generally higher value of science students' grades

did not decline over time for all GRE indicators but only for the Total Score and the Quantitative indicator. It appears that students in the two discipline clusters have become more similar to each other over time in the association between their quantitative abilities and GPAs. There may be a flattening of students' differences between these two discipline clusters with respect to *real* grades evaluated in terms of quantitative ability. Looked at another way, this finding modifies the earlier conclusion that we failed to find evidence of grade inflation over all students when the GRE-Quantitative indicator was used. Rather, we found some evidence of grade inflation, not over all students, but only for students in science majors.

The preceding statement contrasts with Stone's conclusion that grade inflation was greatest in the majors that we have called humanistic. What might this difference suggest? First, a discipline-related trend in a large body of data might be contradicted by local exceptions. However, we propose another intriguing possibility. Stone's observation was based upon data of 1965-1980; therefore, it may be valid for that period only and need revision. The much greater diversity of today's college students might be relevant to that revision. Weak quantitative skills may have served in the past to screen out more students from pursuing majors that placed a premium upon those skills. A more diverse group aspiring to careers requiring such majors might have imposed pressures toward grade inflation, in terms of quantitative ability, within that particular discipline cluster. We propose this hypothesis as worthy of further confirmation study, perhaps applying a similar methodology upon a broader basis than is reported here.

Also for further study, we suggest other institutions utilize the method used in our study to analyze grade inflation in terms of measured achievement. Theoretically, this method could apply just as well to high school research using the SATs and students' grades.

In summary, we have reviewed the application of a new methodology to the study of long-term grade inflation within one institution of higher education. We submit that the exercise has produced a wealth of findings that appear highly relevant to the problem of identifying grade inflation and declining academic standards. For reasons presented earlier in the article, we believe this methodology incorporates a sounder theoretical basis than available heretofore, one that captures the economic nature of the concept and incorporates valid measures of academic achievement. Further, we submit that the method yields a nuanced, multidimensional view of grades interacting with academic abilities through time. The results appear reconcilable with prior research while opening up intriguing possibilities that go well beyond that background.

NOTES

1. Requests for reprints should be sent to the second author. Address: Director of Institutional Research and Planning, Virginia Wesleyan College, 1584 Wesleyan Drive, Norfolk/Virginia Beach, Virginia 23502-5599. E-mail: dridley@cnu.edu. Readers may address any additional requests for information to the first author. Address: JSI FundRaising Systems, Inc., 4732 Longhill Road, Suite 2201, Williamsburg, VA 23188. E-mail: awood @jsifrs.com. The authors gratefully acknowledge the suggestions made by Drs. Robert C. Birney, Paula Szulc Dominguez, Robert F. Grose, and Clinton B. Walker.

2. The point made here does not require that students' "over-aspiration", if indicated by the results, necessarily implies that their plans were unrealistic. Graduate school admissions standards in programs serving the students may have changed in a complementary direction, making these plans realistic. Rather, the terms "over-aspiration" and "under-aspiration" indicate change in mean test results through time relative to grade correlates, regardless of admissions criteria. These two terms are intended merely to describe such relationships and do not imply any judgment on the part of the authors as to which students should or should not aspire to graduate education.

REFERENCES

Anastasi, A. (1976). *Psychological testing*. Fourth Ed. New York: Macmillan Publishing Company.

Bejar, I. I. & Blew, E. O.. (1981). Grade inflation and the validity of the scholastic aptitude test. *American Educational Research Journal, 18*, 143-56.

Ekstrom, R. B. & Villegas, A. M. (1992). College grades: An exploratory study of policies and practices. Final Report. Princeton, N.J.: Educational Testing Service.

Graduate Record Examinations Board. (1997). *Graduate Record Examinations 1997-98 information and registration bulletin.* Princeton, N.J.: Educational Testing Service.

Greenwald, A. G. (1996). *Applying social psychology to reveal a major (but correctable) flaw in student evaluations of teaching*. ERIC Document Reproduction Service No. ED 400 754.

Greenwald, A. G. (1997). *Do easy teachers get high ratings?* Peregrine Publishers. Wakefield, Massachusetts.

Hu, S. & Kuh, G. (1998, November). *Unraveling the complexities of grade inflation between the mid-1980s and mid-1990s.* Paper presented at the annual meeting of the Association for the Study of Higher Education, Miami, Florida.

McKenzie, R. B.,& Tullock, G. (1981). *The new world economics.* Homewood, Illinois: Richard D. Erwin.

Pace, C. R. (1988). *Measuring the quality of college student experiences.* Rev. Ed. Los Angeles: University of California, Higher Education Research Institute.

Ridley , D. R., Quanty, M. B. & Sciabica, M. (1998, November). *Grading standards and course challenge: An analytical-empirical approach*. Paper presented at the annual meeting of

the Association for the Study of Higher Education, Miami, Florida.

Rotfeld, H. (1997). Of grades and standards. *Academy of Marketing Science News, 18*, 9.

Rotfeld, H. (1998). Living grades to avoid grief sells students short. *Marketing News, 2*, 11.

Stone, J. E. (1995). Inflated grades, inflated enrollment, and inflated budgets: An analysis and call for review at the state level. *Education Policy Analysis Archives, 3.*, 11.

Summerville, R. M., Ridley, D. R. & Maris, T. L. (1990). Grade inflation in the eighties: The case of urban colleges and universities. *College Teaching, 38*, 33-38.

Trout, P. A. (1997) What the numbers mean: Providing a context for numerical student evaluations of courses. *Change, 29* (5), 24-30.

Winer, B. (1971). *Statistical principles in experimental design*. (2nd Ed.). New York: McGraw-Hill Book Company.

Wingspread Group on Higher Education. (1993). *An American imperative: Higher expectations for higher education, an open letter to those concerned about the American future.* Racine, Wisconsin: The Johnson Foundation.

TABLE 1.--Descriptive Statistics on GPA and GRE Scores for Time Periods and Disciplines

		GRE Score			
Time/Group/Ns	GPA	Total	Quantitative	Verbal	Analytical
1976-80					
Humanities (N = 20)					
Mn	3.06	1364.0	418.0	482.5	463.5
SD	0.46	246.8	96.0	99.8	103.8
Sciences (N = 12)					
Mn	2.90	1709.2	610.0	536.7	562.5
SD	0.31	239.0	86.2	99.0	87.5
1984-88					
Humanities (N = 34)					
Mn	3.28	1407.3	445.9	479.1	482.4
SD	0.39	218.4	73.2	104.7	89.2
Sciences (N = 21)					
Mn	3.04	1541.4	556.7	466.7	518.1
SD	0.40	255.0	109.6	95.2	124.4
1992-96					
Humanities (N = 30)					
Mn	3.19	1433.7	467.7	471.0	495.0
SD	0.44	277.4	88.8	94.8	126.0
Sciences (N = 30)					
Mn	3.14	1544.7	535.3	456.0	553.3
SD	0.47	239.6	99.1	98.0	87.2

TABLE 2.--Results of Analyses of Covariance: Effects of Disciplines and Time Periods on GPA Adjusted for GRE Covariates

	Covariates (GRE):			
Variable:	Total	Quantitative	Verbal	Analytical
Discipline (A)				
F (1,141)	32.00	62.00	3.74	20.10
p	<.001	<.001	<.06	<.001
Time (B)				
F (2,141)	3.89	1.80	7.66	1.47
p	<.01	n.s.	<.01	n.s.
Interaction (A x B)				
F (2,141)	3.02	3.96	2.21	0.83
p	<.06	<.01	n.s.	n.s.

FIGURE 1: GRE TOTAL SCORES VS. GPA

FIGURE 1-A: GRE TOTAL VS GPA (76-80)

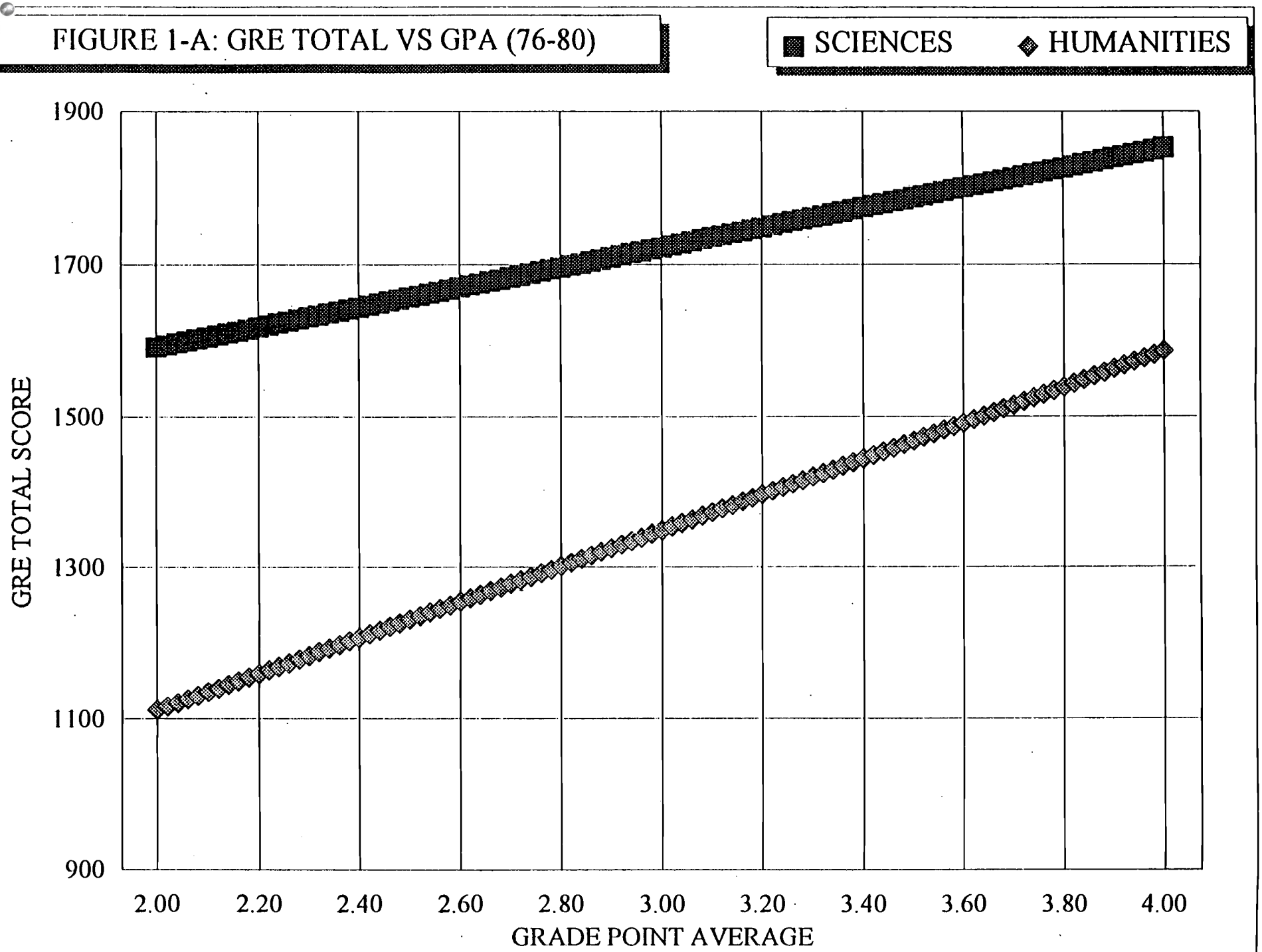


FIGURE 1-B: GRE TOTAL VS GPA (84-88)

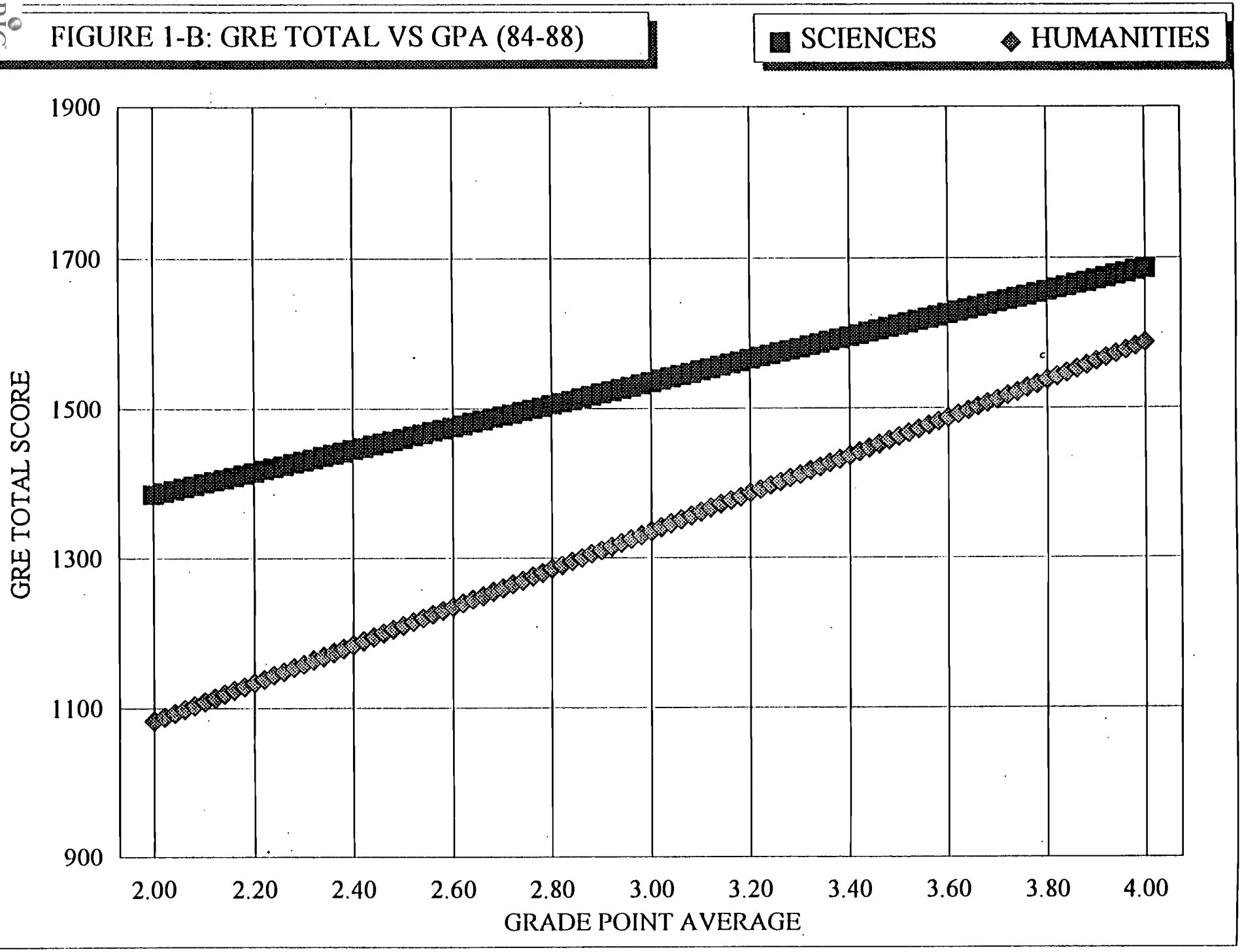


FIGURE 1-C: GRE TOTAL VS GPA (92-96)

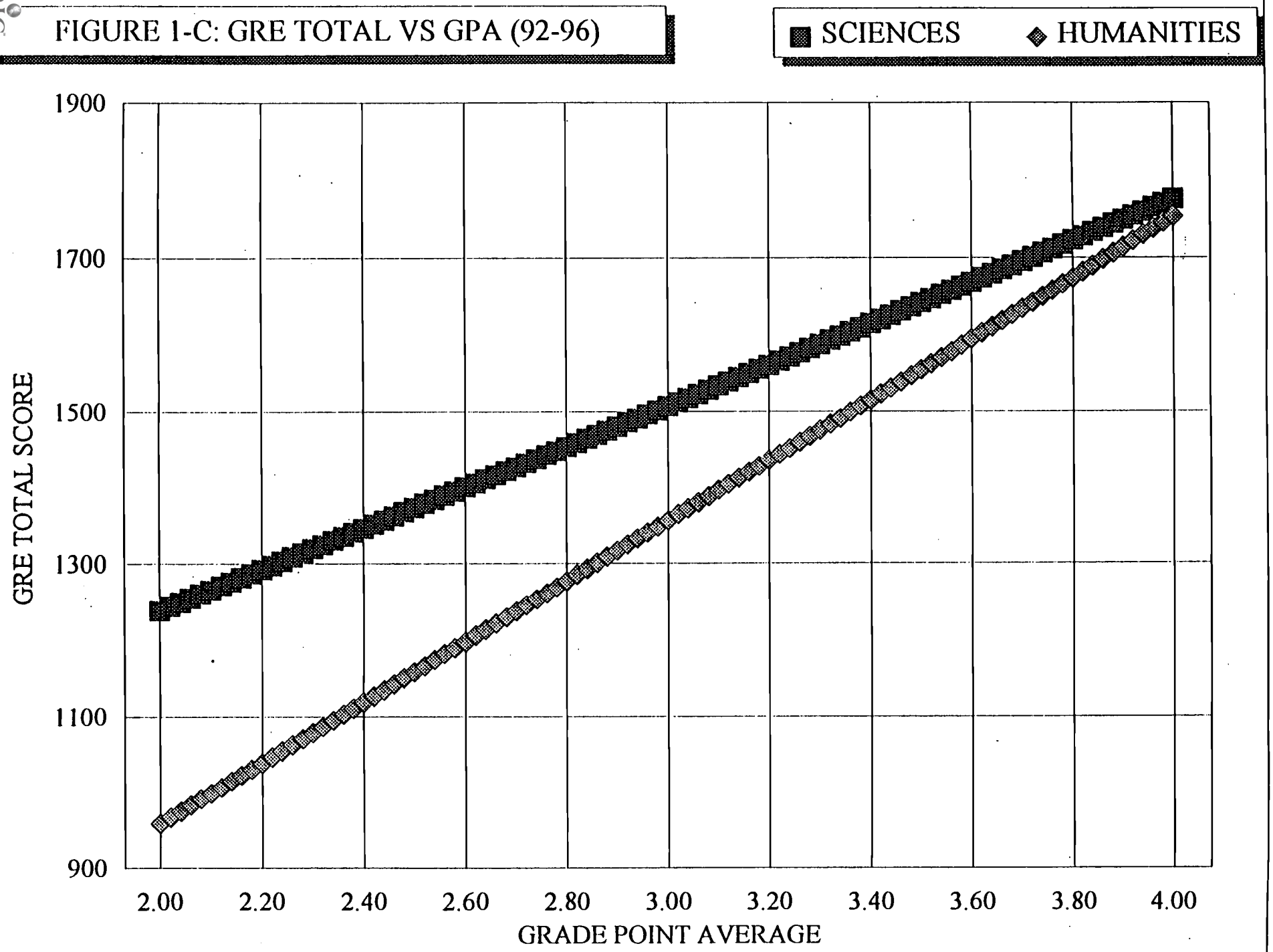


FIGURE 2: GRE QUANTITATIVE SCORES VS. GPA

FIGURE 2-A: GRE QUANT VS GPA (76-80)

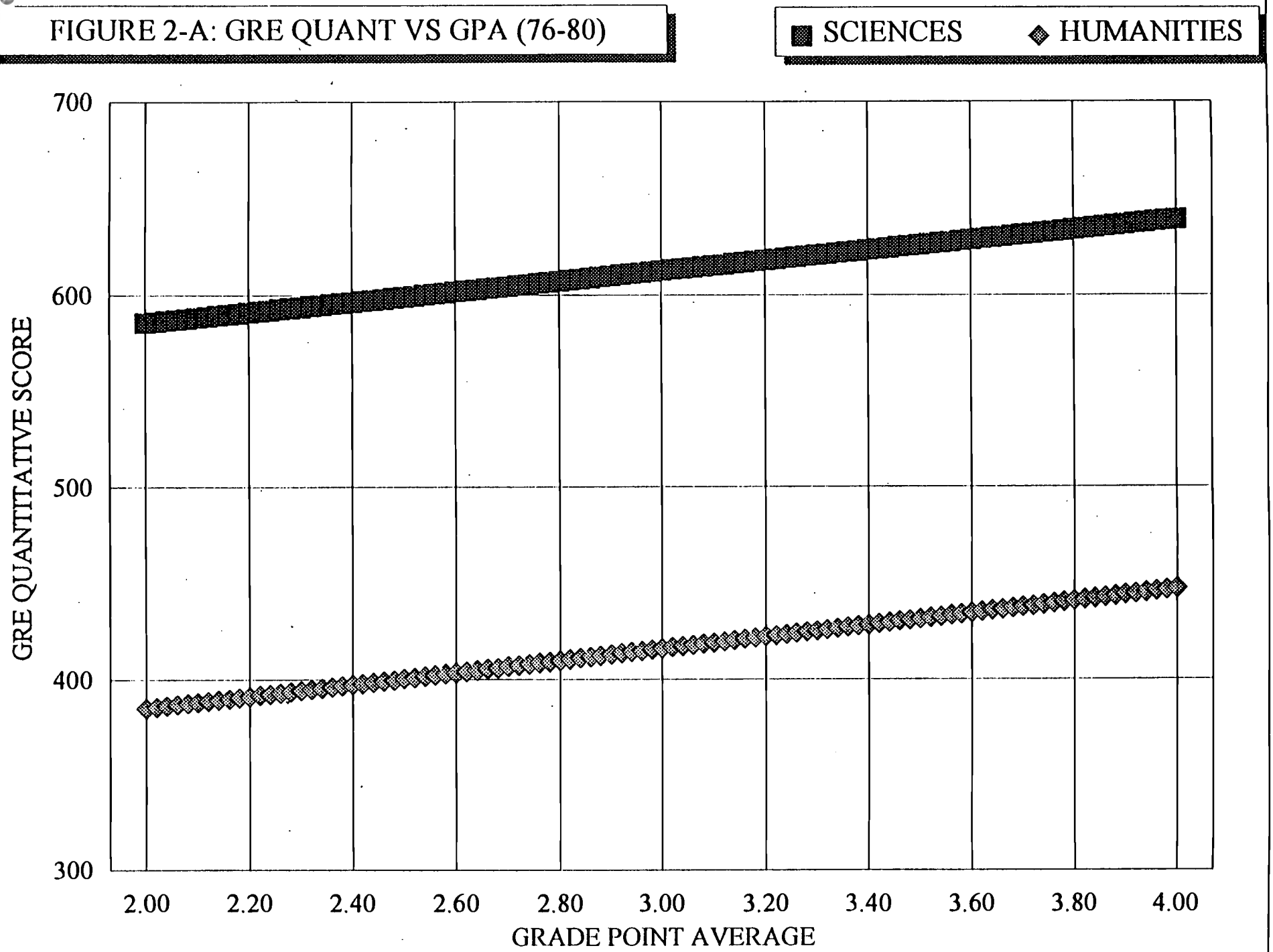


FIGURE 2-B: GRE QUANT VS GPA (84-88)

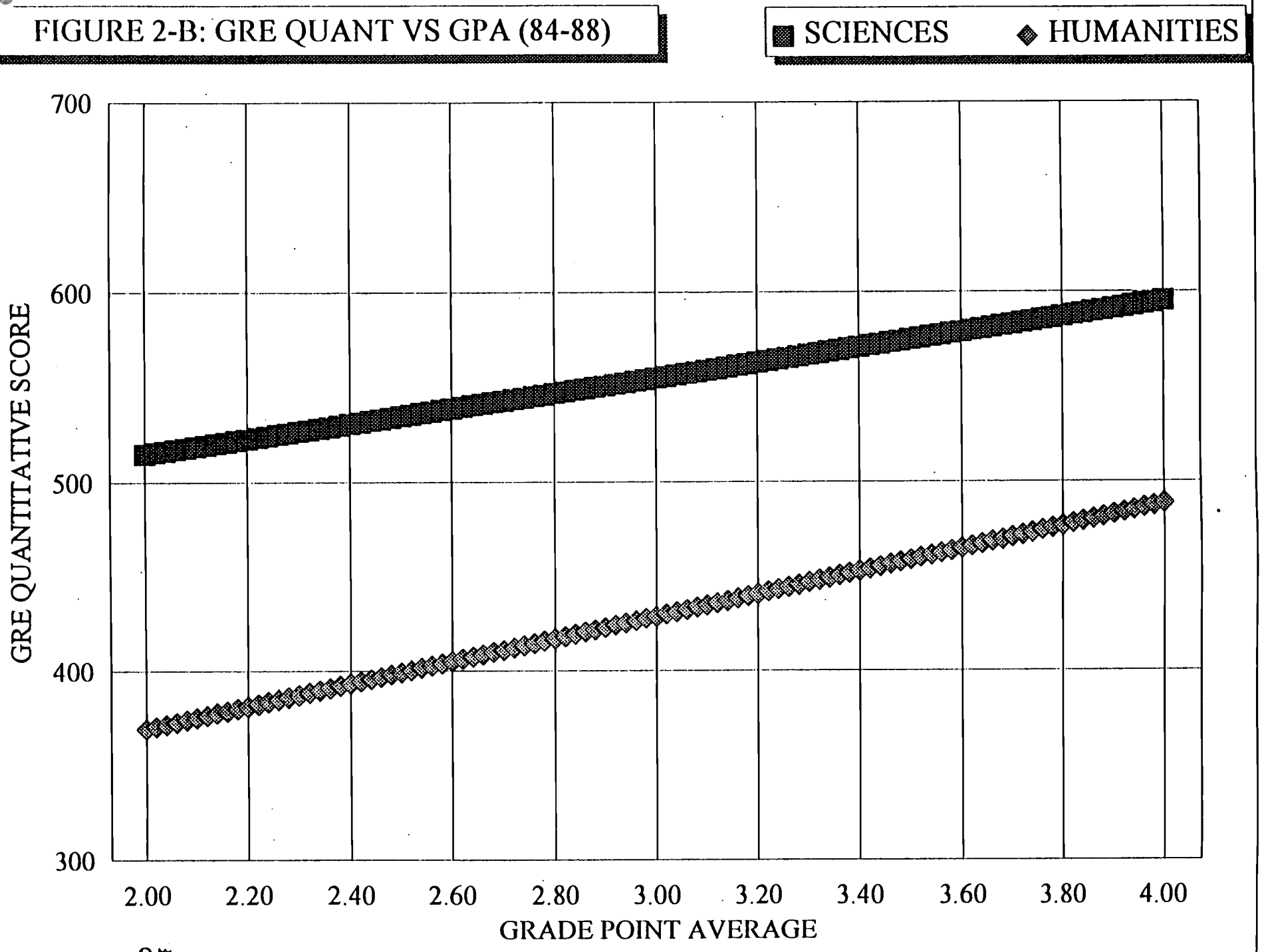


FIGURE 2-C: GRE QUANT VS GPA (92-96)

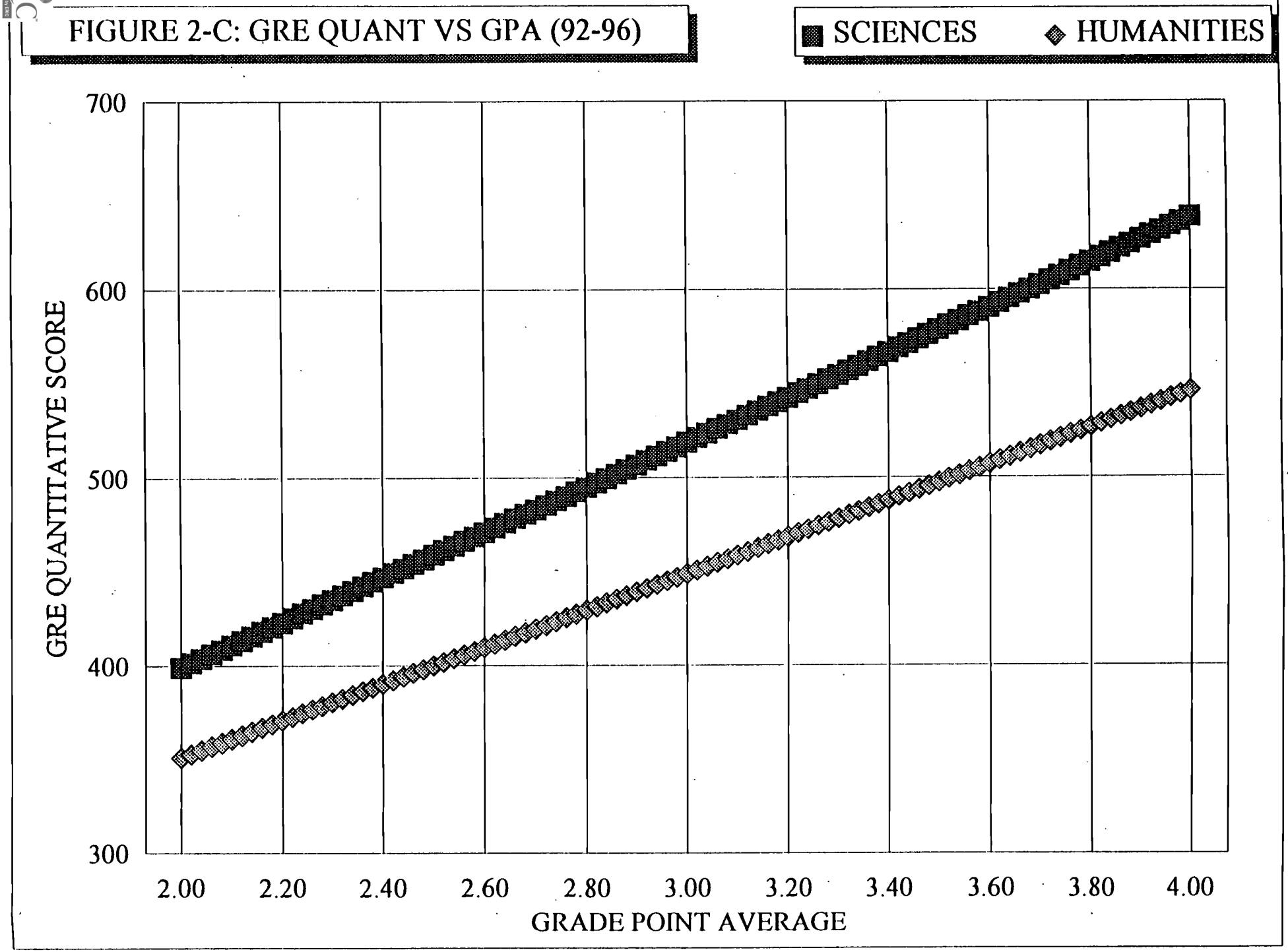


FIGURE 3: GRE VERBAL SCORES VS. GPA

FIGURE 3-A: GRE VERBAL VS GPA (76-80)

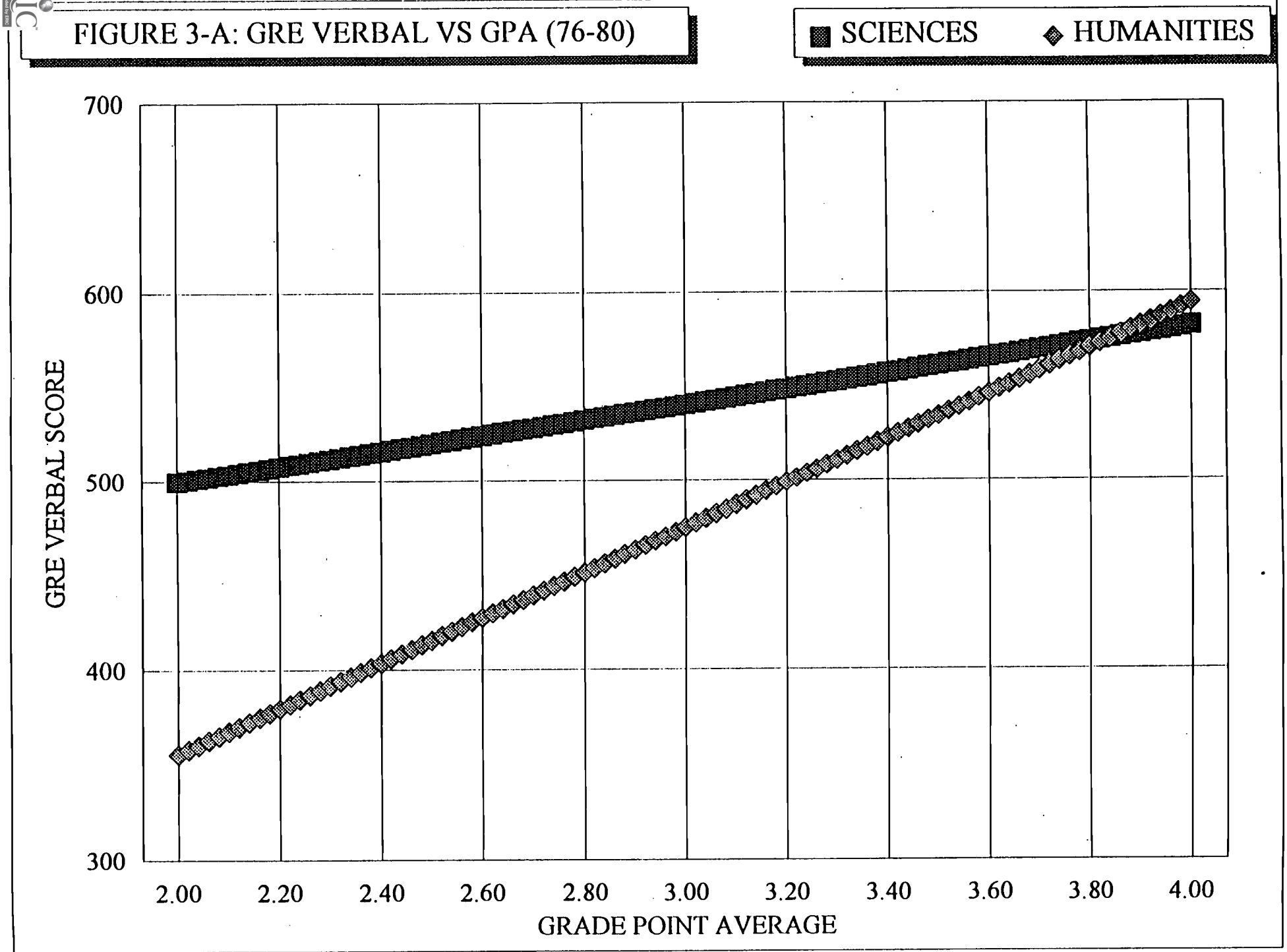


FIGURE 3-B: GRE VERBAL VS GPA (84-88)

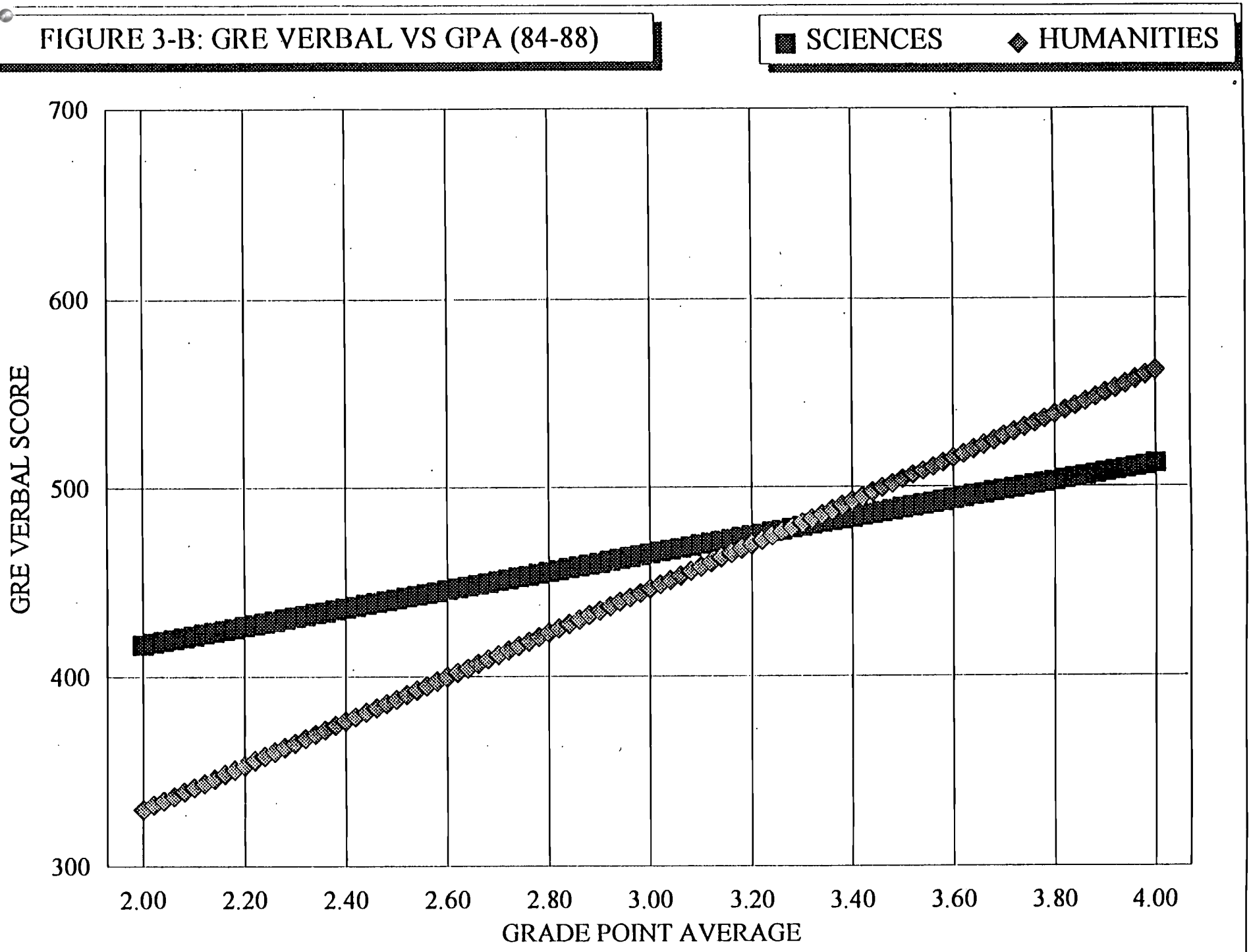


FIGURE 3-C: GRE VERBAL VS GPA (92-96)

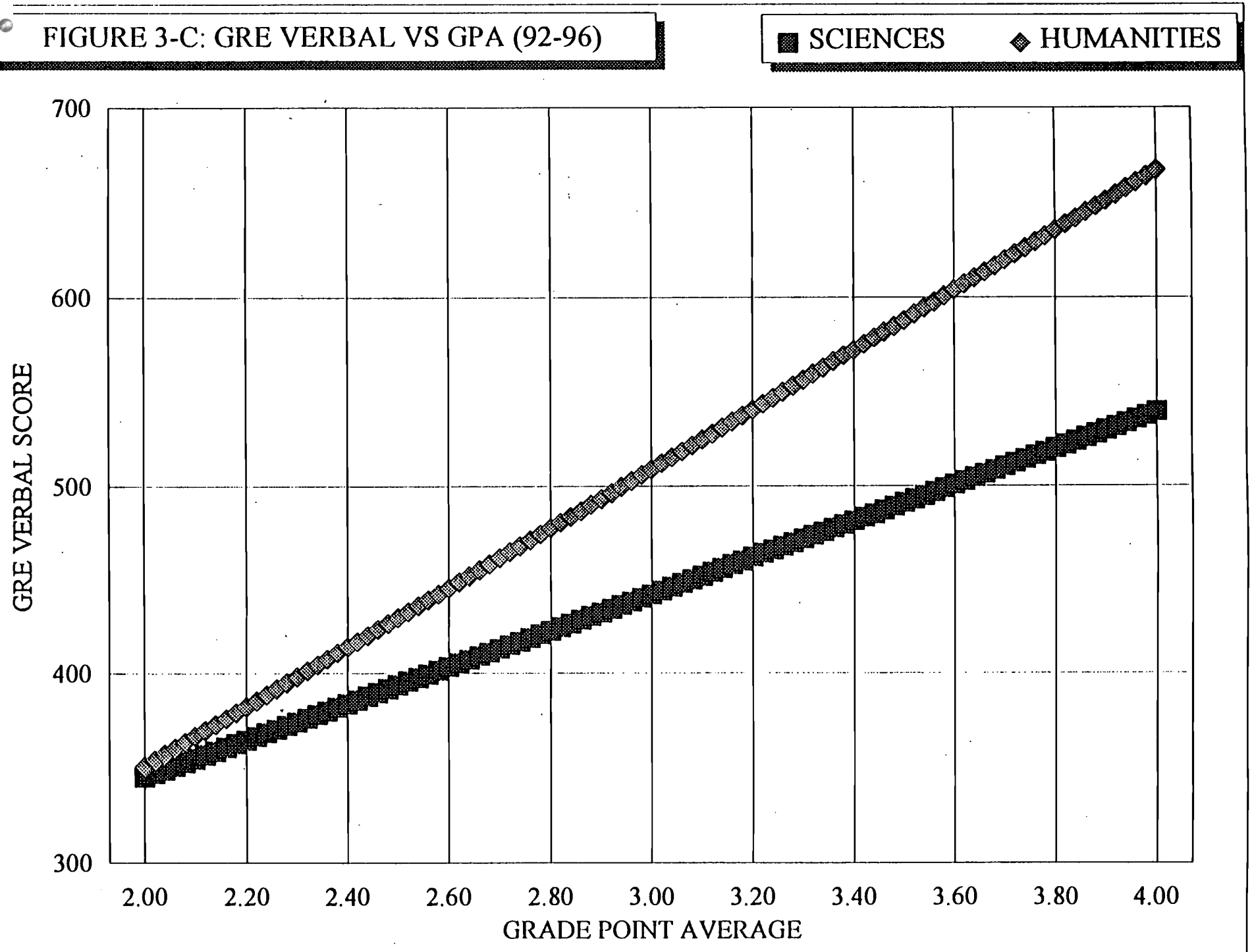


FIGURE 4: GRE ANALYTIC SCORES VS. GPA

FIGURE 4-A: GRE ANALYTIC VS GPA (76-80)

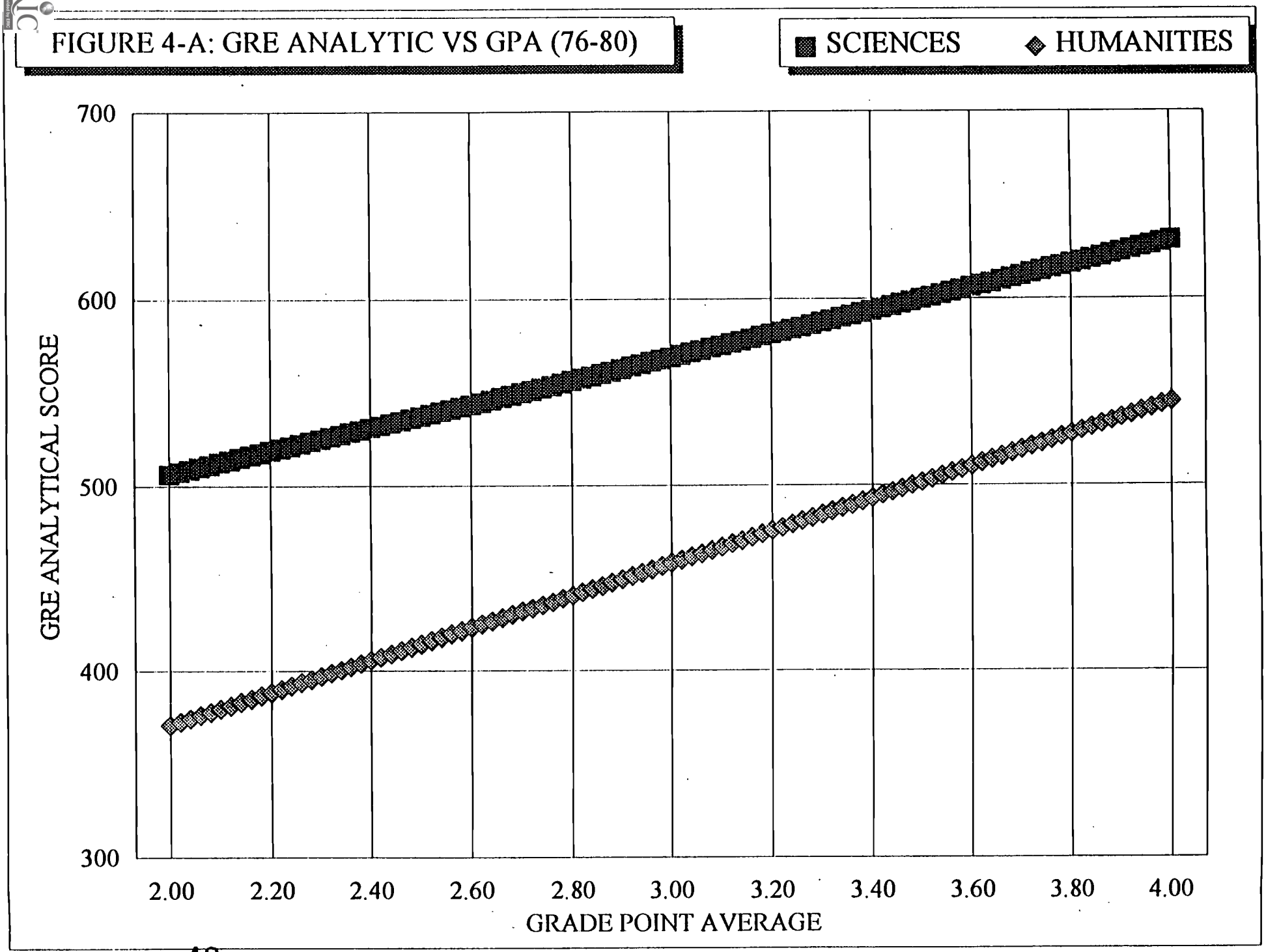


FIGURE 4-B: GRE ANALYTIC VS GPA (84-88)

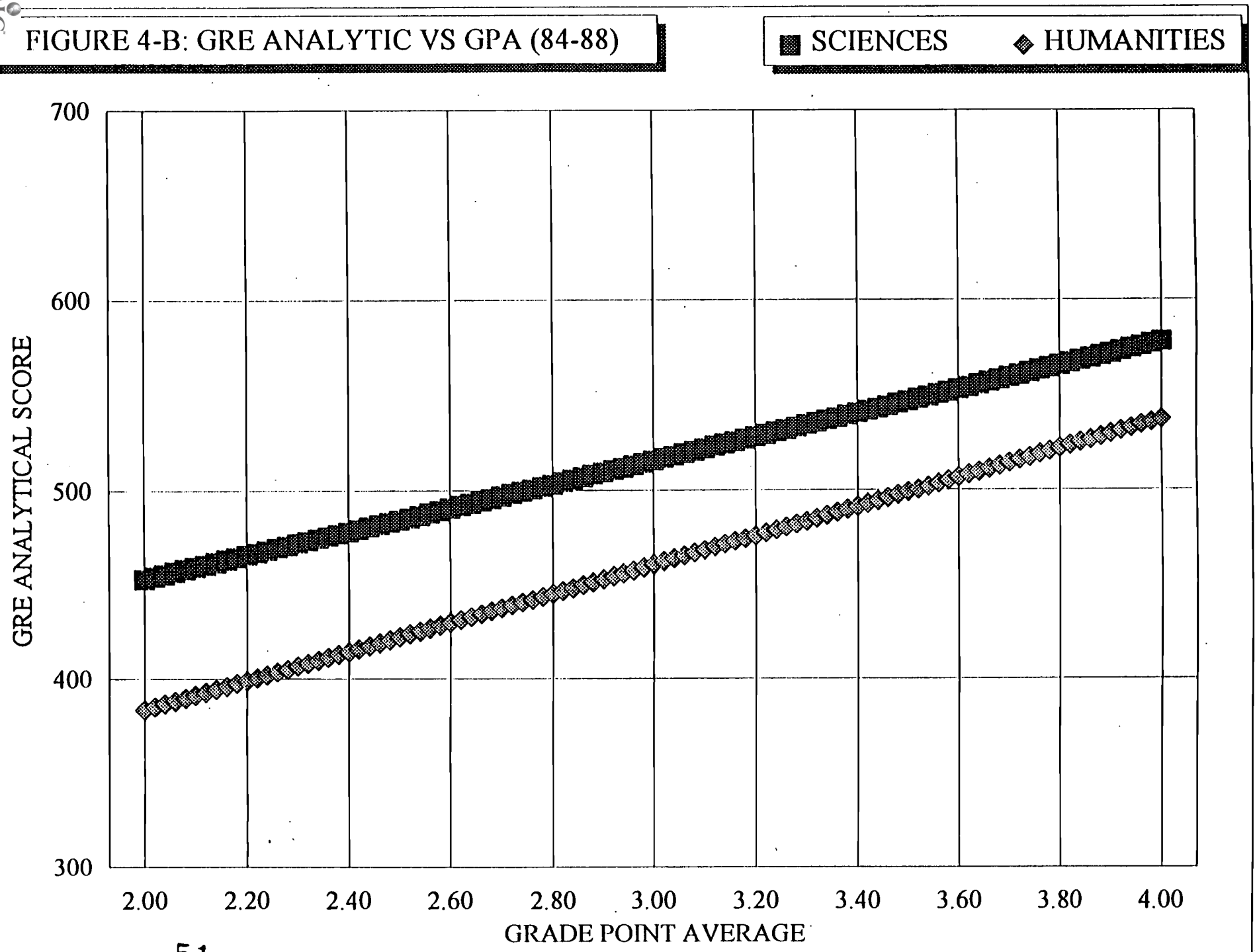
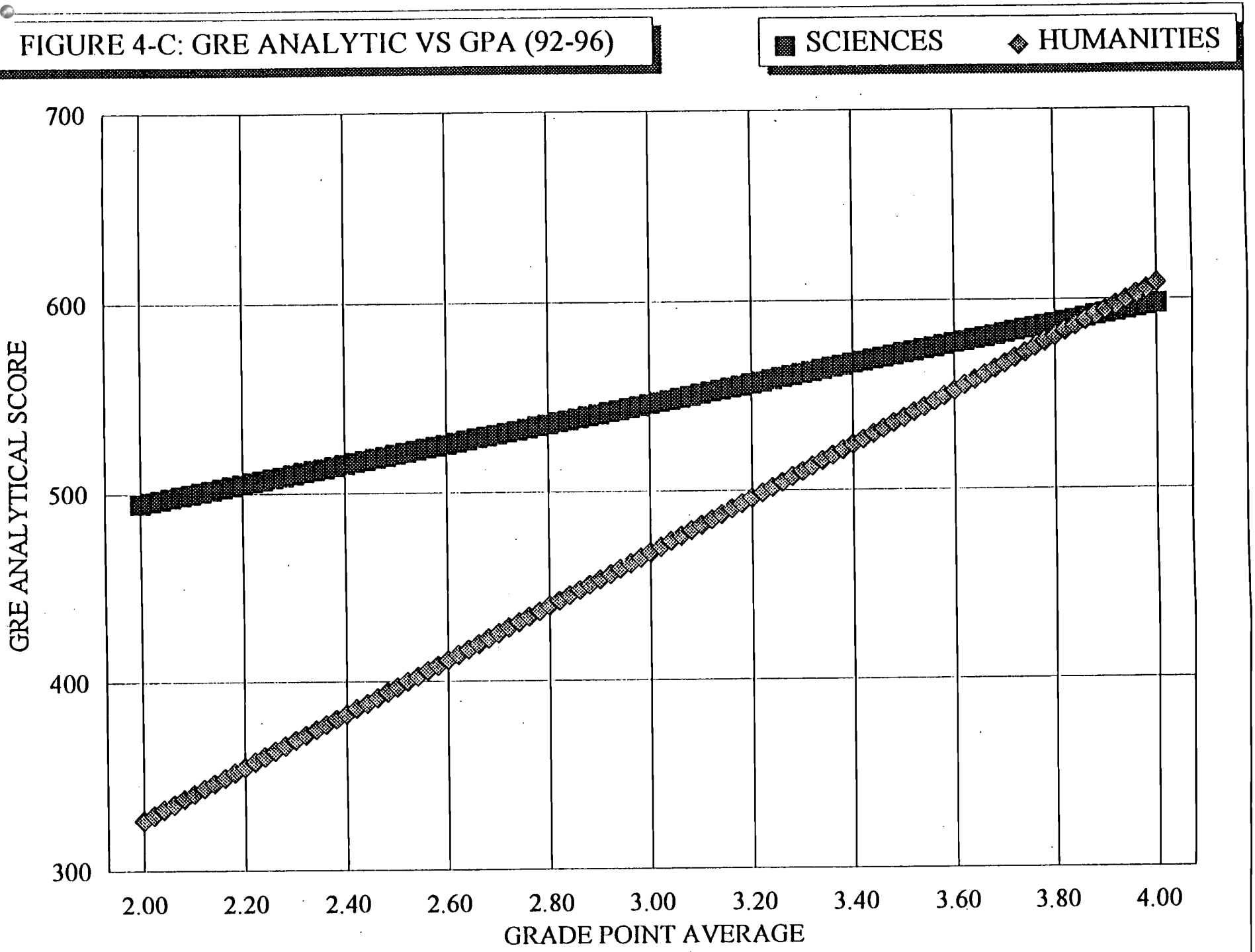


FIGURE 4-C: GRE ANALYTIC VS GPA (92-96)



U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

☒ This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").