Quality Education: Hanged without a Trial

Before the article by Astin on "Undergraduate achievement and institutional 'excellence'" (16 Aug., p. 661) becomes the Project Hindsight of the educational world, it may be well to bring out for public discussion some of the possible sources of error in that study. It seems to us that the use of only a single measure of undergraduate achievement in a study using 69 measures of institutional quality and 103 of student characteristics is inconsistent and exposes the whole study to any weaknesses that single measure may have. We suggest that the Graduate Record Examination Area Tests, which were that single measure, are inadequate as a measure of achievement in ways which are precisely such as to bias the study in the observed directions.

We have examined an area test in the natural sciences. All the informational content examined in physics, chemistry, and biology is presented in high school courses. Although the use of mathematics in physical description is a most important part of material science, there was not a single problem on the examination requiring mathematics beyond the 9th-grade level. In short, the examination measures predominantly information from high school and early college levels, not the level of sophistication or even of information one expects from a science major at a major university. While we are less well qualified to judge the other two sections of the area tests, we believe them to consist similarly only of survey course material.

There are two aspects to education at most colleges and universities. One is the continuance of education on a broad front through introductory and general courses, fundamentally an extension of high school education, to a slightly higher level. The area tests attempt to measure such education. The second aspect is education in depth in a field of specialization. This aspect is the major reason for the existence of colleges and universities as distinct

Letters

from junior colleges or extended high school education. The area tests totally fail to examine education in depth. Astin's measure of achievement has therefore systematically ignored the unique and important feature of university education. At the same time, most of the conventional measures of university academic excellence with which "achievement" is being correlated are related to this education-indepth aspect.

It is then to be expected that Astin finds little correlation between the cost of education and achievement. "Achievement" as measured by the area tests describes only the general and introductory aspect of education. The high cost of good university educationthe need for huge libraries, good laboratory facilities, and teaching bv Ph.D.'s-is chiefly due to the educationin-depth aspect of universities. If an adequate measure of achievement could be found-and we have none to suggest-a study such as Astin's could be quite informative. Astin's use of a biased measure of achievement, however, hangs quality education without a trial. PHILIP W. ANDERSON

Bell Telephone Laboratories, Murray Hill, New Jersey 07974

JOHN J. HOPFIELD

Department of Physics, Princeton University, Princeton, New Jersey 08540

Anderson's and Hopfield's concern that the large number of input and environmental measures "is inconsistent and exposes the whole study to any weaknesses that single measure [the Graduate Record Examination] may have" is difficult to understand. We used a large number of input variables to reduce the chances that variations among institutions in the characteristics of their entering students would result in spurious college "effects," and a large number of environmental variables to maximize our chances of detecting any true environmental effects.

Whether or not one agrees with Anderson and Hopfield that the GRE area tests measure only "survey course

material" and "totally fail to examine education in depth," the important points are (i) that there was considerable variation among the students in their performance on these tests (even among students attending the "highest quality" institutions); and (ii) that little, if any, of this variation could be attributed to differences in the "quality" of the institutions attended. Of course, it may be that the higher quality institutions de-emphasize basic or introductory material in favor of "in-depth" coverage, but it is difficult to see much virtue in such a practice so long as there are still such wide variations among the students in their grasp of the fundamentals. We are currently conducting similar analyses using the advanced tests of the GRE, which presumably measure more of the "in-depth" knowledge that concerns Anderson and Hopfield.

It would be interesting to know if the area tests would have been judged to be "superficial" if our results had shown that attending a "high quality" institution enhances the student's performance on these tests. Unless such judgments about the relevance of evaluative criteria can be made independently of evidence concerning the differential effects of institutions on these criteria, there is a real danger that the folklore about institutional "excellence" will become a self-fulfilling prophecy.

ALEXANDER W. ASTIN American Council on Education, 1785 Massachusetts Avenue, NW, Washington, D.C. 20036

Versatile Genius: Frederick II

In view of the letters (2 Aug., 27 Sept.) concerning the accomplishments of Frederick II of Hohenstaufen in the field of ornithology, it seems unfair to his genius to pass over his feats in other branches of experimental science, as he was no narrow specialist. He was also:

1) A physiologist. He fed two men sumptuously at dinner, and then sent one to sleep, the other to take vigorous exercise. After a sufficient interval, he caused both to be opened in order to judge which had digested better (1).

2) An anatomist. He enclosed a man in a hermetically sealed cask. Since the cask, when opened, showed no soul, but only a corpse, he concluded that man had none (1).

3) A geneticist-linguist. He ordered foster mothers to care for some new-