The Reputations of American Medical Schools

Jonathan R. Cole; James A. Lipton

Social Forces, Vol. 55, No. 3 (Mar., 1977), 662-684.

*The Reputations of American Medical Schools**

JONATHAN R. COLE, *Columbia University*

JAMES A. LIPTON, *Columbia University*

ABSTRACT

This study is one of a larger inquiry into organizational stratification. About U.S. medical schools it asks: How does a sample of full-time clinical and basic science medical school faculty rank 94 medical schools as to quality of faculty and effectiveness of instruction? And: What are the structural correlates of such rankings? The resulting rank order takes on significance as it affects recruitment and placement of students and faculty. Measures of aggrandizement (inflated estimates of worth by insiders) are estimated. Characteristics of medical schools that correlate with perceived quality are: research and publication, eminence of faculty, training and research grants available, size of full-time faculty, and perceived effectiveness of training. While the data support the view that reputation stems from functionally appropriate performance, there is some evidence of a ceiling effect (Harvard) and a halo effect for schools affiliated with universities having national reputations. Regional location is positively associated with perceived reputation in the North and West, negatively in the South. Caveats are entered about interpreting the data.

There is now a history spanning almost 50 years of studies that attempt to measure the quality of graduate departments of Arts and Sciences in the United States. The earliest assessment was by Raymond M. Hughes; three decades later a study was made by Hayward Keniston, followed by two American Council on Education studies, by Allan M. Cartter, and by Kenneth D. Roose and Charles J. Andersen.[1] The last two studies reported rankings for engineering departments, but not for other professional schools. Very little previous work has been done on the reputations, or comparative perceived quality, of professional schools generally, and medical schools specifically.[2] This paper reports findings of a study designed to examine the reputational standings of American medical schools as assessed by the community of physicians and scientists who work in them.[3]

While reputation should not be equated with quality, it also should not be dismissed as an insignificant part of the social reality of the medical community.[4] Reputation makes a difference because it has multiple consequences for students, for faculty members, and for medical schools. Students are concerned with the reputation of medical schools when they elect to apply to some rather than others. They are aware that the reputation of their *alma mater* has an impact on their subsequent career mobility; they perceive the medical school as a first, but critical stepping stone in the medical career, opening or closing future opportunities.

*This study was supported by a grant from the National Science Foundation to the Columbia University Program in the Sociology of Science, NSF–SOC72–95326; by the Center for Advanced Study in the Behavioral Sciences, where the first author was a Fellow in 1975–76 ; and by a Guggenheim Fellowship, held by the first author in 1975–76. We thank Bernard Barber, Peter M. Blau, Stephen Cole, and Robert K. Merton for their helpful comments on an earlier draft of this paper.

Further, the reputation of their school influences students' self-esteem and affects perceptions of their ability within their significant reference groups. Faculty members are interested in reputations of schools when they consider appointments and promotions, not only because these reputations affect their own visibility and perceived ability in the larger medical community, but more basically because they hinder or enhance opportunities to obtain resources and facilities necessary for research. Correlatively, medical schools are concerned with their reputation because it affects their success in recruiting able faculty and outstanding students, and in obtaining resources to carry out basic and clinical research. In short, general reputation has much to do with the actual quality of a medical school.

Previous work in the sociology of science has demonstrated that the reputations of scientists depend largely on meritocratic or universalistic bases. The quality of research performance and the honors received by scientists explain a large portion of the total variance on reputational standing. Sheer quantity of research has little independent affect on reputations (Cole and Cole, c; Gaston). The process of accumulating advantage also operates to reinforce and enhance the reputations of scientists who make important discoveries early in their careers (Merton, c; Zuckerman). The difference in reputation between the "rewarded" and the "unrewarded" increases over time because resources for future research are disproportionately placed in the hands of scientists who have made discoveries in the past. Most of this earlier work focused on individuals. In this paper, we examine schools, and ask whether the basic stratification processes that affect reputations of scientists are similar to those affecting the standing of medical schools.

The data reported in this paper are, then, part of a larger ongoing study of stratification processes within a variety of social institutions. They are intended to promote further inquiries by students of the social stratification system within the medical profession. More immediately, we center on two questions: What is the distribution of American medical schools in terms of their perceived quality and their visibility among a stratified random sample of full-time faculty members on the medical and basic science staffs of these schools? What are the organizational and structural correlates of these rankings?

METHOD

A short questionnaire was sent to full-time medical school faculty members within all clinical and basic science departments in 87 American medical schools. All schools in the United States approved by the American Medical Association as of 1971–72 were included in the population of schools from which the sample of respondents was drawn. There were 94 such schools in all, but 7 were eliminated from the sample: 4 because available catalogs did not include the names of full-time faculty members after 1970, and 3 because they did not differentiate their faculties by academic rank.[5] The sample was stratified by three academic ranks—full professor, associate professor, and assistant professor—and included every twentieth faculty member within each academic rank. The questionnaire was designed to obtain an appraisal of the faculty quality and the effectiveness of training in all 94 fully approved institutions.

The questionnaire has two parts. Part I asked the medical faculty to rate medical schools on two dimensions of reputation: the perceived quality of its medical faculty, and the perceived quality or the effectiveness of its medical training program. This questionnaire format is comparable to that used in the American Council on Education studies of Graduate Arts and Sciences departments.[6] Faculty members were asked: ''Please circle the number under the term that corresponds most closely to: A. Your judgment of the quality of the medical faculty at each institution. B. Your rating of the effectiveness of the medical training program at each institution.'' Seven response categories for A were presented to the physicians and scientists: Distinguished (6); Strong (5); Good (4); Adequate (3); Marginal (2); Poor (1); Insufficient Information (0).[7] Five response categories were provided for B: Extremely Attractive (4); Attractive (3); Acceptable (2); Not Acceptable (1); Insufficient Information (0).[8]

Part I contained an additional question: ''What factors did you consider important when evaluating the quality of the medical faculty—and the effectiveness of medical training—for the schools listed?'' Answers to this question were used to compare the subjective criteria of evaluation with quantitative correlates of perceived quality. Part II of the questionnaire requested information about the social background of respondents, such as their age, academic rank, and medical or scientific specialty.

It was not practical to ask each faculty member to rate all of the 94 schools. Instead three forms of the questionnaire were designed. Each contained the names of 40 medical schools arranged alphabetically by the state in which they are located. Thirteen randomly selected schools appeared on all three forms, in order to test for comparability of rankings by faculty members responding to the various forms. We received 186 usable questionnaires for Form I, 193 for Form II, and 204 for Form III. Rankings on the different forms are highly consistent. Johns Hopkins and UCLA medical schools are two typical examples drawn from the 13. The quality-of-faculty scores (sometimes referred to as ''perceived quality scores'') for Johns Hopkins were 5.11, 5.13, and 5.11, respectively, on each of the three forms; for UCLA they were 4.66, 4.55, 4.62, respectively. In short, the respondents to all three forms of the questionnaire answer in roughly the same way.

In all, 2,049 physicians and scientists were sent questionnaires; 30.3 percent responded. This rate of response and its distribution requires examination. There were 583 usable questionnaires; 19 which could not be used; and 61 which were returned because the faculty member was no longer affiliated with the medical school. The response to the questionnaire by faculty members of different ranks was not fully representative of the population. Full professors represented 29 percent of the respondents, but only a fifth of the sample; associate professors, 36 percent of the respondents but only a quarter of the sample; and assistant professors, 35 percent but roughly half the sample. There are, of course, several plausible reasons for this unrepresentative response rate. It may be that full and associate professors, with longer experience in the medical science community, are more likely to have extensive information about programs at other medical schools. They are apt to feel an obligation to know what is going on at other schools; it is a role-appropriate response. Perhaps younger scientists and physicians, more engrossed in surviving within the competitive world of American medical schools, are less likely to allocate time for answering questionnaires. Or, age and academic rank may be inversely related to skepticism about the value of social science research focusing on medical schools. Other factors may be operative as well.

In any case, though different ranking scientists and physicians respond in disproportion to their numbers in the population, the critical question is: Do they rate medical schools differently? As it happens, the differing response rate among the variously ranked faculty members creates no special problem in measuring the reputation of medical schools. The pattern of evaluations is almost exactly the same for all academic ranks. The rank-order correlation between the rankings of schools by full professors and associate professors is $r = .99$; between full and assistant professors, $r = .96$; and between associate and assistant professors, .97.[9] The extraordinarily high correlations indicate that there is no significant bias in terms of professional rank; rank simply does not influence the rating of the schools.

Of course, there are many possible bases of differential awareness and evaluation of medical schools. We examined several of these, including those that might result from differences in age, quality of Ph.D. departments among scientists in the schools, and scientific or clinical speciality. Here no significant rank-order differences were obtained.[10] There was, however, one factor that did significantly influence evaluations: there were self-aggrandizement effects in rating medical schools.

To examine self-aggrandizement, the rating of Columbia, for example, by physicians and scientists currently affiliated with Columbia would be compared ideally with ratings by non-Columbia faculty. Since there were not enough respondents from each school to make individual comparisons, we aggregated respondents affiliated with similarly ranked schools. The procedure followed requires review. For each school in a group, the average rating given to a school by all those faculty affiliated with it was obtained. These ratings were summed over all raters in a group, and a group mean was computed. This mean is the average evaluation given by insiders to their own schools. For instance, 40 respondents were affiliated with medical schools ranked among the first 10 in quality of faculty. The sum of the ratings they gave to their own school, 216, was divided by 40, giving an average rank of 5.40. This insider score was then compared to the average rating of the same schools as judged by faculty members not affiliated with them (outsiders).Of course, in comparing a group of, say, 10 schools, a faculty member may hypothetically be an insider only once, but an outsider as many as nine times.[11] The grand mean for affiliated and non-affiliated respondents was computed for each group. The difference between the mean perceived quality scores of insiders and outsiders is taken as a rough estimate of self-aggrandizement resulting from current affiliation.[12]

A second type of self-aggrandizement was estimated, that resulting from receiving medical training at a school. The results of this analysis may be seen by comparing the columns of Table 1.

There are two consistent patterns to the data. First, faculty members rate their own schools significantly higher than do others in the medical community. The average difference in means between insiders and outsiders is .67. Among the top 10 medical schools, for example, insiders rated their schools at 5.40 compared to 4.93 among outsiders. Self-aggrandizement is particularly strong among faculty members affiliated with lower-ranked medical schools, from 61st to 94th: insiders rated these schools 3.86; outsiders, 3.00. These are statistically significant differences.

Second, there is more self-aggrandizement in rating *alma mater* than in rating current affiliation. Stratifying the faculty members by the medical school from which they received their M.D. training produced average differences between insiders and outsiders of .73. Faculty members who received their medical education from the lower-ranked group of schools are most apt to overrate their *alma maters*; correlatively, products of the highest rated schools are least likely to overrate them.[13]. Although we find self-aggrandizement in scores, it does not significantly distort the overall rank ordering of medical schools. The rank-order correlation between perceived quality scores for insiders and outsiders is .84; between the rankings of schools by the entire sample and outsiders it is .998. In sum, there appears to be an extraordinary degree of consensus within the medical school community about the relative standing of the faculty at the 94 schools.

These results are consistent with the Cartter study findings on self-aggrandizement. But they run counter to self-aggrandizement findings reported in occupational prestige studies. At every level of prestige in the occupational structure, people rate their own occupations almost exactly as others rate them.[14] How can we account for the different patterns?

Perhaps the absence of significant self-aggrandizement in estimating occupational prestige results from continual reinforcement of the actual prestige position of an occupation. Attempts by individuals to inflate the prestige of their own occupation are continually negatively sanctioned in interaction. There are fine-tuned social thermostats which constantly feed back to incumbents the relative prestige of their occupation. Fine-tuned feedback mechanisms may not exist in the medical school community. Physicians and scientists in variously ranked medical schools tend to associate predominantly with others in the medical community who are in similarly ranked schools.[15] This pattern of differential association may reinforce tendencies toward self-aggrandizement.

These faculty quality estimates assume that those who respond represent the population. We know that this was not so in the medical school study, but the slight unrepresentativeness of the sample among different ranks apparently makes little difference in the rankings obtained. It does remain possible, however, that all of the respondents, regardless of academic rank and other social characteristics, differ significantly in their assessments from nonrespondents.[16]

Table 1. ESTIMATES OF SELF-AGGRANDIZEMENT AMONG PHYSICIANS AND SCIENTISTS LOCATED IN AMERICAN MEDICAL SCHOOLS

	Mean Perceived Quality Ratings of Medical Schools			
Rank of Medical School	*Affiliated with the School**	*Not Affiliated with the School*	*Attended the School**	*Did Not Attend*
1 - 10	5.40 (40)	4.93 (2,246)	5.60 (93)	5.08 (2,426)
11 - 20	5.37 (49)	4.50 (2,087)	5.13 (30)	4.53 (1,900)
21 - 40	4.60 (47)	3.92 (2,204)	4.76 (33)	3.94 (2,023)
41 - 60	3.97 (31)	3.51 (2,062)	4.20 (25)	3.54 (1,859)
61 - 94	3.86 (43)	3.00 (3,687)	4.06 (17)	3.01 (3,440)

*The number of faculty included under the affiliated and attended columns does not equal the total sample size, because there were three forms to the questionnaire, and in some cases respondents returned questionnaires rating schools other than the one at which they were currently located.

The unit of analysis is the medical school, not the individual faculty member. There are two dependent variables in this study: the perceived quality and the visibility of medical schools.[17] Perceived quality will be discussed in detail, but visibility only cursorily. A school's visibility score is taken as the percentage of all respondents who felt that they had sufficient information to rate it. It is obtained by dividing the total number of assessments (6 to 1), regardless of the quality ascribed, by the total number of respondents who returned the questionnaire. Perceived quality of faculties of medical schools is measured by considering only the evaluations of respondents who actually assessed the quality of a school's faculty. Among those who made judgments, the perceived quality score is the mean rating of the respondents. It has been noted that these scores vary from a high of six (''Distinguished'') to a low of one (''Poor''). The number of cases on which perceived quality scores are based varies among schools of differing degrees of visibility.

We collected data on thirty-seven characteristics of each medical school. Some of these variables were computed by aggregating information obtained for each of the faculty members within the school.[18] For example, data on the number of papers published in scientific journals in 1972 by members of the clinical and basic science faculty of each medical school were collected from the *Source Index* of the *Science Citation Index* (SCI).[19] The aggregated number of publications was used as a characteristic of the medical school. Similarly, the eminence of the faculty of each school was estimated by aggregating the number of faculty members elected to the Association of American Physicians (AAP), ''an organization limited to individuals who have made distinguished contributions to medical science.'' A second aggregated indicator of eminence was the total number of chairmen and deans of medical schools that any school had produced. Other characteristics of each school, not based on aggregation procedures, included the total funds it received in 1969 from the National Institutes of Health, and from the Department of Health, Education and Welfare; the total funds received from these sources for research and development; and the total for training graduate as well as post-graduate students and fellows. Data were obtained on faculty size, on social and individual characteristics of students, such as the sex composition and the applicant-to-acceptance ratio, as well as on the achievements of the school's graduates. All of these data were obtained from available sources (American Association Medical Colleges; Association of American Physicians; Council on Medical Education; Dube; Giza and Burns; Institute for Scientific Information; Singletary; Theodore).

FINDINGS

A basic hypothesis of ours is that the reputations, or perceived quality, of medical schools will depend largely, in fact almost entirely, on their performance as scientific and research organizations. This assumes, of course, that reputation is largely a consequence of rational processes and of the application of performance criteria of evaluation. To test this hypothesis we will want to consider several questions. How much do the reputations of medical schools reflect their functional performance in research and their contributions to the growth of knowledge in the clinical and basic sciences? How strongly is reputation influenced by the level of federal funding of research? How closely is reputation related to the number of eminent stars on a school's faculty? Do irrational, or functionally irrelevant, characteristics of a school, such as its geographical location, or its sex and racial composition, influence its reputation? Finally, what social processes contribute to the persistence or change in reputations of American medical schools?

To address these questions, we turn to the results obtained in the survey. Table 2 presents the medical schools in rank-order of perceived quality of their faculty. The itemized list is presented in detail both for its intrinsic interest and for its possible use in future research. The analysis in this paper, however, deals rather

Table 2. PERCEIVED QUALITY AND VISIBILITY SCORES OF AMERICAN MEDICAL SCHOOLS

Medical School	*Perceived Quality Score*	*Std. Dev.*	*Visibility Score*	*Number of Raters*
Harvard	5.71	0.54	87.3	509*
Johns Hopkins	5.11	0.92	84.7	494*
Stanford	5.11	0.84	81.2	151
California, San Francisco	5.01	0.76	75.1	145
Yale	5.00	0.79	82.0	478*
Columbia	4.93	0.86	79.2	462*
Duke	4.77	0.82	82.4	159
Michigan	4.74	0.82	76.2	147
Cornell	4.71	0.80	76.9	143
Washington, St. Louis	4.68	1.00	80.3	155
U. of Pennsylvania	4.66	0.84	75.6	146
Minnesota	4.62	0.82	69.0	402*
UCLA	4.61	0.81	74.4	434*
Albert Einstein	4.60	0.94	70.1	143
U. of Chicago, Pritzker	4.52	1.06	57.0	110
U. of Washington, Seattle	4.52	0.81	69.5	405*
Case Western Reserve	4.41	0.86	76.7	148
Rochester	4.37	0.79	69.4	134
Colorado	4.36	0.86	71.5	133
California, San Diego	4.27	1.01	60.3	123
Mount Sinai	4.22	0.94	67.9	131
NYU	4.18	1.03	60.8	113
Texas, Southwestern	4.11	1.00	48.2	93
Vanderbilt	4.10	0.90	61.8	126
North Carolina	4.07	0.76	59.8	122
Baylor	4.06	0.95	64.7	132
Tufts	4.05	0.91	68.6	140
U. of Wisconsin	4.02	0.85	59.3	121
Northwestern	3.98	0.87	67.7	126

Table 2, continued

Medical School	*Perceived Quality Score*	*Std. Dev.*	*Visibility Score*	*Number of Raters*
Emory	3.95	0.97	63.2	129
Boston University	3.95	0.94	65.1	121
Iowa	3.93	0.82	63.7	123
U. of Virginia	3.89	0.85	58.0	112
Ohio State	3.79	0.84	64.0	119
Alabama	3.78	1.00	48.9	91
U. of Florida, Gainesville	3.77	0.88	60.1	116
Dartmouth†	3.73	0.91	59.5	347*
Illinois	3.70	0.89	59.7	111
Tulane	3.68	1.00	62.7	128
Georgetown	3.68	0.96	75.1	145
Utah	3.68	0.82	50.0	93
Cincinnati	3.68	0.87	55.9	114
California, Davis	3.66	0.85	50.5	94
Penn State	3.64	1.03	46.8	87
Pittsburgh	3.64	0.80	59.7	111
Vermont	3.62	0.80	50.0	102
Virginia Medical College	3.62	0.85	60.2	112
Oregon	3.58	0.85	48.2	93
State University of New York, Syracuse (Upstate)	3.57	0.89	58.6	109
Michigan State	3.54	1.20	49.5	101
Indiana	3.54	0.87	52.0	106
Buffalo	3.53	0.84	59.1	114
Texas, Galveston	3.48	0.94	48.9	91
St. Louis	3.48	1.19	60.8	124
Temple	3.46	0.79	61.3	125
Miami	3.43	0.92	65.6	122
Medical College of Wisconsin	3.43	1.05	52.8	102
Maryland	3.41	0.82	64.2	124
Kansas	3.40	0.84	52.2	97
Albany	3.37	0.75	56.4	115
Bowman Gray	3.37	0.87	59.7	98
Arizona	3.34	0.95	37.7	77
Missouri	3.33	0.87	50.5	94
California, Irvine	3.32	1.04	45.6	93
George Washington	3.31	0.95	67.2	125
State University of New York, Brooklyn (Downstate)	3.31	0.84	57.4	117
Texas, San Antonio	3.25	0.75	34.7	67
Wayne State	3.24	0.89	54.8	102
Chicago Medical School	3.21	1.21	55.2	322*
Oklahoma	3.20	0.76	43.9	256*
Kentucky	3.20	0.82	52.9	97
Jefferson	3.19	1.02	61.1	118
New Mexico	3.17	0.80	50.0	93
Georgia Medical College	3.12	0.81	42.0	81

Table 2, continued

Medical School	*Perceived Quality Score*	*Std. Dev.*	*Visibility Score*	*Number of Raters*
Tennessee	3.12	0.92	45.2	84
Louisiana State	3.05	1.07	45.7	85
Arkansas	3.05	0.78	39.9	77
Connecticut	2.96	0.95	44.6	91
Louisville	2.94	0.90	52.8	102
Medical College of Pennsylvania	2.94	1.07	47.8	89
Hahnemann	2.94	0.96	52.5	107
Loma Linda	2.93	1.00	46.6	90
West Virginia	2.92	0.81	36.7	214*
Nebraska	2.92	0.86	38.7	79
New York Medical College	2.92	1.00	59.0	344*
South Carolina	2.91	0.83	36.3	70
Mississippi	2.86	0.95	41.5	80
Ohio, Toledo	2.84	1.02	29.9	61
Howard	2.69	0.93	52.9	108
Loyola	2.64	1.02	42.2	86
Creighton	2.48	1.04	45.7	85
New Jersey	2.40	0.90	47.7	92
Puerto Rico	2.29	0.75	17.2	35
Meharry	2.23	0.85	47.8	239*

*These 13 medical schools appeared on all three forms of the question-naire. This accounts for the larger number of raters.

†Dartmouth, of course, provided education only in the basic sciences at the time of survey.

with those characteristics of schools which are associated with the distributions of these rankings than with the absolute values of assessments for particular schools.

The perceived quality scores range from a high of 5.71 to a low of 2.23; the mean score is 3.68 with a standard deviation of .70.[20] To guide interpreting differences in perceived quality scores, we computed the standard error of the estimate for each score.[21] These errors varied very little from one score to another. The mean standard error for the entire set of rankings is .082. It would be a mistake to attribute any statistical significance, therefore, to scores differing by less than twice the standard error, that is, by .164.[22] For example, it would be an obvious mistake to take the difference in scores between, say, Stanford (5.11) and Johns Hopkins (5.11), on the one hand, and Yale (5.00) on the other as either statistically or substantively significant. However, differences between school ratings of 5.00 and those of, say, 4.00 would be statistically significant. Of course, even scores that differ by more than twice the standard error may not be substantively significant.

Medical schools vary significantly in terms of their visibility to others within the medical school community, that is, in the percentage of faculty from other schools who feel they know enough about the designated school to appraise its faculty. The mean visibility score is 57.7; the standard deviation, 13.3.[23]

We come then to the central question: What characteristics of medical schools are correlated with their perceived quality; and what variables have the strongest independent effect in predicting these scores?

We hypothesized that the stratification of medical schools' perceived quality was largely a result of their functional performance in research.[24] Schools with the greatest resources in support of research, with faculty producing the most research, and with faculty recognized and honored for their research performance, would be rated most highly.[25] If the hypothesis has any face validity there must be fairly strong correlations between indicators of research performance and the perceived quality of schools. How much support is there, in fact, for this hypothesis?

In arriving at their appraisals, medical faculty gave quality of research far greater importance than teaching and other features of schools. Coding criteria given for the assessments of faculty research performance (180 mentions) and eminence of faculty (203 mentions) ranked just behind personal knowledge of or acquaintance with faculty members at other schools (214 mentions) as the most frequent bases given for evaluations. Teaching performance, listed next most frequently, was mentioned only 61 times. Are these bases for appraisals consistent with the correlation between perceived quality scores and objective measures of research output and eminence of faculty?

Table 3 presents a set of zero-order correlations of selected independent variables with the perceived quality of medical schools.[26] The data suggest a strong association between the publication of scientific research and the perception of quality of the schools at which the work originated. The primary indicator of productivity, as noted above, is the aggregated number of papers listed in the 1972 SCI, published in scientific journals by both the clinical and the basic science faculties of each medical school.[27] The correlation between perceived quality and this indicator of scientific output is .87, thus itself accounting for 75 percent of the variance on perceived quality.

The association between the number of research publications and perceived quality scores for the 94 medical schools is represented in the scattergram of Figure 1. This association corroborates the criterion that the physicians and scientists reported having used in arriving at evaluations. If the unit of analysis had been the individual physician or scientist, rather than the medical school, the correlation between perceived quality and productivity would have been lower than .87 because there is, of course, significant variability in the number of scientific papers published by individual faculty members in schools of varying perceived quality.

While productivity is in general strongly associated with the perceived quality of medical schools, the various indicators of research productivity presented in Table 3 yield differing correlations. As noted, the correlation for the faculty as a whole is .87. But if we standardize for size of faculty, thus producing average productivity per faculty, the same correlation is .58. This reduced correlation reflects the association between faculty size and perceived quality, and the skewed

Table 3. CORRELATION COEFFICIENTS BETWEEN SETS OF INDEPENDENT VARIABLES AND THE PERCEIVED QUALITY OF MEDICAL SCHOOLS

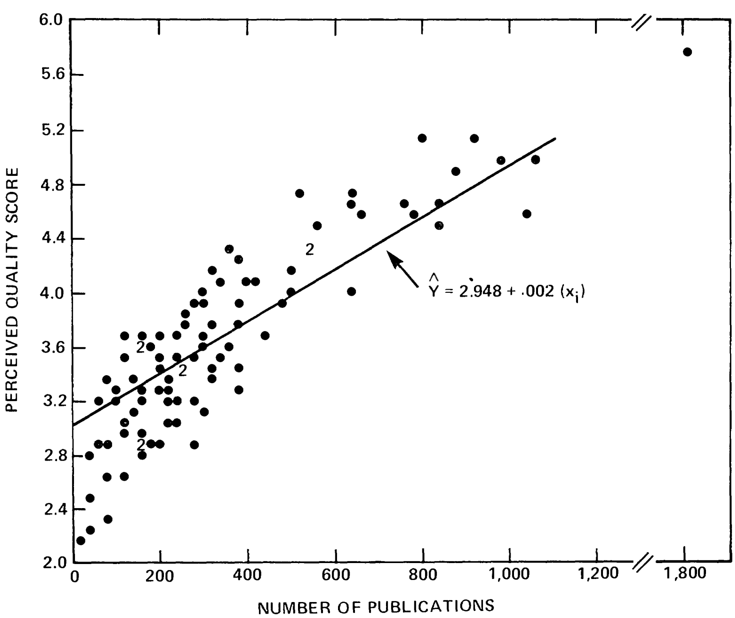
Independent Variables	*Dependent Variable* *Perceived Quality*
Productivity	
Total papers (clinical + basic science faculty - 1972)	.87
Papers/FTF (clinical + basic science)	.58
Papers (clinical faculty only - 1972)	.36
Eminence	
AAP members (total - 1972)	.79
Elites (deans & chairmen in 1972)	.60
Resources (1969)	
NIH total funds	.61
NIH graduate training & fellowship funds	.84
NIH research & development funds	.85
DHEW total funds	.63
DHEW graduate training & fellowship funds	.82
DHEW research & development funds	.84
Size (1971-72)	
FTF: total full-time faculty, clinical & basic science	.59
Total number of applicants	.23
Total number of students enrolled	.15
Percent female	.11
Total number of graduates	.17
Percent first year students from same state as medical school	-.34
Effectiveness of Training	.99

distribution of productivity within most medical schools. Indeed, it points to the possibility that the reputation of schools is more a function of the performance of a few stars located at the school than of the performance of most of its faculty.

Still further, the data suggest what we might expect, that there are different rates of scientific publication in the pre-clinical basic sciences and the clinical departments. Considering only the publications of clinical faculty members, less given than their basic science colleagues to work on research leading to publication, the correlation between scientific output and perceived quality is reduced to .36. This datum suggests that the perceived quality of medical schools is influenced more by the productivity of basic science than by clinical faculty.[28] In fact, the structural organization of medical schools may influence, in part, their perceived quality. Although most universities locate pre-clinical departments within the structure of the medical schools, some do not. Schools that do have affiliated pre-clinical departments benefit from the esteem associated with work produced by faculty members in those departments. Thus, it may seem questionable to include publications produced by pre-clinical faculty in the school totals, since some schools do not have well-funded basic science departments to contribute to their total. But the presence or absence of these departments in the medical school is part of the social reality of that school as viewed by the medical science community, and it is this social reality that we are trying to capture.

If the functional hypothesis is supported by the data on research productivity, it is given additional weight by data on the honorific achievements of a school's faculty. Total membership in the Association of American Physicians, one indicator of faculty eminence, is highly correlated with perceived quality, $r = .79$. A second indicator of faculty eminence was obtained by finding the number of medical school deans, and chairmen of departments of psychiatry, medicine, and surgery, in the academic year 1972–73, who had graduated from the school. The position of department chairman and dean carry more prestige in medical schools than they do in Arts and Sciences departments. Schools were, therefore, classified by their record in producing chairmen and deans. This indicator of eminence is correlated .60 with perceived quality. These associations between faculty eminence and perceived quality suggest that it is the leading figures, the relatively few eminent men and women on the basic science and clinical faculties, who are largely

Figure 1. SCATTERGRAM OF PRODUCTIVITY OF MEDICAL SCHOOL FACULTY BY PERCEIVED QUALITY



responsible for a medical school's general reputation. As we shall see, however, the presence of a significant number of eminent faculty is itself correlated with other attributes of medical schools.

Given the strong correlations between the perceived quality of schools and the research performance and recognition of their faculties, we should also find a strong association between the financial resources available for research at medical schools and their ratings.

Indeed, this is the case. We collected data on the amount of NIH and Department of Health, Education, and Welfare funds available for research and development and for educating doctoral and post-doctoral scientists, at each medical school in 1969. These totals are strongly associated with perceived quality. Total NIH funds, for example, is correlated .61 and total NIH research and development funds, .85. Although we do not present the figures in the correlation matrix (Table 4) for total HEW support of the various medical schools, the correlations between the HEW figures and perceived quality do not differ significantly from the perceived quality and NIH correlations. To cite only one example, the zero-order correlation between HEW research and development funds and perceived quality is .84.

These then are the basic zero-order correlations that lend support for the functional performance hypothesis. Is there any evidence to support an alternative hypothesis that irrational or functionally irrelevant criteria influence the reputations of medical schools?

Let us return to the data on faculty research performance and examine them more closely. The linear regression equation for the relationship between the production of scientific papers and perceived quality, illustrated in Figure 1, is:

$$\hat{Y} = 2.948 + .002X_i$$

where $\hat{Y}$ = predicted perceived quality score; and
X_i = scientific productivity of the medical faculty

This equation estimates how well faculty productivity predicts the perceived quality of a school. The high zero-order correlation means, of course, that it predicts perceived quality quite well. But there are some notable deviations from the least-squares regression line, a finding that gives us a clue to a functionally irrelevant basis of evaluation. It turns out that only 4 of the 94 schools had residuals greater or less than twice the standard error. Therefore, we lowered the criterion for identifying deviant cases, examining the list of schools that had estimated scores, or residuals, greater than or less than one standard error of the estimate, or $\pm .35$ from their actual scores. Clearly, we are searching for substantively interesting clues to behavior, rather than statistically significant patterns.

On the basis of faculty productivity, 27 schools had predicted scores higher than their actual scores; 22 had scores lower than predicted. Harvard is the most notable school among those whose actual perceived quality is less than predicted by scientific productivity. Although Harvard ranked first among all medical schools, its faculty is so extraordinarily prolific that we would predict its perceived quality to be

1.06 points higher than it is in fact—a score that would exceed the upper bound of the scale. Only one other medical school associated with a major university, UCLA, had a perceived quality score substantially underestimated by its faculty productivity.

Of the schools with scores lower than predicted by faculty productivity, 8 of the 22 are located in the South. None of the 8 is associated with better known southern universities. Consider a concrete example. On the basis of faculty productivity alone, the University of Florida, Gainesville, should have a perceived quality score .36 higher than it actually received. Similar discrepancies obtain for the Medical College of Georgia, L.S.U., University of Mississippi, South Carolina, Tennessee, and West Virginia. A number of southern schools with national reputations, including the University of North Carolina, Duke, University of Virginia, Emory, and Vanderbilt, had scores which were higher than we would predict on the basis of their faculty productivity.

That the predictions are higher than the actual ratings does not mean, of course, that the faculty at these southern schools are as prolific as those at higher ranked ones.[29] The residuals do suggest that halo-effects influence the perceived quality of medical schools. Those schools that are part of universities with national reputations have actual scores which either are predicted well or are somewhat higher than predicted by faculty productivity (Harvard and UCLA excepted). Correlatively, schools affiliated with universities without national standing tend to be victimized by their association—their actual scores are lower than predicted. Since these data do not control for the quality or impact of papers published by faculty members, the deviations from predicted values might well be lower if these factors were incorporated in the model.

The residuals suggest, then, a significant relationship between the geographical location of schools and their perceived quality. Regression analysis, in which region was recoded into a set of dummy variables, indicates that medical schools in the Northeast and the West have higher perceived quality scores, on average, than schools in other regions.[30] Location in the Northeast is correlated .19 and in the West, .17, with perceived quality. Correlatively, location in the South was negatively correlated with perceived quality scores, $r = -.20$. These effects of region, especially in the South, are not significantly reduced by the addition of performance variables to the regression equations.

Thus, we have at least one indication that the reputations of medical schools are not based solely on functional research performance. Are there other bases of evaluation which depart from the rational ideal? The data we have collected suggest that variables that could be taken as indicators of functionally irrelevant bases of evaluation have only minor effects. For example, the sex composition of the student population, which varies very little, is almost uncorrelated with perceived quality scores ($r = .11$). Similarly, privately endowed schools tend to have slightly higher scores than those that are state-supported ($r = .20$). Older schools, with longer traditions in medical education, have somewhat higher scores than younger schools ($r = .16$).[31] These correlations are quite weak, and in fact, since each of these independent variables is significantly correlated with the performance variables, their independent influence on perceived quality after controlling for faculty performance is minimal. In sum, geographic location of schools is the only factor that had an independent effect on perceived quality after controlling for research performance of faculty. The zero-order correlations suggest, then, that there is strong support for the research performance hypothesis.

We turn from the efficacy of the functional performance hypothesis to the relative predictive effects that variables associated with performance have on perceived quality scores. To estimate relative effects, we regressed perceived quality on six predictors: faculty productivity; faculty eminence; faculty size; total NIH graduate training and fellowship funds; total NIH research and development funds; and total post-doctoral students, fellows, and graduate students in clinical departments. The correlation matrix (Table 4) indicates that faculty productivity is strongly correlated with other performance-related variables. In fact, the addition of the five predictor variables increases R^2 from .75—that is, the total variance explained by faculty productivity alone—to .80.[32] The relative weights of these six variables are strongly influenced by the high intercorrelations between the predictors, since there is strong multicollinearity in these data. For example, the regression coefficient in standard form for productivity is .38; but for faculty eminence, which is correlated .79 with perceived quality, it is only .05. The level of NIH funding, however, does have some independent predictive effect. The beta coefficient for total research and development funds received by schools is .23; it is .16 for total NIH training and fellowship funds. The relative effects of these six predictor variables are presented in Table 5. It should be emphasized that given the high level of multicollinearity in these data, other regression equations produce roughly the same amount of total explained variance. But the estimates of R^2 are not influenced significantly by the restricted range in the dependent variable. Probit regression estimates yielded similar results.

We have avoided causal language, since it is in fact difficult to determine the causal sequence of the functional performance variables. The social processes that determine perceived quality are undoubtedly mutually interactive and self-reinforcing. By and large the same processes that determine quality also tend to maintain it. Although there may be notable shifts for a few universities that make extraordinary efforts at structural change and at faculty recruitment, there is probably much stability in these scores over time.

Consider how the current sharp differentiation in facilities, resources, and production of original scientific discoveries tends to maintain existing reputation distinctions. Past outstanding research performance by medical faculty increases the probability of financial support for current and future projects. But a high level of funding also makes high productivity and the production of original ideas feasible, especially in an age when basic science and medical research depend so heavily on adequate facilities and resources. It is this reciprocal combination—resources and performance—that attracts physicians and scientists of the first rank, as well as

Table 4. CORRELATION MATRIX OF SELECTED CHARACTERISTICS OF AMERICAN MEDICAL SCHOOLS

	Q	*V*	*P*	*N*	*M*	*F*	*E*	*S*	*G*
Q Perceived quality									
V Visibility	.846								
P Productivity - total[1]	.866	.749							
N NIH - total funds[2]	.608	.582	.681						
M NIH - research & development funds[3]	.846	.745	.880	.773					
F Faculty size[4]	.594	.610	.612	.573	.643				
E AAP members - eminence of faculty[5]	.790	.641	.879	.642	.805	.555			
S Number of post-doctoral students in clinical departments[6]	.683	.607	.706	.490	.614	.470	.622		
G NIH - graduate training & fellowship funds[7]	.843	.716	.861	.765	.936	.632	.827	.684	
Mean	3.68	57.7	346.6	6017[8]	3203[8]	314.7	3.7	60.9	1244
Standard deviation	0.70	13.3	288.4	5339	2952	207.2	4.9	94.2	1231

Notes 1-7: For a list of the sources of these variables, see note 26.

8: In thousands of dollars.

quality medical students. Since these institutional features of medical schools also result in added prestige, schools which already are superior are more apt to remain so.

DISCUSSION

The value of the earlier Cartter and Roose–Andersen studies of the quality of graduate departments is often obscured by the negative reactions to some of the gross measures used in their design. Those studies have been properly faulted for not specifying and measuring the multiple dimensions of the comparative quality of departments, and also for depending largely on mere estimates of prestige (Elton and Rodgers; Knudsen and Vaughan; Lewis; Magoun). The rough measures encouraged readers to assume that the small differences in scores could be equated to actual differences in quality. The numerical rankings tended to become reified.

With all their limitations these studies assessing the quality of departments had, to use Paul Allison's felicitous phrase, "potential for elaboration," and in fact contributed greatly to studies in the social organization of science. For instance,

Table 5. CORRELATIONS AND STANDARDIZED REGRESSION COEFFICIENTS OF THE DETERMINANTS OF THE REPUTATIONS OF AMERICAN MEDICAL SCHOOLS

	Dependent Variable: Perceived Quality of Medical School	
Predictors:	*Zero-Order Correlation*	*Regression Coefficient*
Productivity-total	.866	.38
NIH graduate training & fellowship funds	.843	.16
Size (number of fellows, graduate & post-doctoral students in clinical science departments)	.683	.12
NIH research & development funds	.846	.23
Size of full-time faculty	.594	.03
Eminence of faculty (number of members in AAP honorific society)	.790	.05
Multiple R		.89
R squared		.80
Residual		.45

studies of social mobility among academic scientists would have been difficult to do without even rough estimates of the quality of academic affiliations, in as much as these studies measured a form of recognition associated with prestigious academic affiliation (Cole and Cole, a; Crane; Gaston). The Cartter ratings made it possible to study in detail this feature of the reward system of science. Further, the Cartter and Roose–Andersen rankings provided an opportunity to investigate the social correlates of the evaluations, to uncover the variables that are associated with the rise and fall of academic departments (Hagstrom). Testimony to the impact of these ranking studies may be found in the number of citations to them in the *Social Science Citation Index* (SSCI). In 1974, the Cartter study, published eight years earlier, received a total of 25 citations. When this total is compared to the average number of citations received by any cited paper, which is just over one, the relative impact of the report should be clear.[33] Further, of these 25 citations, 23 came from different authors publishing in sociological, economic, and political science journals. The total of 25 citations does not include, of course, the many citations to papers that made extensive use of the ratings.

Nonetheless, studies assessing the quality of schools—the present study of medical schools, the Cartter study, and others—are subject to several limitations. Let us note the limits to the data reported here. First, we do not investigate the *quality* of medical schools *per se*, but only their *perceived quality* (Cole and Cole, c). This is not a trivial distinction. The perceptions of these schools by the physicians and scientists may be based on inadequate or false information; their images may be distorted or dated. The perceived quality may reflect a halo effect generated by the reputation of the larger university with which the medical school is associated. When future research on the quality of medical schools, using different indicators of quality, is accomplished, it will be useful to compare the several measures.

Second, the rankings presented in this paper are not fine measurements. Differences in rank between any adjacent ranking schools are not necessarily substantively significant. In short, let us not reify these numbers. Third, the data are the product of a single measuring instrument at one point in time. Other measurements might change some of the obtained rankings. Moreover, slightly different wording of the questions to which the medical faculty responded might yield somewhat different results. Fourth, our study does not deal with the problem of intersubjectivity. We did ask physicians and scientists to indicate the criteria they used to arrive at their judgments. But they may have used differing subjective criteria and levels of actual knowledge to reach the same judgment that the quality of a particular school is, for example, distinguished. We have no way of telling the relative weights assigned to evaluation criteria by respondents reporting their judgments.[34] Finally, the questions on which our rankings are based ask only for evaluations of the quality of the faculty of a medical school as a whole.

Clinical and basic science departments within each institution are not assessed separately. This makes for rough overall estimates, since there is, of course, extensive variability in both the actual and the perceived quality of different departments within the same school. In large part, physicians and scientists base their evaluations on knowledge of their own specialties or related ones in other schools, rather than on actual knowledge of the entire faculty. The measures of perceived quality and of visibility are for the medical schools as a whole, not for specific specialties.

What then can we tentatively conclude from the data reported in this paper? Medical schools in contemporary American society vary widely in terms of their reputations, that is, in terms of their perceived quality and visibility. The characteristics of medical schools that predict these ratings, and are likely to determine them, are associated with basic science and clinical research performance: faculty productivity, eminence of the leading members of the faculty, and resources available for research. In sum, the hypothesis that functionally relevant criteria of evaluation of medical schools predominate in the formation of assessments of their quality is strongly supported by the data. There is some limited evidence, however, that functionally irrelevant characteristics of schools and of raters influence evaluations. Medical schools located in different geographical regions receive somewhat varied evaluations, and ratings by alumni of their *alma maters* and by current faculty of their present affiliation produces some self-aggrandizement.

NOTES

1. Actually, estimates of quality date back to 1911 when the Bureau of Education prepared a rating of 344 institutions at the request of the Association of American Universities. The study by Hughes in 1925 was repeated by the American Council on Education in 1934. The results of the 1934 study were reported by Wilson.
2. Margulies and Blau and Blau and Margulies report evaluations of American professional schools, based on questionnaire responses by deans of the various schools. Each dean was asked to list the top five institutions in the profession. Of course, for each discipline only a limited number of schools were mentioned. For medicine, only 5 schools were listed, and these ratings were based on a sample of eight deans responding to an initial survey. Based upon a followup survey, 11 schools, rated by 51 Deans, were listed in rank order.
3. The absence of systematic estimates of the rank of medical schools has created multiple research problems for sociologists of medicine. Some of these are discussed in Barber et al.
4. Extended discussions of the significance of reputational rewards in science are found in Merton (a).
5. The available catalogs for Loma Linda, Loyola, Harvard, and Michigan State were dated prior to 1970; Howard, New Mexico, and Texas–San Antonio did not distinguish the ranks of faculty members in their bulletins.
6. The ACE format was adopted in order to compare the scale scores for the medical schools and those for Arts and Sciences departments.
7. The actual questionnaire used a scale range of seven to one, with "insufficient information" being scored "one." The category "insufficient information" is here scored "zero," and the scale has been transformed into a six-point scale.
8. It turns out that the correlation between the faculty scores and the training scores is .99. In effect, there are no differences in the ratings of medical schools along these two dimensions. Therefore, the training scores need not be discussed.
9. These rank-order correlations are for the responses to (Form II). The same pattern is found for the other two forms.
10. On differential awareness of the performance of scientists, see Cole and Cole (b).
11. For those schools in the top 10, there were a total of 2,246 responses, since each respondent could theoretically rate at least 9 institutions, and 10 if he or she was not affiliated with a school in this group. When the sum of the ratings was divided by 2,246, the overall mean was computed as 4.93 for the 10 schools.
12. The insider and outsider perspectives are developed by Merton (c).

13. These results are not attributable to a ''ceiling effect''; the average scores do not approach the scale's upper bound.
14. Religious groups tend to aggrandize their own reputations. D. J. Treiman reported these unpublished data to me in private communication.
15. The idea of differential association must remain speculative until more research is done on the interaction patterns of physicians and scientists at different schools.
16. An attempt to check on nonresponding faculty by a limited telephone inquiry found that faculty members who did not answer in the first place would not respond to further queries. Consequently, bias resulting from differences between those who responded and those who did not remains a possibility.
17. Many interesting questions about the variations in individual responses to the set of medical schools will not be addressed here. For example, are the perceptions of schools influenced by the eminence of the observer? For discussions of problems similar to these in the larger scientific community, see Cole and Cole (b).
18. On types of aggregation as a method of constructing new variables, see Lazarsfeld.
19. Publication counts are limited to the journals abstracted and listed by the Institute for Scientific Information, in the Source Index of the SCI. More than 2,500 journals are abstracted annually.
20. The omission of the University of Southern California medical school was the result of a clerical error in printing the questionnaire. Its perceived quality and visibility scores are not presented in Table 2. Data on the characteristics of Southern California and its faculty were collected, however, and on the basis of a prediction model discussed below, which explains more than 90 percent of the variance on perceived quality, Southern California had an estimated perceived quality score 3.96. A small follow-up study could be designed to assess the accuracy of this prediction.
21. The standard error is, of course, easily obtained by dividing the standard deviation of the ratings for a school by the square-root of the number of raters, i.e., $s / \sqrt{n}$.
22. When examining scores in terms of their standard errors, two standard errors difference is the conventional rule for establishing statistical significance. Thus, scores which do not differ by more than this should be viewed cautiously in terms of their statistical significance.
23. There will be no further detailed discussion of visibility scores and their correlates. The determinants of visibility are much the same as those for perceived quality.
24. For the functional principle from which this hypothesis derives, see Barber; Cole and Cole (c); Merton (b).
25. The relationship between scientific productivity and scientific recognition has been a major focus of attention in the sociology of science for the past 10 years. In particular see Barber et al.; Cole and Cole (c); for a discussion of this relationship for faculty members in two types of hospitals.
26. There are only two significant differences in the way visibility and perceived quality scores correlate with the independent variables presented in Tables 3 and 4. Indicators of research productivity, faculty eminence, and federal resources are consistently, but only slightly, more highly correlated with perceived quality than with visibility; correlatively, indicators of size of schools, student populations, and numbers of graduates are more strongly correlated with visibility than with perceived quality.
27. Citation indicators of quality of output were not collected, since that would have meant collecting data for each faculty member at every medical school. Publication counts could be more conveniently obtained because the Source Index of SCI lists all publications by the affiliation of the authors as well as by authors' last names. Thus the count for productivity of schools was easily obtained. Adequate controls for the impact of scientific papers produced by faculty at the medical schools would require the count of citations for individual faculty members and the aggregation of these individual scores. On the correlation between productivity counts and citation counts, see Cole and Cole (c).
28. Productivity counts for clinical faculty vary greatly among and within the different schools.
29. If the actual scores of these southern schools were closer to those predicted by their faculty productivity, the span between highest and lowest scores would be further reduced. A comparison of medicine and other professions in terms of the distance between high- and lower-ranked schools should be done.
30. Five dummy variables were constructed, in which a region was given a value of one and all other regions were scored zero. In the regression equations we omit one dummy variable in each equation. Cartter's (Appendix E) categorization of a school's geographic region was used.
31. The age of medical schools is to some extent related to the eminence of its graduates. Clearly, medical schools that only recently opened could not have many graduates who are deans or in AAP.
32. The maximum amount of variance explained on perceived quality by a subset of the 37 school characteristics is .947, with a total of 13 independent variables; and .999 with 16 variables in a step-wise

regression equation. These prediction equations did not include either the school's visibility or its rated effectiveness in training students. Of course, as the explained variance approaches 1.0, the beta weights become unstable. The variables included in the 16-variable regression equation are: publications; total NIH funds; year of founding of the school; number of department chairmen and deans; percentage of graduates who are not a member of any specialty board; percent of first-year class completing four years of college; total HEW funds; number of Fellows, post-doctoral and graduate students enrolled in the clinical sciences; total number of graduates; total number of graduates alive as of December 31, 1967; number of "elites" per graduate; percentage of graduates who are not federally employed; total number of full-time faculty; HEW research and development funds; total number of students; geographic region: East; not East.

Since we have a small number of cases we corrected the total R^2 explained for degrees of freedom. This correction is obtained with the following equation (McNemar).

$$R = 1 - [(1 - R^2)(N - 1)] / (N - n)$$

where: N = number of cases;
n = number of variables, independent and dependent.

The sample size does not significantly affect the R^2. For the six-variable regression discussed in the text, the adjusted R^2 is .79; for the 16-variable equation, it is .9988.

33. For impact of the evaluations as primary sources, we examine works that refer to the authors of the studies in the Science Citation Index; for a more extensive measure of impact we would count citations to works that have made extensive use of the ratings.

34. Standardized criteria for evaluation plainly requires more research. In fact, consensus on standardized evaluation criteria is a problem faced in many settings, but is particularly significant in the process of refereeing journal articles and in peer review decisions.

REFERENCES

American Association of Medical Colleges. 1972. *Directory of American Medical Educators*, 1972–73. Washington, D.C.

Association of American Physicians. 1972. *Transactions*, Volume 95: 85th Session.

Barber, B. 1975. "The Limits of Equality: Social Stratification in Complex Societies." Paper presented for a conference, More Equality as a Goal of Public Policy. Princeton: Institute for Advanced Study.

Barber, B., J. J. Lally, J. L. Makarushka, D. Sullivan. 1973. *Research on Human Subjects*. New York: Russell Sage.

Blau, P. M., and R. Z. Margulies. 1974. "The Reputation of American Professional Schools." *Change* 6(December):42–7.

Cartter, A. 1966. *An Assessment of Quality in Graduate Education*. Washington: American Council on Education.

Cole, S., and J. R. Cole. a:1967. "Scientific Output and Recognition: A Study in the Operation of the Reward System in Science." *American Sociological Review* 32(June):377–90.

———. b:1968. "Visibility and the Structural Bases of Awareness of Scientific Research." *American Sociological Review* 33(June):397–413.

Cole, J. R., and S. Cole. c:1973. *Social Stratification in Science*. Chicago: University of Chicago Press.

Council on Medical Education, American Medical Association. 1972. "Medical Education in the United States, 1971–1972." *Journal of American Medical Association* (November):921–1081.

Crane, D. 1965. "Scientists at Major and Minor Universities: A Study of Productivity and Recognition." *American Sociological Review* 30(October):699–714.

Dube, W. 1973. "Study of U.S. Medical School Applications, 1971–1972." *Journal Medical Education* 48:395–403.

Elton, C. F., and S. A. Rodgers. 1971. "Physics Department Ratings: Another Evaluation." *Science* 174(November 5)565–68.

Gaston, J. 1973. *Originality and Competition in Science*. Chicago: University of Chicago Press.

Giza, R., and C. Burns. 1971. "DHEW Obligations to Medical Schools, Fiscal Years 1967–1969." Washington: Department of Health, Education and Welfare.

Hagstrom, W. 1971. "Inputs, Outputs and the Prestige of the American University Science Departments." *Sociology of Education* 44(Fall):375–97.

Hughes, R. M. 1925. *A Study of the Graduate Schools of America*. Oxford: Miami University.

Institute for Scientific Information. 1973. *Science Citation Index, 1972*.

Keniston, H. 1959. "Appendix: Standing of American Graduate Departments in the Arts and Sciences." In *Graduate Study in the Arts and Sciences at the University of Pennsylvania*. Philadelphia: University of Pennsylvania.

Knudsen, D. D., and T. R. Vaughan. 1969. "Quality in Graduate Education: A Reevaluation of the Rankings of Sociology Departments in the Cartter Report." *American Sociologist* 4(February):12–19.

Lazarsfeld, P. F. 1958. "Evidence and Inference in Social Research." *Daedalus* 87:99–130.

Lewis, L. S. 1968. "On Subjective and Objective Rankings of Sociology Departments." *American Sociologist* 3(May):129–31.

Magoun, H. W. 1966. "The Cartter Report on Quality in Graduate Education." *Journal of Higher Education* 37(December):481–92.

Margulies, R. A., and P. M. Blau. 1973. "America's Leading Professional Schools." *Change* 5(November):21–7.

McNemar, Q. 1969. *Psychological Statistics*. 4th ed. New York: Wiley.

Merton, R. K. a:1960. " 'Recognition' and 'Excellence': Instructive Ambiguities." In N. Storer (ed.), *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press, 1973.

———. b:1968. *Social Theory and Social Structure*. New York: Macmillan.

———. c:1972. "The Perspectives of Insiders and Outsiders." In N. Storer (ed.), *The Sociology of Science: Theoretical and Empirical Investigations*. Chicago: University of Chicago Press, 1973.

Roose, K. D., and C. J. Andersen. 1970. *A Rating of Graduate Programs*. Washington: American Council on Education.

Singletary, O. (ed.). 1972. *American Universities and Colleges*, 11th ed.

Theodore, C. 1968. *Medical School Alumni, 1967*. Chicago: American Medical Association.

Treiman, D. J. Forthcoming. *Occupational Prestige in Comparative Perspective*. New York: Academic Press.

Wilson, L. 1942. *The Academic Man: A Study in the Sociology of a Profession*. New York: Oxford University Press.

Zuckerman, H. Forthcoming. *The Scientific Elite*. New York: Free Press.