Evaluating Journalism & Mass Communication Education:

Are U.S. Efforts Applicable Globally?

By

Lee B. Becker

James. M. Cox Jr. Center for International Mass Communication Training and Research
College of Journalism & Mass Communication
University of Georgia
Athens, GA 30602-3018, USA
(Phone 706 542-5023; Fax 706 542-5036; E-mail lbbecker@arches.uga.edu)

&

Gerald M. Kosicki

School of Journalism and Communication
Ohio State University, Columbus, OH 43210-1107, USA
(Phone 614 292-9237; Fax 614 292-3809; E-mail kosicki.1@osu.edu)


ABSTRACT: Government agencies, educational policy planners, students and prospective students, and media industries are among the many groups in society today seeking accountability of institutions of higher education. All want to know if the educational institutions are doing their jobs; many want to know which ones are doing those jobs better than others.

Answers to these questions depend on a central issue: measurement of university and specific program performance. How can one know which universities and which programs within those universities are the best? How can performance be observed, measured and quantified so others can understand the standards and the criteria used to evaluate them?

In the United States, commercial firms, educational leaders, and media leaders have developed different procedures for the assessment of quality in journalism and mass communication education. The Gourman Report and U.S. News & World Report are two commercial organizations that have evaluated journalism education in recent years. Another type of evaluation is provided by the Accrediting Council on Education in Journalism and Mass Communications (ACEJMC). Academics in the field of educational research have developed yet additional approaches to program evaluation, including of journalism and mass communication programs at the university level.
Each of these evaluators applies its own methods in reaching its conclusions about which are the best universities and university programs in the country. In some cases, the methods overlap. Others are unique. None employs the full range of evaluation possibilities.

This report provides a methodological and conceptual overview of the current state of evaluation in the field of journalism and mass communication education. It makes empirical comparisons of the outcomes of various techniques and provides a critical assessment of the alternatives used. It also draws on experiences of other fields and on general issues in measurement of performance in higher education. It offers a critique of the current system and makes suggestions for the future.

The study draws on the general literature from outside the U.S. and offers a critical assessment of the U.S. experience with journalism and mass communication program assessment. Particular attention is given to the U.S. efforts to set a single standard and accredit programs accordingly.

I. Academic Literature on Evaluations

Educators who study quality in higher education approach the issues from at least three distinct, but related perspectives. Some focus on indicators of quality and their measurement. Another group concerns itself with the program review process. A third group is interested in what are termed performance indicators in higher education.

Quality Indicators

George D. Kuh, a professor of education at Indiana University, reviewed and summarized the academic literature on indicants of the quality of the undergraduate experience in a monograph published by the American Association for Higher Education. Kuh identifies four distinct criteria--those related to context, input, involvement and outcome. These are shown in Table 1.

Context, or environmental factors, include institutional size. Generally, size is a negative indicant of quality, both at the departmental and university level, Kuh argues. Another context factor is institutional purpose. Narrow and clear statements are treated as indicants of quality. Faculty and student morale are indicators of quality. Kuh also advocates assessing financial resources, arguing that use rather than amount, is important. He also advocates developing measures of the quality of the students' living environment.

Input factors include measures of the ability of the students, such as rank in high school and entrance test scores, and what Kuh labels biographical characteristics. Diversity is considered an indicant of quality. Kuh also treats what he calls "nonintellective" characteristics, such as student participation in high school extracurricular activities, student aspirations and student interests, indicators of quality.

Involvement factors include student participation in extracurricular activities. Various measures of instructional quality suggest themselves, such as a match between student need and instructional style. The amount of informal faculty-student interaction is important, as is the quality of student involvement in instruction.

Outcome factors including persistence, measured by the rate of retention, and achievement, measured by gains in substantive knowledge by the students, are suggested. Grades can be one possible indicant of
achievement. Kuh suggests use of standardized indices of intellectual and social emotional development as another outcome indicator. Employment and compensation of graduates are indicants of the quality of the educational experience.

Kuh argues that both quantitative and qualitative techniques should be used in developing the measures of quality.

**Academic Program Reviews**

Related to quality indicators is academic program review. Conrad and Wilson conclude that the high level of current interest in academic program review, including various forms of accrediting, is a response to fiscal constraints and to external constituencies' expectations for accountability of institutions of higher education. Program review is concerned with quality, and Conrad and Wilson offer four distinct perspectives on how quality should be defined; (1) the reputational view, (2) the resources view, (3) the outcomes view, and (4) the value-added view.

The reputational view, Conrad and Wilson conclude, assumes that quality cannot be measured directly and is best inferred through the judgments of experts in the field. The resources perspective focuses on the human, financial and physical assets available to programs. The assumption is that quality is present when an institution has excellent students, productive and highly qualified faculty and modern facilities and equipment. Proponents of the outcome perspective argue that the product of the institution should be given prominence. Faculty publications, student accomplishment after graduation and employers' satisfaction with graduates are seen as indicators of quality. The value-added view focuses on what the institution has contributed to the student's education. The concern is with what the student has learned, and programs are evaluated highly if they can demonstrate increased knowledge and personal development.

Measurement of quality, regardless of perspective, is difficult, Conrad and Wilson note. "The problem is that no one has yet found a way to measure quality directly," they write (p. 50). The measurement assumption of the reputational perspective, Conrad and Wilson say, is that "the optimum way to assess quality is to seek a consensus of informed opinion" (p.51). Typically, opinion is assessed through a survey. The strength of the method is that the raters are supposedly those who know best what quality is. Weaknesses of the approach, according to Conrad and Wilson, are that they focus on scholarly productivity and reputation, rather than the instructional program, and that reputations can be out of date.

The assumption of those holding the resources view is that quality exists where the "resources" of bright students, excellent faculty, adequate budgets, strong research support, strong libraries and good facilities are present. Actual measures include student test scores, proportion of the faculty with a doctorate, grant support, and the number of volumes in the library. Conrad and Wilson note that the advantage of such an approach is that such measures are readily available at most institutions, that the measures are current and that comparisons can be made across institutions. There is little research evidence, however, that resources equate with student learning, according to Conrad and Wilson. There also is a false sense that there is some ceiling to quality, i.e., that quality peaks out when all of the faculty have doctorates or when all of the students are at the top of the range in test scores.

The outcomes perspective assumes that the proper way to measure quality is through the results of the program. Typical indicators are faculty productivity, students' accomplishments following graduation, employer satisfaction with graduates, and institutional contributions to the solution of local, state or national problems. Specific measures include the number of faculty publications in scholarly journals,
the number of graduates admitted to leading graduate or professional schools, employer surveys, percentage of graduates finding employment soon after graduation, and lifetime earnings of graduates.

An advantage of the approach is that many of the measures can be applied across institutions. Disadvantages, according to Conrad and Wilson, include the difficulty of determining how much of the outcome is attributable to the institution itself, rather than such things as the quality of the students admitted to the program. The measures also focus more on the past than on current circumstances at the institution.

The value-added perspective holds that quality is reflected in the institution's ability to affect its students in a favorable way by making a difference in their intellectual and personal development. Evaluation should focus on what students learn while enrolled in a program. A typical measure, according to Conrad and Wilson, is achievement tests at the time of enrollment and the time of graduation. The chief advantage of this approach, the authors say, is that it takes account of the quality of students upon entry. Institutions are judged by how much they help students, regardless of where the students were when they entered. Among the limitations are expense of administration of such tests, and difficulty in reaching consensus on what a student should learn and how it should be measured. Finally, the effects of other factors, such as maturation, employment, and other outside forces cannot easily be controlled.

Table 2 summarizes the observations by Conrad and Wilson about program evaluation.

**Performance Indicators**

In a 1994 review of the literature on performance indicators, Gaither, Nedwek and Neal note an increasing demand for quantifiable evidence that higher education is worth the cost and time involved. They attribute this increasing demand to social and economic change in society. First, they argue, higher education has become more important in landing and retaining good-paying jobs. Second, consumers generally have become more sophisticated and less willing to accept the effectiveness of higher education on faith. Third, the cost of higher education has increased greatly, and people want to know where they should spend their money.

What society and its members want are performance indicators, that is, ways to judge the quality of the higher education process. Gaither, Nedwek and Neal identify three major types of such indicators, those that focus on results, those that focus on inputs, and those that focus on the process itself. These are shown in Table 3.

Results indicators include the percentage of students graduating, the number of degrees granted, the percentage of undergraduate students going on to graduate school, and the number of publications by faculty. Results measures that focus on quality versus mere quantity might include the percentage of students going to the best graduate schools, the average starting salary of the graduates in relation to the salaries of graduates of other schools, or the number of faculty publications cited in the publications of other scholars.

Input indicators include entrance exam scores for incoming students, the rank or class of the enrolling students, the number of merit scholars, the rank of faculty members' graduate degree programs, the research grants of incoming faculty, or the number of refereed publications of newly hired faculty. The presumption is that if the quality of input is high, the quality of the output also will be high.

Process indicators include the frequency with which students use the library, participate in study groups, hold positions of leadership, write papers, or consult with their academic advisers. Process indicators
also would include the number of hours faculty used to prepare for class, their use of new technology in teaching, or how they validate their assessments of students.

There are many pitfalls in using indicators of quality, Gaither, Nedwek and Neal note. One commonly used indicator of quality is the percentage of lower-level courses taught by senior faculty. There is little research evidence, however, that senior faculty are, in fact, better teachers. The assumption that more learning takes place in smaller classes also is not supported by the literature, though students do believe they learn more when they are in smaller classes.

Gaither, Nedwek and Neal recommend the use of multiple indicators of quality and offer the standards of reliability and validity as a means of evaluating the indicators used. Reliability is a question of repeatability or replicability. A reliable measure does not vary across time or between measurers unless real change has taken place. Validity is a question of the link between the indicator and the underlying concept of quality. If classroom size is not related to learning, it cannot be used as an indicator of learning and, by extension, quality of instruction. "Bad data are worse than no data when indicators become the primary tool for managing higher education or allocating resources," the authors write (p. 22).

**National Research Council Graduate Rankings**

The 1995 release of the National Research Council (NRC) 1993 study of Research-Doctorate Programs in the United States caused quite a bit of reaction around the country, particularly at the major research universities. The 1993 study was actually an update of a 1982 study by NRC, and employed a similar methodology. The central purpose of both studies was "to assess the quality of individual doctoral programs in terms of their effectiveness in preparing graduates for careers in research and scholarship..." (p. viii). The NRC study is a concrete example of evaluation carried out within the academic community itself.

NRC, the working arm of the National Academy of Sciences and the National Academy of Engineering, examined doctoral programs in 41 fields. Included were 11 fields in Arts and Humanities, seven in the biological sciences, eight in engineering, eight in the physical sciences and mathematics, and seven in the social and behavioral sciences. All 32 fields included in the 1982 survey were included in the 1993 survey. New fields were added if the field awarded a minimum of about 500 degrees in about 50 programs for the years 1986 to 1990. Programs in communication or mass communication were not included either in 1982 or in the 1993 study.

The NRC study included two types of data, reputational evaluations and descriptive data on characteristics of the evaluated doctorate programs. Three hundred universities were identified as having at least one doctoral program actively producing doctoral students in one of the 41 fields selected for study. Those 300 universities were asked to participate in the NRC study; 284 agreed to do so. Working with a coordinator, usually the graduate dean, at each of these 284 institutions, the NRC obtained faculty lists for the program at each university eligible for inclusion in the study. The listed faculty were included in the reputational survey, termed the National Survey of Graduate Faculty, and used to generate statistics about faculty research and publication activities. A total of 3,634 programs in the 41 selected field at the 274 universities ultimately became a part of the NRC study.

The NRC conducted its reputational survey, or National Survey of Graduate Faculty, via questionnaires mailed to a probability sample of faculty in the selected fields in the Spring of 1993. Respondents were presented with a list of approximately 50 programs, randomly selected for each individual respondent, and asked to rate the "scholarly quality of (the) program faculty" on a six-point scale from

http://www.grady.uga.edu/annualsurveys/IAMCR98/iamcr98a.htm
"Distinguished" to "Not sufficient for doctoral education" and to rate the "effectiveness of (the) program in educating research scholars/scientists on a four-point scale from "Extremely effective" to "Not effective." Respondents also were asked to indicate the level of their familiarity with the program faculty and with graduates of the program and could indicate they did not know the program faculty or graduates well enough to rate a particular program. Respondents were asked to respond to questions about their own background. Each program evaluated appeared on at least 200 of the questionnaires. The NRC set as its goal a total of at least 100 ratings per evaluated program. In the end, the number of raters varied from a low of 82 in Comparative Literature to a high of 424 in Biochemistry and Molecular Biology. A total of 7,926 respondents returned questionnaires, or 51% of those considered to be eligible to do so. Return rates varied from a low of 38% in Molecular and General Genetics to a high of 64%, achieved in both Materials Science and in Geography.

From the two survey questions—one on program faculty and the other on program effectiveness—the NRC created two separate reputational measures. The former (faculty quality) has been shown to correlate with research and publications of faculty, according to the NRC report (p. 22). The effectiveness measure "is believed by many to correlate closely with the career outcomes of program graduates," though this "hypothesis" was not examined by the NRC. The NRC report acknowledges several limitations of reputational rankings. First, such rankings are related to program size. Second, while the rankings do correlate with level of faculty involvement in research and scholarly activities, they do not necessarily indicate anything about program structure, the environment for the students, or the actual success of the students in finding jobs upon graduation. Reputational rankings of programs are influenced by the evaluation of a prominent individual in a program and by the prominence of the university in other fields. The NRC also acknowledges that the rankings do not indicate the quality of faculty teaching or of faculty involvement in other service to the university or the community.

The NRC supplemented the reputational data with measures of program characteristics. Included were data on library holdings and expenditures, number of faculty in a program, the percentage of faculty holding the rank of professor, percentage of faculty with research support, the percentage of faculty receiving an honor or award, percentage of faculty publishing in the period 1988 to 1992, an index of citations to work of faculty in a program, the number of students, the percentage of students who were female, the number of number of degrees granted, the percentages of degrees awarded from July 1986 to June 1992 to women and to minorities, and the median time elapsed between entering the doctoral program and earning the doctorate.

The NRC provided extensive detail of the data obtained by field. The data reported for Political Science, for example, showed that Harvard had the highest reputational score for faculty quality among the 98 evaluated programs and ranked third in terms of the effectiveness measure. The correlation coefficient (a statistic that varies between -1.00 and +1.00, with the former indicating complete and perfect disagreement and the latter perfect agreement) between the score on faculty quality and the score on effectiveness was +.97 for all of the 98 programs. Both measures correlate highly (+.74 for faculty quality and +.69 for effectiveness) with the size of the program, measured by number of faculty. The measures correlate more moderately with several measures of faculty research productivity and with size of the program as measured by the number of graduate students enrolled. For example, faculty quality correlated +.52 with the percentage of faculty in the program with federal research support in the 1988-92 period and +.52 with the percentage of faculty publishing in the same time period. The correlation also was +.52 with number of students in the program and +.59 with number of degrees granted. Faculty quality correlated not at all with the percentage of degrees granted to women (+.03) and slightly negatively with the percentage of degrees granted to minorities (-.28). There also was a slight negative relationship between median years required to earn a degree and faculty quality (-.24). There is variability among the fields on these relationships. In sociology, for example, the relationship between faculty quality and faculty size is +.59. The correlation between faculty quality and the percentage of
faculty with federal research support is +.46. There is no relationship (-.01) between faculty quality and percentage of degrees granted to minorities. What determines this variability is not known at this point.

II. Commercial Evaluations

Several individuals or organizations have recognized the commercial value of evaluations of academic programs. These individuals or groups gather data and make evaluations that are then reported in publications to be sold to prospective students and their parents. These evaluations are, for the most part, applications of the principles discussed in the academic literature on evaluation.

Gourman

Of the commercial rating services, The Gourman Report may be most famous. Authored by Dr. Jack Gourman, The Gourman Report, A Rating of Undergraduate Programs in American & International Universities, is now in its ninth edition. The companion report on graduate education is in its seventh edition. The volume on undergraduate programs contain ratings of Journalism and Mass Communications programs and Communication programs. The graduate volume contains ratings of Journalism programs.

The undergraduate volume contains only an outline of the methodology used in the evaluations. Educators and administrators at the institutions being evaluated provide information to Gourman. In addition, Gourman uses records on funding, university reports on faculty makeup, fields of study, and physical plant. Finally, Gourman relies on "individuals, associations and agencies whose business it is to make correct projections of the success graduates from given institutions and disciplines will enjoy in the 'real world'" (page 2).

The Gourman Report lists 18 criteria used in reaching its final rating of undergraduate programs. These are shown in Table 4. How these data are used to actually create a numeric ratings of from 2 to 5, shown in the report, is not specified. In the preface to the seventh edition of the report, Gourman also lists a series of "shortcomings" that lead to low scores. These are shown in Table 5. The graduate volume contains no independent listing of criteria used to evaluate programs, though it, too, lists in the current edition (ninth) preface "shortcomings" that lead to low scores (Table 6).

Table 7 summarizes the Gourman journalism ratings in the most recent volumes. The schools listed are those that Gourman considers to have the top undergraduate journalism and mass communication or top graduate journalism programs. Shown are the ranks and the raw scores. In addition, the overall rank for the university is shown, if it is the top 50 undergraduate programs, according to Gourman. Only the top 50 universities are ranked overall. Finally, the table shows the overall score for the university, as reported in the undergraduate volume. As noted above, these scores range from 2-5.

Gourman ranks five graduate programs on his list of the top 22 that either do not have undergraduate programs (Stanford, Michigan and Columbia) or do not have undergraduate programs Gourman considered to be among the top 31 undergraduate journalism programs. Even among the 17 universities with both graduate and undergraduate programs ranked, the relationship between the underlying scores of these two programs is weak (Pearson R of +.40). The Syracuse undergraduate program illustrates the lack of a strong relationship. It is ranked as third best in the undergraduate Gourman listing with a score of 4.82, but 17th in the graduate listing with a score of 4.49.

In fact, there is a cluster of four schools out of the 17 ranked in the undergraduate and graduate rankings that cluster together as anomalies, Figure 1 shows. (A low score represents a high rank in the chart.)
These schools, shown in the upper, left-hand side of the chart, are Syracuse, North Carolina, Ohio State and Michigan State. Otherwise, there appears to be a clear relationship between undergraduate and graduate rankings.

The relationship between the score of the 31 undergraduate journalism and mass communication programs and the overall university score is stronger (Pearson R of +.61). Still, there are significant disparities. Missouri and Syracuse have the number 1 and 3 undergraduate journalism programs respectively, according to Gourman, yet those universities are not even ranked among the 50 undergraduate universities by Gourman. The 13 schools ranked in the top 50 overall by Gourman and ranked in the journalism listing are shown in Figure 2. There is quite a spread to these rankings. Clearly it isn't necessary to have a topnotch university to have a top ranked journalism program, using the Gourman criteria.

The relationship between the rating of the 22 graduate journalism programs and the rating of the undergraduate score for the university is nearly the same as the relationship between the rating of the undergraduate journalism programs and the undergraduate university score (Pearson R of +.57). Gourman reports no overall university graduate rankings or scores.

The scatterplot of the rankings of the 15 universities that Gourman considers to be among the top 50 in the country and which also were ranked in terms of their graduate journalism programs is shown in Figure 3. Once again, there is quite a spread to the scores.

**U.S. News & World Report**

Gourman is alone among the rating services in providing evaluations of undergraduate journalism programs. *U.S. News & World Report* for the first time in 1996 evaluated graduate journalism programs. (9)

*U.S. News* fielded a reputational survey in late 1995 to create its rankings of the graduate journalism programs. The magazine sent a questionnaire to "340 deans and leading faculty" at the 170 schools identified through the Annual Survey of Journalism & Mass Communication Enrollments. (10) How the "leading faculty" were selected was not specified. Nor was it specified that one was selected from each of the 170 schools, though that seems to have been the case. The term "dean" seems to have been used generically to refer to unit head.

The *U.S. News* survey actually contained two distinct components. The questionnaire asked the evaluators to "rate each school according to one of five levels of academic quality." The categories were: Distinguished (5), Strong (4), Good (3), Adequate (2) and Marginal (1). The respondents were encouraged to mark "Don't Know" if they were "unfamiliar" with the program. The person completing the survey was then asked to list the "10 best graduate degree programs" in each of four "fields of Journalism and Mass Communications." These were: Print Journalism, Radio/TV, Public Relations and Advertising. The programs were not to be ranked, but merely listed. Fewer than 10 programs could be listed. The selection was to be "based on your opinion of the school's academic excellence, strength of faculty and research achievement in each specific field."

The magazine conducted a separate reputational survey of "a random sample of 190 print journalists, 180 broadcast journalists, 150 public-relations executives and 150 advertising executives."

The magazine reported that 40% of the 370 educators surveyed returned the instrument. It did not specify how many universities were represented by those 148 individuals. Only 13% of the 670
communication professionals responded.

Some of the results of these two surveys were reported in the March 18, 1996, issue of *U.S. News & World Report*. The article and some additional editorial material not drawn from the surveys appeared in the 1996 Edition of the magazine's publication, *Best Graduate Schools*. The magazine presented a rank of 15 graduate programs in each of the four areas, advertising, print journalism, public relations and radio/television, based on the educator responses. It did not release the results of the overall rankings and has refused to release those data subsequently. It listed, but did not rank, five programs in each of the four areas based on the responses of the 87 communication professionals, arguing that the low response rate did not "permit statistically representative rankings." Such a deficiency, of course, would make reporting of any data--not just rankings--suspect.

**Table 8** shows the rankings and listings from the *U.S. News* report. Six of the schools are ranked in the top 15 in all four areas. These are the University of Florida, Northwestern University, the University of Texas at Austin, the University of Missouri at Columbia, Syracuse University, and the University of Wisconsin at Madison.

Others get more specialized recognition. For example, the University of Illinois is ranked only in Advertising (where it is at the top of the list). Three other universities also only have a top 15 ranking in advertising: The University of Tennessee at Knoxville, the University of Colorado at Boulder, the University of Nebraska at Lincoln. The University of California at Berkeley, the University of Kansas, and the University of Minnesota received top scores only in Print Journalism. San Diego State, California State at Fullerton, and Boston University received top 15 rankings only in Public Relations. The University of Southern California, Southern Illinois University and Temple University get top rankings only in Radio/Television.

The few working communication professionals who responded listed five Advertising programs from the top 10 on the educators' list, five out of the top 10 in Print Journalism, three of the top 10 in Public Relations, and four of the top 10 in Radio/Television. The fifth Radio/Television program was at Stanford, which was not listed by educators and which administrators at Stanford report does not exist at all.

**Table 9** shows the *U.S. News* rankings and the most recent graduate rankings from *The Gourman Report*. *The Gourman Report* does not differentiate programs by speciality and should be equivalent to the overall evaluation obtained in the *U.S. News* survey but not reported. Gourman's top 22 graduate programs include seven universities not listed on any of the top 15 lists. Four of these were given a rank of 15 or lower by Gourman.

To provide summary statistics comparing the rank data in **Table 7** with the Gourman data, we selected only the 29 universities ranked by *U.S. News* in **Table 7**. In those cases where a university did not get a rank, it was assigned the rank of 22, or halfway between 15 and 29. The Gourman ranks for those same 29 universities were recomputed by eliminating missing ranks (7, 8, 11 and 13). This, too, produced ranks for 15 universities. The remaining were assigned a rank of 22.

**Table 10** shows the Kendall's *tau b* correlation coefficient for these measures. The coefficient can vary between 0 and 1, and, in fact, most relationships shown in **Table 9** are slight. Advertising and Public Relations programs get somewhat similar rankings in the *U.S. News* reputational analysis. Print Journalism and R/TV programs also get somewhat similar evaluations by the educators. The Gourman rankings are most similar to the print rankings and, to a lesser extent, the R/TV rankings. Gourman is
not related in any appreciable way to the *U.S. News* Advertising and PR rankings.

In reputational rankings the graduate experiences of the evaluator can be important. It isn't possible to know from the *U.S. News* reports what role graduate degree may have played. What is clear, however, is that only nine of the 30 programs listed in Table 7 do not offer a doctorate. Only one of the six universities ranked in the top 15 on all four criteria does not have a doctoral program: Northwestern. Two historically important doctoral granting programs--Wisconsin and Minnesota--both have made the list, the former ranking in the top 15 in all four areas.

The *U.S. News* rankings were quite controversial, with the Stanford ranking by professionals a particularly glaring mistake. The low response rate from educators (40%) and professionals (15%) make inferences from the data problematic. For 1997, *U.S. News* dropped journalism from the graduate ratings.

The *U.S. News* rankings of graduate programs in advertising are quite consistent with those obtained in a separate reputational study by Richards and Taylor, conducted in 1993 but also published in 1996. Richards and Taylor sent questionnaires by mail to the program or sequence head for each of the 143 universities offering advertising education in some form, based on the Ross and Johnson study of advertising education. Usable responses were obtained from 68 schools, representing a response rate of 48%. Respondents were asked to list, in order, the three best advertising programs at the undergraduate level and the three best advertising programs at the graduate level. In addition, respondents were asked to list the single program the respondent "would most like to emulate." Respondents were not allowed to list their own program in any of the responses.

The data from the Richards and Taylor survey are shown in Table 11. Ranks in the case of undergraduate and graduate program evaluations are based on the weighted score, which assigned a value of three for a top ranking, a value of two for a second ranking, and a value of one for a third place ranking. Also shown are raw number of mentions.

Michigan State and the University of Illinois at Champaign-Urbana received the top ranking at the undergraduate level. The University of Texas followed closely. At the graduate level, Northwestern University received the top ranking in the Richards and Taylor survey. Next was the University of Illinois and the University of Texas at Austin. The *U.S. News* rankings of graduate programs in advertising from 1996 are shown in the Table 11 as well. The *U.S. News* survey was conducted in late 1995, before the Richards and Taylor data were published. There is a high level of agreement, nonetheless, in the responses to the two surveys. Illinois is top-ranked in the *U.S. News* survey, and Northwestern is rated number 3. The Spearman Rank Order Correlation Coefficient (rho) comparing the rankings of the Richards and Taylor and *U.S. News* is +.74 (out of a possible +1.00). By comparison, the Spearman Rho between the Richards and Taylor undergraduate rankings and the *U.S. News* graduate rankings was a more modest +.43. Within the Richards and Taylor data there is a discrepancy between undergraduate and graduate evaluations. The Pearson Product Moment Correlation Coefficient between undergraduate and graduate scores for the nine programs mentioned in both is +.63 (out of a possible 1.00). Number of mentions to the question on emulation was related more highly to mentions at the graduate level (Pearson coefficient of +.57) than at the undergraduate level (Pearson coefficient of +.25). In other words, respondents were more likely to want to emulate the programs evaluated highly for their graduate programs than those evaluated highly for their undergraduate programs.

The ranking procedures that had been used by *U.S. News* for journalism in 1996 contrasted in their simplicity with those used for other fields. Law school rankings, for example, were based on student selectivity, placement success, faculty resources, and reputation. Student selection was measured by
median undergraduate grade point average (weighted at 40%), median Law School Admission Test Score (50%), and proportion of applicants accepted (10%). Placement success was based on proportion of the graduating class employed full and part time in legal and nonlegal jobs at graduation (30%), proportion employed full and part time in legal and nonlegal jobs six months after graduation (60%), and the median starting salary of graduates employed in the private sector (10%). Faculty resources was based on total expenditures per student for instructor, library and supporting student services (65%), financial aid, indirect expenditures and overhead per student (10%), total number of volumes, microfilm, microfiche and titles in the law library (5%), and current student-to-teacher ratio (20%). Reputation ranking was based on two surveys, one of four deans and faculty members per school, the other of 2,167 practicing lawyers, hiring partners and senior judges. Results of the two surveys were considered separately. Overall rank was determined by converting into percentiles the scores achieved by each school in the five categories. The highest score in each category was given a value of 100%. The scores for all of the other schools were converted to percentiles by comparing them to the score achievements by the top school. Law schools were then ranked in each of the five categories. The number ranks were weighted as follows: academic reputation and student selectivity (25% each); placement success (20%), faculty resources and nonacademic reputational ratings (15% each). Final rankings were determined by totaling the five weighted number-ranked scores. The weighted score for the top school was given a value of 100%, and scores for the others were determined as percentage of the score of the top school.

Business schools were evaluated on student selectivity, placement success, and two reputational surveys. Medical schools were evaluated on student selectivity, faculty resources, two reputational surveys, research activity (total National Institutes of Health research grants), and Primary-care rate (average percentage of graduates going into primary care residencies). Health professions were rated via a survey of deans and top officials of accredited programs. (Response rates varied from a high of 77% in dentistry to a low of 52% for social work.) Education schools were evaluated on student selectivity, faculty resources, research activity, and reputation (via two surveys). Master's level programs in library science, film, drama/theater, architecture, music and fine arts were evaluated via a single questionnaire sent to deans, top administrators and senior faculty of the accredited schools in those disciplines. (Response rates ranged from 60% for architecture to 34% in the fine arts).

Doctoral programs in the 12 disciplines with the largest graduate enrollments in doctoral programs also are ranked. Questionnaires were sent to the department heads and directors of graduate studies at schools granting at least five degrees in the field over the last five years. Respondents were asked to evaluate the programs based on their reputation for scholarship, curriculum and the quality of the faculty and graduate students. Response rates varied from 52 to 30%.

U.S. News also rates undergraduate programs in a Best Colleges volume. The beginning point for the evaluations is the classification of the 1,422 accredited four-year schools by the Carnegie Foundation by the Advancement of Teaching. The magazine mailed surveys to college presidents, deans and admission directors at these institutions. Those receiving the survey instrument were to rate "by quartile" all schools in the same Carnegie category as their own. (U.S. News reports that it simplified the Carnegie classification by combining some categories and subdividing others by region.) "Some 65%" of those receiving the surveys completed them, the 1997 Best Colleges directory reports.

In addition to the reputational rating, the magazine evaluates selectivity of the institution, faculty resources, financial resources, retention, something termed "valued added," and alumni giving. Selectivity was determined by acceptance rate of applicants, percent of those accepted who attended, SAT or ACT scores, and high school class standing of entering freshmen. Faculty resources included student/faculty ratios, percent of full-time faculty with doctorates or other terminal degrees, percent of faculty with part-time status, average salary for full-time faculty, and class size. Financial resources
included a school's total fiscal expenditure for instruction, student services and academic and instructional support divided by the number of students. Retention was based on the average percent of students who graduated within six years and the average percentage of freshmen who returned for the following autumn. Value added was determined by the difference between the expected graduation rate of the institution based on the test scores of its entering students and actual graduation rates. Alumni giving was determined by the average percent of a school's living alumni who contributed in an annual fund drive. The components were weighted, with reputation counting for 25%, selectivity counting for 15%, faculty resources counting for 20%, financial resources counting for 10%, retention counting as 20% for national schools and 25% for regional schools, value added counting as 5% for national schools, and alumni giving counting as 5% in the final score. Scores were assigned to each school in each of the eight areas. The top school received a score of 100 in each category. Other schools received scores representing a percentage of the top score. Final scores, after weighting, were similarly adjusted so the top score was 100 and other schools were ranked as a percentage of that score.

Table 12 shows some key data from the U.S. News Best Colleges 1997 volume for the national universities. Journalism education has prospered at large institutions, such as those fitting into this category. Northwestern and Columbia, both with prominent journalism programs (the latter only at the graduate level), are at the top of this list. Other universities with prominent journalism programs in the top 50 national universities, according to U.S. News, are North Carolina, University of California at Berkeley, New York University, Boston College, University of Wisconsin--Madison, University of Washington, Syracuse University and the University of Illinois. These 50 universities make up the top tier in the U.S. News rankings, having earned overall scores of from 100 to 71.2. The table also shows the reputational rankings for these top 50 national universities.

The Princeton Review

Another entry in the university review market is called The Princeton Review Student Advantage Guide to the Best Colleges.(18) The 1997 edition includes 310 colleges and universities.

The Princeton Review ratings, in part, are based on input from students. For the 1997 volume, more than 56,000 students at the 310 colleges and universities rated answered questions about those universities. "Except at some extremely small schools (undergraduate enrollment below 1,000)," responses were obtained from at least 100 students at each of the campuses covered. Selection of the students, however, was not probabilistic. The data come from students who enrolled in SAT and graduate test preparation courses operated by The Princeton Review.

The survey asked students to answer 70 questions on such things as social life, quality of food, and quality of teaching. Additional data were obtained from "independent college counselors," who were asked for comments on the school's academic departments, admissions and financial aid policies, and "hospitality" toward minorities. From these data, the Review developed a number of ratings, including one for academics. "Factors weighed included how many hours students studied, how vigorously they did assigned readings and attended all classes, and the quality of students the school attracts." Also considered were "the student/teacher ratio, the students' assessments of their professors' abilities and helpfulness, and the students' assessment of the school's administration."

The Princeton Review does not rate individual areas of speciality, such as journalism. The Princeton Review scores for universities, however, are theoretically comparable to the university scores of Gourman and U.S. News & World Report. Table 11 lists ranking data from these three sources. The table is constructed with the U.S. News data as the base, since U.S. News is more comprehensive than the other two rating sources.
Money Guide by Money Magazine

Money Magazine also ranks colleges in its annual Money Guide. Unlike the Gourman, U.S. News and Princeton Review guides, however, the Money Guide does not focus on academic strength of the rated universities directly. Rather, Money attempts to identify the "best buys," that is, the "schools that deliver the highest-quality education for the tuition and fees they charge" (p. 23).

The 1996 guide ranked 100 schools and provided data on another 949. The top ranked school was the New College of the University of South Florida in Sarasota, followed by Rice University (Texas), Northeast Missouri State University, Trenton State (N.J.), and the California Institute of Technology. None of these have journalism programs ranked by either U.S. News or Gourman. The University of North Carolina at Chapel Hill is ranked sixth by Money and appears on both the U.S. News and Gourman graduate and undergraduate lists. The University of Illinois at Champaign is ranked ninth by Money and also appears on the U.S. News and both Gourman lists.

Money uses 16 "measures of educational quality." These are listed in Table 13. These measures are combined in some unspecified fashion and "compared" with the college's "tuition and fees to arrive at a value rating." Out-of-state tuition is used where such a differential exists.

The measures shown in Table 13 help illuminate this publication's approach to educational evaluation and, to a certain degree, the basic approach of all the commercial services.

Comparison of Rankings

Table 14 shows the current (1997) university rankings for the three general commercial services. The services report rating scores differently, and full data are available for only a small number of the universities, since U.S. News reports raw scores only for the top 50 national universities. Of these 50 universities, The Princeton Review rated 49, and Gourman rated 48. As Table 12 showed, U.S. News ranked Yale as the top university, while Princeton assigned Yale a score of 93, lower than the score for several other universities. Gourman put Yale in the fourth position in its ranking. The University of Illinois at Urbana-Champaign gets a ranking of 50 by U.S. News, a score of 77 by Princeton, and a rank of 21 by Gourman.

One way to compare these various ratings and rankings is via a correlation coefficient. A correlation coefficient of 0 would mean there is no relationship between the evaluation by one of these three commercial services and the evaluation of another. The Pearson Correlation Coefficient between U.S. News rating scores and Princeton rating scores is a very strong +.91. In other words, these two sources produce very comparable assessments of the 49 universities. (Only the academic component of Princeton is being considered.) The coefficient comparing U.S. News rating scores and Gourman rating scores is a lower +.61, meaning there is significant but less agreement between these two sources.

On the other hand, the U.S. News reputational rankings alone, that is, that component of the overall score based solely on reputation and shown as a rank, correlates +.56 with the Princeton ranking for the 133 universities evaluated in common. The U.S. News reputational rank is correlated +.86 with the Gourman overall rating score for those 224 universities evaluated by both services. Among the 49 schools ranked by Gourman and for which there is a reputation rank for U.S. News, the correlation is a nearly comparable +.87. The Princeton rating and the Gourman rating for those 131 universities jointly evaluated is +.50.
The data suggest that Gourman includes a stronger reputational component in its evaluations than does either the overall *U.S. News* evaluation or the *Princeton Review* rating. The overall *U.S. News* score and the Princeton academic score, at least among those top schools, seem to be measuring much the same thing. These analyses are somewhat limited by the *U.S. News* procedure of only reporting data on the top 50 universities. Disagreements might well be higher with a larger set.

Another possible comparison is among the evaluations of the research doctorate programs by the National Research Council (NRC) and the evaluations of graduate programs in the same areas by *U.S. News* and by Gourman. All three sources rated the following programs: economics, English, history, political science, psychology and sociology. Since these are fields that are important to both the undergraduate and graduate curricula of journalism and mass communication, we selected them for analysis. The NRC rating is wholly reputation, as is the rating of these graduate programs by *U.S. News*. Gourman is offered as a rating that includes more than reputation. In each case, data are reported as raw scores.(22)

**Table 15** shows the Pearson Product Moment Correlation Coefficients for the programs rated by each of the three sources. The rating for 26 economics programs evaluated both by the NRC and by US News and World Report in 1996 is nearly identical, as indicated by the Pearson coefficient of +.96. The same can be said for the other fields as well. Only in the case of the 27 psychology programs is the relationship lower than +.90. The NRC ratings also match well with those of Gourman. For the 39 economics programs evaluated by both sources, the scores correlated +.98. The correlations in the five other fields also are quite high. Consistent with the analysis of undergraduate programs discussed above, the *U.S. News* reputational measures of graduate programs correlate highly with the ratings of those programs by Gourman. For economics, the correlation was +.91. The lowest correlation coefficient is still a high +.75 for the 25 programs in sociology evaluated by both.

The final two columns in **Table 15** show that there is great consistency in the evaluation of these programs from year to year by *U.S. News* and World Report. The relationship between actual raw ratings scores for reputation by *U.S. News* in 1996 correlates nearly perfectly with the raw ratings scores for reputation in 1997. In other words, the method being used by the magazine to obtain its ratings of graduate programs in these areas, at least, appears to be highly reliable or repeatable year-to-year.

**Standardized Data Gathering**

The *U.S. News & World Report* rankings and those of the other commercial sources have come under close scrutiny in recent years because of the suspicion that universities do not always provide accurate information to the evaluators.(23) *U.S. News* and other college guide books, including those such as Peterson's and the College Board guides that list schools without rating them, rely heavily on information provided directly by the colleges and universities. Officials at many universities admitted to *The Wall Street Journal* in 1995 that they reported false or partial data to improve their image and give them an advantage in recruiting students. Average SAT scores of entering freshmen were reported by several in a way to overstate the quality of students at the university.

*U.S. News* has joined with the College Board, Peterson's and Wintergreen/Orchard House, publishers of *College Admissions Data Handbook* and suppliers of data to *Money Magazine*, to standardize questions posted to colleges and universities and the definitions used in answering them.(24) The common data set and questions will be used in 1997 to gather data for next year's guidebooks.

*U.S. News* had to admit yet another problem with its 1997 Graduate School rankings. The magazine reported in its March 17, 1997, edition that the previous edition of the magazine containing the first
report of the graduate school rankings for 1997 contained a "miscalculation" leading to incorrect law school rankings.\(^{(25)}\) As a result, the magazine said, 44 of the law schools were misranked, including 33 in the top 50, and information in the table on placement success rate was incorrect for 159 of the 174 schools.

While \textit{U.S. News} provided no additional details on the error and how it occurred, \textit{The Chronicle of Higher Education} in its March 14, 1997, edition reported that two columns of numbers--those reporting unemployed students looking for work and those reporting unemployed students not looking for work--were transposed at some step in the analysis.\(^{(26)}\) Unemployed students not looking for work are not used by the magazine in figuring the job-placement success rate shown in the magazine.

\textit{U.S. News} reported in the March 17, 1997, edition that it had "called back all copies of the guidebooks with inaccurate data." The magazine rack from which we purchased the March 17 issue of the magazine, however, contained immediately to the right of the magazine itself a copy of the guidebook with the uncorrected rankings included.\(^{(27)}\)

The \textit{Chronicle} reported in the March 14, 1997, story that dental schools across the country, in protest against the \textit{U.S. News} rankings, organized an effort to boycott them. The American Association of Dental Schools urged its 54 members not to participate when the magazine sent out the questionnaires from which it sought to gain information. According to the \textit{Chronicle}, \textit{U.S. News} abandoned its plans to rank Dental Schools when "only about 35% of its surveys had been returned." Dental school deans, the paper reported, also sent letters complaining about the rating methodology.

The president of Alma College also has called for a boycott by college presidents and other administrators of the reputational rankings of liberal-arts colleges,\(^{(28)}\) and a group of Stanford students has led a movement by students at other universities across the country asking university administrators to withhold information from \textit{U.S. News} so it cannot complete its rankings.\(^{(29)}\) Stanford President Gerhard Casper announced in April of 1997 that Stanford was attempting to create an alternative to the \textit{U.S. News & World Report} rankings by asking interested colleges and universities to post relevant data on a World-Wide Web site.\(^{(30)}\) Stanford posted graduation rates, tuition and housing costs, class sizes, and data on faculty quality on its web site.

\section*{III. Accrediting as a Means of Program Evaluation}

Academic institutions engage in a process of accrediting as a means of specifying that they meet certain minimal, agreed upon standards. The accrediting process necessarily involves a process of program evaluation. Accrediting is a type of what Conrad and Wilson have called academic program review.\(^{(31)}\)

\textit{Accrediting in Journalism Education}

The Accrediting Council on Education in Journalism and Mass Communications (ACEJMC) is recognized by the Commission on Recognition of Postsecondary Accreditation and the U.S. Department of Education as the accrediting authority for professional education in journalism and mass communication in institutions of higher education in the U.S.. The Council is made up of 37 members plus a president and represents industry and the academic community. ACEJMC traces its history back to 1945, when the American Council on Education in Journalism was formed.\(^{(32)}\) The name was changed to its current version in 1980.

The Council has established what it terms "standards" for accreditation. A program must meet what is
termed "the minimum standards" (p. 5). In actual practice, programs must meet an unspecified number of these standards in order to be accredited.

Preliminary evaluation is made by a site team, which reports to an Accrediting Committee, which reports to the Council. The site team relies on a report by the program applying for accreditation as well as a visit to the campus of the program in reaching its recommendation for accreditation, provisional accreditation, or denial of accreditation. This recommendation, with an accompanying report, is evaluated by the Accrediting Committee, made up of 15 members selected by the Council. The Committee hears reports from all site teams evaluating programs in a given year and attempts to produce some commonality in evaluation. The Committee then recommends to Council that the program be accredited, provisionally accredited, or denied accreditation. Council receives the recommendation of the Committee, minutes of discussion of the site team report by the Committee, and the site team report before making its decision on accreditation. Provisional accreditation is granted when "the Council has found deficiencies that can be corrected in a relatively short time, perhaps a year" (p. 14). Accreditation is for six years.

For academic year 1996-97, 105 programs were accredited by ACEJMC. These range from Texas Woman's and Wisconsin River Falls (with 95 students and 3 full-time faculty and with 122 students and 2 full-time faculty respectively) to Syracuse (with 1,897 students and 56 full-time faculty) [33]. Size is not a criterion for accrediting. Programs are to be evaluated according to the 12 standards with recognition "that each institution has its own unique situation, particular mission, and special resources, and that this uniqueness is an asset to be safeguarded" (p. 5). At the same time, the Council recognizes that "(t)here are limitations...to this latitude. Programs...must be professional programs aimed at students seeking careers in the broad area of journalism and mass communications" and they must meet the "minimum standards" established by the Council (p. 5).

While ACEJMC does not speak of its 12 standards as research concepts, that is how they might be viewed. Each of the programs under review is evaluated on each of 12 characteristics. Evidence is gathered about where the individual programs fall along a continuum, representing levels of each of the 12 characteristics. There is a presumption that at some point along that continuum the program has met a "standard" of minimum performance, that is, that an acceptable level has been attained.

Crucial to the process, then, are three elements: measurement of each of the concepts, agreement on the basic level of achievement that must be reached for each concept for a program to be judged "in compliance with the standard," and agreement on how the 12 standards come together to produce an overall evaluation of accreditation.

Site teams "assess quality through interviews with faculty, administrators, students, and representatives of other schools; class visits, examination of students records, alumni records and communications, equipment, and buildings; and analysis of budgets" (p. 11). Much raw data are provided the team by the program in advance.

Accredited Journalism and Mass Communications Education 1995-96 lists each of the 12 standards and the evidence that a site team, the Accrediting Committee and the Accrediting Council are expected to use in reaching their conclusion about accreditation. (In fact, only the site team actually examines the evidence. It writes a report summarizing its conclusions once it has completed its visit.) Those standards and the list of evidence for each are shown in Table 16.

The first standard deals with two related concepts: Amount of faculty governance and quality of administrative leadership. The difficulty of assessing either of these in a short visit is rather obvious.
The list of types of evidence gives much weight to faculty reports on governance and supervisory comments on the performance of the administration. No sampling scheme for gathering these data or standards for measurement is offered. Obviously, programs do vary across a continuum in terms of these two characteristics. What level of governance is required as the minimum or what level of administrative competence is mandated isn't clear.

The second standard deals with budget, or, more precisely, the amount of money available to the unit under review. The team is to make comparisons of the unit's budget against others, though the precise comparisons are not specified. The expectation that students provide team members with data on student financial assistance seems particularly problematic.

The third standard seems to be quality of the curriculum. Evidence of quality is to come from student interviews, though the methods of selection of students or the questions to be asked are not specified. Other evidence is to be drawn from course syllabi and visits to classes. No sample procedure for the latter is specified.

The fourth standard is completeness of student records. Students are relied on for evidence regarding advising, though no sampling procedure or questions are specified.

The fifth standard is quality of instruction. Classroom visits and interviews with students are expected to provide some of the evidence.

The sixth standard is quality of the faculty. Faculty vitae and interviews with faculty are to be used to make the evaluation.

The seventh standard is quality of internships. Student interviews again play a key role.

The eighth standard is quality of the equipment and facilities. Team members inspect the facilities, visit the library in search of basic reference materials, and observe the use of equipment. Student interviews again are used.

The ninth standard is amount of and quality of scholarship. Team members are to examine vitae, discuss research with faculty members, and ask local practitioners about the professional activities of the faculty.

The tenth standard is amount of and quality of public service. The team is to look for records of the activities and talk with practitioners and alumni about public service of the unit.

The eleventh standard concerns itself with alumni and graduates of the unit. It isn't concerned with the actual success of the unit in placing its graduates, but only with whether the unit keeps records on its graduates.

The twelfth standard deals with minority and female representation in the unit. The concepts underlying the standard seem to be both the amount of effort by the unit to increase its female and minority representation and the actual match between the outcomes of those efforts and a standard: the characteristics of the population being served by the university in which the unit is located. Evidence of amount of effort of the unit invested in increased diversification is presence of a plan of action, inclusion of materials in course syllabi, and recruitment procedures. Evidence that the unit matches a standard comes from a comparison of characteristics of the faculty and students in the unit against the characteristics of the relevant population.
It is obvious from an examination of Table 16 that there is considerable conceptual ambiguity in the 12 standards. There also are many problems of measurement. The cut off point to be used to determine if a unit "passes" the test of a standard is wholly unspecified.

That conceptual imprecision, problems in measurement, and lack of specification of the decision rule for meeting the standard present problems for the site team, for the Committee and for the Council is to be expected. An examination of the team reports on the 13 programs evaluated for accreditation in academic year 1995-96 shows significant variation in the level of detail and attention paid to the evidence to be used in reach conclusions about accrediting.

Table 17 presents a summary of the reports for the 13 programs on a single standard, Standard 12 dealing with minority and female representation. (Three additional reports of teams revisiting Schools provisionally accredited a year earlier are not shown in the table.) An attempt was made to classify each of the reports in terms of the six types of evidence specified for Standard 12.

The first evidence requirement is that each unit has a written plan for recruitment, retention and advancement of women and minority faculty and students. For 11 of the programs evaluated in 1995, the reports contained clear statements about the existence of such a report. In one case, the report is unclear on this matter. Another report makes no mention of this requirement at all.

The second requirement is that evaluated units must "promote" course offerings "of academic vigor" across the curriculum to help prepare students to understand multi-culturalism. Two of the reports make no mention of this requirement at all in their reports. Three others report that the evidence of the existence of these courses is mixed, that is, some evidence supports the conclusion that the requirement is met, while other evidence contradicts it.

The third requirement is that units document the number of minority students enrolled and graduated from the program. No mention is made in the accrediting document that anything more than reporting of these data is to be done. The reports, however, often compare the reported data with state or service area data and with general university data, which units also must report, and there is a presumption of some linkage between these data. Table 15, for this reason, classifies the reports in terms of evidence presented about the match between the enrollment and service area data. Only three of the 13 programs were as diverse as the identified service area, based on the reports. Another five programs did not have student composition matching the service area, according to the data in the reports. No specification of the relevant data was provided in four of the reports. In one report (the University of Alaska at Anchorage), the report contained two statements at odds with each other: "Minority representation among students is not as high as the minority representation in the general population or in the population that attends the university (p. 32)" and "The unit's minority student profile matches that of the state and exceeds that of the university as a whole (p. 34)." Interestingly, in the case of California State University at Fresno, where the report shows the student data do not match either the university or service area data (p. 47), the report also contains the following statement of praise: "CSUF has one of the most diverse student bodies the site team has ever experienced (p. 53)." The report for San Jose State University also contains inconsistencies. While there is no mention of a specific comparison, note is made of the "relatively low percentage of Asian-American students in the School (p. 23)" while it is praised (p. 25) for "A diverse student body with large minority enrollments."

The fourth requirement is that units must ensure that faculty searches are conducted so as to consider minorities and women. Five of the site team reports contain no mention of this evidence; the remaining eight report evidence this requirement had been met.
The fifth requirement is that units must document that they have made "substantial, good-faith efforts" to find and recruit minority and women faculty and enhance their opportunities for promotion and tenure. All reports contained statements by the site team indicating this to be the case.

The final requirement is that units must make an effort to hire minorities and women if they make adjunct appointments. Eight of the reports make no mention of this requirement. The remaining five reports contained evidence the requirement had been met.

Only one unit, Oklahoma State University, was found to be out of compliance on Standard 12 by its site team. Oklahoma State was distinct from the other programs, according to the evidence in Table 15, only in that there was some uncertainty in the report about the existence of a written plan of affirmative action. None of the other variables is linked clearly to a noncompliance finding.

When the Accrediting Committee reviewed these 13 reports at its March 16, 1996, meeting in Chicago, conversation varied widely. Official minutes of the meeting show that the Committee asked few questions and devoted little discussion to Standard 12 (Minority/Female Representation) for Oklahoma State University, found out of compliance on that standard but on no other. There was much discussion of this standard for The Ohio State University and Indiana University, both of which were found in compliance. The Committee vote was unanimous, without abstentions, to reaccredit Oklahoma State despite the noncompliance finding on this standard. The vote to reaccredit Ohio State was 10 to 4, without abstentions; for Indiana the vote was unanimous, with one abstention, for reaccreditation. Ohio State had been found out of compliance on Standard 10 (Service). Indiana was found in compliance on all standards. In each of these three cases, the site team had recommended reaccreditation.

Similarly, minutes of the Accrediting Council meeting in May 1996, during which the team and Committee recommendations were reviewed, show similar discrepancies. At this session, the Committee recommendation for Oklahoma State University, found out of compliance on Standard 12 (Diversity) was carefully and critically reviewed, unlike at the Chicago meeting. The final vote for reaccreditation was 22-8. Consistent with the discussion in Chicago by the Accrediting Committee, the discussion of Indiana University (found in compliance on Standard 12) focused on weaknesses of the program in terms of diversity. Council voted unanimously (with two abstentions) for reaccreditation. Minimal attention was devoted to Standard 12 for Ohio State University (found in compliance with the standard), though the Accrediting Committee in Chicago had debated this issue extensively. Discussion instead focused on a merger underway at Ohio State. The Council voted 17-10 for reaccreditation.

Table 18 shows only weak correspondence between team findings, Accrediting Committee votes and Accrediting Council votes on the 13 schools under review in 1996. Three programs were found by the site team not to meet one of the standards, and two were found not to meet two of those standards. For the three programs out of compliance on a single standard, each team had recommended accreditation. The votes by Committee for accreditation were reported as unanimous in one case, 10 in favor and 4 opposed in a second, and unanimous with two abstentions in the third. Vote by Council on those three cases respectively were 22 for and 8 opposed, 17 for, 10 opposed and 1 abstention, and unanimous with two abstentions. In the case of the two programs out of compliance on two standards, team recommendations differed. One team recommended provisional accreditation, while another recommended accreditation. The Committee voted 8 for and 6 against for the program recommended for accreditation and 12 for, 1 against and 1 abstention for the program recommended for provisional accreditation. The Council voted 21 for and 9 against in the case of the program recommended for accreditation and 22 for and 6 against in the case of the school recommended for provisional accreditation. All 13 programs were granted accreditation, though one was granted the provisional accreditation recommended by the site team. That program was reviewed in a year and considered by the Committee and Council at its next meetings.
Criticism of the Accrediting in Journalism

Accrediting as practiced in journalism education has many critics, some from within ACEJMC itself. Robert Giles, editor and publisher of the Detroit News and president of ACEJMC, used a February 1995 roundtable discussion on journalism education to launch an attack on accrediting itself. Giles said that accrediting "no longer give(s) the assurance" that accredited programs are quality programs. He said "some accredited programs today are not quality programs." One solution, according to Giles, is to designate schools that exceed the basic qualifications of accrediting, in other words, set up at least two levels of accrediting.

Another recent critic is Betty Medsger, who completed a study of journalism education for The Freedom Forum. According to Medsger accrediting has contributed to what she believes is an overly heavy reliance on faculty members with doctorates by journalism programs through Standard 9 dealing with Faculty Scholarship/Research/Professional Activities. Medsger contends that an examination of the accrediting records from 1989-90 through 1995-96, involving 105 programs, shows that site teams preferred research that was aimed at other academics over research for popular consumption or traditional journalistic products. Medsger also contends that teams encouraged programs to hire faculty members with doctorates so as to improve scholarly productivity. Evidence to support both contentions is not systematically presented, and Medsger acknowledges that the "most superlative comments about the quality of research" were made about two programs where little research aimed at academics is produced.

For Medsger, a journalism program that is high in quality is one that has relatively few faculty members holding a doctorate. Medsger suggested three other criteria for evaluating journalism programs. A good program should have a high percentage (50%) of its journalism (as opposed to public relations and advertising) graduates remain in the field for at least three years, a high number (20) of its graduates working in journalism (as opposed to public relations, advertising, or some other occupation) three years after graduation, and a high number (20) of its graduates winning the monthly William Randolph Hearst Journalism Awards during the 1989-90 to 1994-95 period. Those accredited programs with high percentages and numbers of graduates in journalism three years after graduation are below average in terms of percentage of faculty with doctorates, Medsger reports. The same is true for the programs with a high number of winners of the Hearst awards. The actual data are not precisely reported.

Why Medsger used doctorates, as opposed to amount of nonacademic experience of faculty, also listed in ACEJMC documents, as her predictor variable in these analyses is not clear. She also does not indicate that ACEJMC provides no guidelines on how programs are to assess the activities of their graduates three years after graduation (though Standard 11 does mandate such an assessment of the experience of its graduates) or guidance to teams on how to evaluate the data reported. The three-year time period is one used by ACEJMC for unspecified reasons.

Medsger does credit ACEJMC, through its Standard 12 on minority/female representation, with increasing the diversity of the study body at accredited programs, though her evidence in support of this contention is indirect. Criticism of a rather different sort was leveled at Standard 12 by Ted Glasser in an article in the Journal of Communication. Glasser takes ACEJMC to task for focusing in its interpretation of Standard 12 on counting of minorities and women and providing evidence that diversity is included in the curriculum. Glasser argues in favor of different type of diversity, i.e., a diversity of perspective, which is not necessarily indexed by these measures. In response to changes in the climate for affirmative action programs nationally, ACEJMC changed Standard 12 in 1997, to go into effect in 1998.
Robert Blanchard and William Christ have offered a more general critique of accrediting. They contend that "accreditation works against the intended result of providing students with a broad-based education" by making what they believe is a false classification of all journalism coursework as outside the liberal arts. Blanchard and Christ called for relaxation of Standard 3 on curriculum to allow some journalism and mass communication courses to be classified as part of the nonjournalism curriculum. ACEJMC, in fact, made such a change in late 1992 and early 1993.

Accrediting in Other Fields

Accrediting is not unique to journalism education. An examination of documents for only two quite different fields, both of which use accreditation, gives some sense of the variability in approaches taken to the task. The fields selected are counseling psychology and engineering. The former is more scholarly in focus than journalism education. The latter focuses on more specific skills than does journalism education. As such, they offer a relevant framing for accrediting in journalism education.

Counseling Psychology

The American Psychological Association, through its Committee on Accreditation, reviews programs in Professional Psychology. Professional Psychology is defined as "that part of the discipline in which an individual with the appropriate education and training provides psychological services to the general public. Currently, these services primarily involve health and human development." Included in the purview of APA accrediting are doctoral graduate training programs and internships building on the skills acquired during doctoral training.

The key APA documents, Guidelines and Principles for Accreditation of Programs in Professional Psychology (p. 2.) spell out guiding principles of accreditation as well as procedures to be used. Standards for programs are general, stating that the program: (1) must have a clearly specified philosophy, objectives and curriculum plan, (2) that resources appropriate to achieve goals must be available, (3) that the program must recognize the importance of cultural and individual differences, (4) that the program must demonstrate mutual respect and courtesy between students and faculty, (5) that the program uses self-assessment to achieve goals, and (6) that the program describes itself accurate and completely in public documents. Similar standards are specified for the internships.

Applicants for accreditation begin the process by submitting a self-study report. Instructions for preparing the report are submitted by the Office of Program Consultation and Accreditation of APA. Upon receipt of the application and self-study, the Office of Program and Consultation and Accreditation will assign a primary and second reader to the report. The review will be in terms of the responsiveness of the applicant to the specified guidelines and procedures. The Committee on Accreditation can next authorize a site visit, ask for additional information, or deny a site visit. The Committee on Accreditation "relies on the expertise and judgment of the site visit team to verify the self-study report, to further explore the unique characteristics of the program, and to report additional information relevant to the Guidelines and Principles (p. 22)." The team can be directed by the Committee on Accreditation to obtain specific information the Committee feels is necessary.

The site visit team delivers a report to the Office of Program Consultation and Accreditation within 30 days of the visit. The site team may make evaluative comments, but it makes no recommendation regarding accreditation. The program receives a copy of the report, to which it may reply. The final accrediting decision is made by the Committee on Accreditation, at a scheduled meeting. It makes decisions on accrediting by a simple majority.
**Engineering**

Accrediting in Engineering is under the authority of the Accreditation Board for Engineering and Technology, Inc. (ABET). To be accredited, a program must include "engineering" in the program title and the program must be designed to prepare graduates for the practice of engineering. Included are programs in bioengineering, ceramic engineering, chemical engineering, electrical engineering, and other such specialities.(42) Accrediting for ABET is handled by a standing commission, called the Engineering Accrediting Commission (EAC).

ABET spells out eight criteria that are to be used in evaluating a program. The institution seeking accreditation must demonstrate to ABET that it meets each of the criteria. First, the institution must show that it evaluates, advises and monitors students in the program to determine if the program is successful in meeting its objectives. Second, each program to be accredited must have published educational objectives that are consistent with the institution's mission statement, a process for evaluation of the objectives, a curriculum and process that ensures achievement of these objectives, and a system for implementation of the results of the evaluation. Third, programs must demonstrate that their graduates have specific abilities. For example, students must know how to apply knowledge of mathematics, science and engineering, be able to design and conduct experiments and analyze and interpret data, be able to use the techniques, skills and modern engineering tools, and be able to formulate and solve engineering problems. Included in the list of requirements is that students be able to communicate effectively, have a knowledge of contemporary issues, and recognize the need for life-long learning. Evidence of success may include student portfolios, nationally normed subject content examinations, alumni surveys, employer surveys and placement data.

The fourth criterion on which ABET evaluates engineering programs is termed the professional component. ABET specifies subject area, but it does not prescribe specific courses. The faculty must demonstrate only that adequate attention and time to each component, consistent with the objectives of the program and institution. The professional component must include one year of college-level mathematics and basic sciences, some with experimental experience, one and one-half years of engineering topics, to include engineering sciences and engineering design, and a general education component that complements the technical content of the curriculum and is consistent with the program and institutional objectives.

The fifth criterion is the faculty. ABET specifies that the faculty must be of sufficient number and must have the competencies to cover all of the curricular areas of the program. There must be sufficient faculty to "accommodate adequate levels of student-faculty interaction, student advising and counseling, university service activities, professional development, and interactions with industrial and professional practitioners as well as employers of students." The faculty must be qualified, based on such things as education, diversity of backgrounds, engineering experience, teaching experience, ability to communicate, level of scholarship, and participation in professional societies.

The sixth criterion is facilities. ABET states: "Classrooms, laboratories and associated equipment must be adequate to accomplish the program objectives and provide an atmosphere conducive to learning...Programs must provide opportunities for students to learn the use of modern engineering tools. Computing and information infrastructures must be in place to support the scholarly activities of the students and faculty and the educational objectives of the institution." Criterion seven calls for evidence of institutional support and adequate financial resources for the program. The final criterion is related to the specific program in place. A program with a particular engineering specialization must have a curriculum and faculty in place to guarantee competence in that specialization.

The programs are evaluated on the basis of data submitted by the institution to ABET in the form of a self-study questionnaire. Among the data reported are syllabi (using a standardized format), curricular requirements, detailed descriptions of facilities, and faculty vitae (also in standardized format). Data from the questionnaire are supplemented by a report of an on-site visit by a team representing ABET. The purpose of the on-site visit is three-fold, according to the ABET Accreditation Yearbook (1994 edition). The visit should "assess factors that cannot be adequately described in the questionnaire," such as intellectual atmosphere, the morale of the faculty and students, the caliber of the staff and student body, and the outcome of the education offered. The team also should help the institution assess its strong and weak points. Finally, the team should examine in further detail the material compiled by the institution related to such things as control of the accredited unit, programs offered, age of the institution, requirements for admission, number of students enrolled, teaching staff and teaching loads, physical facilities, finances, curricular content, and representative samples of student work.

The visiting team reports its preliminary findings and recommendations in writing to ABET. ABET then forwards the report to the institution for response. The institution may provide supplemental materials that may be helpful to ABET.

EAC, accrediting commission of ABET, meets annually to consider the applications for accrediting. The decision is based on the team recommendation, the response of the institution, and the data provided by the institution in the initial questionnaire.

IV. Concluding Comments

Academic Literature

The academic literature makes clear that evaluation of higher education is extremely complicated. There are different notions of what is meant by quality, and different views about how that quality should be measured. Ironically, the National Research Council study of quality in doctoral programs, examined here as an example of evaluation by academics of academic programs, is very limited in its focus, giving most attention to reputation as an indicant of quality.

In the end, the academic literature is very clear in making one point: Quality is not unidimensional. It can be indexed by characteristics of the institution, by characteristics of the faculty, by characteristics of the students, by activities of the students, and by outcome, to name just a few options. There is a striking difference between an evaluation that focuses on what students bring to the academic environment, such as entry SAT or ACT scores, and one that focuses on changes in knowledge from entry to departure from the university. In fact, one can imagine that, under ideal circumstances, those institutions admitting weaker students might be able to demonstrate more knowledge gain than those admitting strong students because the weak students have the most room to grow. (In actual practice, this is unlikely to be the case because the weak students most likely will have deficiencies that will make it difficult for them to acquire new knowledge.) Indicators of quality focusing on process, rather than input and output, might well not be sensitive to the differential needs of students. What works well for one student, such as extensive hands-on learning, might hamper another student who excels in abstract work environments.

The conclusion that seems most appropriate from the academic literature is the one offered by Gaither, Nedwek and Neal: use multiple indicators of quality, and concern oneself with the procedures of measurement. A measure that isn't reliable, i.e., repeatable, isn't valuable. Nor is one that has no connection to quality itself, though it may be easy to actually gather data relevant to it.

The National Research Council study, with its almost total reliance on reputational indicants of quality,
does not meet Gaither, Nedwek and Neal's requirement that multiple indicators of quality be used. As Conrad and Wilson have noted, a reputation can be outdated, and, at best, it focuses on scholarly productivity and peer evaluation, rather than the actual experience of those in the program.\(^{(44)}\)

The NRC study raises another important issue: What happens when programs are not evaluated and, for that reason, cannot be compared with others within the university in terms of reputation? The Speech Communication Association (since renamed the National Communication Association), conducted its own reputational study of communication doctoral programs because it feared the consequences of the NRC decision not to include communication in the national study.\(^{(45)}\) The method was similar to that used by NRC. Though Mass Communication was one of the specialities included in the study, data were not reported for this speciality because mass communication programs associated with journalism schools were not included in the survey. It seems likely that programs not included in NRC do face a significant threat, as they are viewed as marginal and are not easily able to argue their standing nationally in a way to garner resources within the competitive university environment.

**Commercial Evaluations**

As noted, the commercial evaluations have come under criticism in recent years because of their dependence on information provided by those institutions being studied. Because institutions have little incentive for providing accurate information other than their own pride and fear of being detected if they mislead, there is good reason to be suspicious of the information provided. Efforts at standardizing the information gathered would seem to help some. The commercial firms need data to produce their evaluations, however, and actually are better served by weak data than by no data at all. In fact, there isn't much incentive for the evaluators to invest beyond the minimum in gathering and analyzing the data.

At present, the commercial evaluators use a hodgepodge of methods to reach their evaluations. As shown, *U.S. News & World Report* even uses a variety of methods for the various fields it evaluates. There also is much variation in the information provided about these evaluations. *The Gourman Report* provides almost nothing to help the reader understand the scores presented. As the *U.S. News* evaluations of journalism graduate programs shows, the evaluators can be disingenuous in discussing deficiencies in the data actually available to them. The error in the 1997 law school rankings is another case in point.

None of the commercial services deal adequately with questions of sampling in their reputational and other surveys. Perhaps *The Princeton Review's* use of students who enroll in its own classes to prepare for graduate admissions tests is most egregious. That these students are a biased sample of the population of graduates is highly likely. Even the noncommercial NRC reputational survey, however, is poorly documented in terms of return rates. It isn't possible with the published data, for example, to know exactly how many evaluators actually rated a program, such as the Political Science Department at The Ohio State University. In any case, the number is likely to be small, since the goal was set at approximately 100 evaluators per program.

The commercial evaluators, however, do provide concrete examples of the kinds of data that can be gathered that seem to be of interest to those interested in actually enrolling in the programs—consumers of the product. (We assume that commercial services that don't fill a need will vanish for lack of financial resources.) The *Money Magazine* list of criteria, shown in Table 13, is a particularly good example of precision in the specification of the data gathered. Someone can dispute the importance of a given criterion, but the data are at least at hand for individual assessment. One can imagine that a clever consumer would build his or her own index by scanning or downloading the matrix and computing his
Accrediting

Accrediting is a very prominent activity in journalism and mass communication education in the United States. Much energy is invested in it. It has much to offer. As a method of program evaluation, it can offer consumers of the journalism education product much helpful information about various programs available and set minimum standards against which those programs can be evaluated.

Clearly each of the 12 characteristics of journalism and mass communication programs selected by ACEJMC for evaluation of programs is relevant to the overall quality of those programs. There is room for debate about the completeness of the list, however, and the relative importance of each. ACEJMC documents provide no suggestion that any one of the "standards" is more important than others, though, as noted, discussion at the March 1996 Committee meeting and the May Council meeting focused heavily on Standard 12. Clearly Standard 12 has been the center of much discussion in recent years. Whether changes on the national scene in the area of affirmative action will increase or decrease the prominence of Standard 12 in the short run is hard to say.

The problem with accrediting as it is currently practiced in journalism education seems to be more an issue of measurement and a lack of specification of the decision rules to be used to determine when a standard is met than with the standards themselves. The measures take no notice of what is normative or existent at other programs. In addition, there is little focus on outcomes, that is, the consequences of what is taught at the university. In addition, calibration of the data for the number of students served would be helpful. At present, for example, the evaluation does not specify appropriate student-to-faculty ratios, and the accrediting documents do not even ask for calculation of this key statistic.

Current attempts to improve accrediting by creating a two-tiered system or by designating certain programs as superior to others seem doomed to failure. Without clear measurement procedures and without specification of what does and what does not constitute "compliance" it is impossible to take the next step calling for even more detailed decisions on what constitutes "exemplary compliance."

This not to argue, however, that the current approach of accrediting all programs by the same standard is appropriate. An institution that attempts to offer undergraduate through doctoral instruction might appropriately be evaluated differently from a program offering only undergraduate programs or even one offering only graduate programs. A program that focuses on only one aspect of journalism education might be evaluated differently from one that attempts to bridge the many fields.

As a starting point, ACEJMC might want to divide programs seeking accrediting by size or according to the characteristics of the overall university. The Carnegie classification scheme has enjoyed wide acceptance and might be adopted. Such a distinction would lead to an assessment of the research contribution of programs within research universities and an emphasis on the teaching mission of institutions dedicated to undergraduate liberal arts curricula. Another option would be to use the U.S. News classification of national universities, national liberal arts colleges, and regional colleges. The present one-standard-fits-all approach seems not to recognize inherent differences between programs across a variety of dimensions. Providing more homogeneous groupings would make more sense to the programs themselves and provide more information to prospective students and their families.

ACEJMC currently relies on the Accrediting Committee to provide standardization of review. The central office of ACEJMC develops common reporting techniques and assembles the site team, but it does not evaluate the materials received or the reports of the teams. Since the Committee has only the
team reports to go on and does not, in fact, even see the self-study produced by the program, the Committee has limited ability actually to standardize the process. The team has tremendous importance.

In fact, the reliance almost entirely on an outside team has some disadvantages. Conrad and Wilson, in their review of program evaluation, note that most program reviews rely on reviewers external to the program, but two different approaches exist. One is to use reviewers from the same discipline who are outside the university. The other is to use reviewers from the same university who are outside the discipline. The advantage of use of outside reviewers from the same discipline is that there is an understanding of disciplinary needs and peculiarities. Reports also may be more likely to be seen as objective by those evaluated. Among the limitations is a lack of knowledge of the local context, lack of adequate time to assess the program, a tendency to focus on insignificant issues, and a tendency to focus on additional resources as the solution to problems. The use of local evaluators offers the advantage of individuals who know the local context and norms. The lack of knowledge of the discipline is a disadvantage as is a tendency to allow previous familiarity to influence the evaluation.

Obviously, it is possible to include both. ACEJMC site teams do interact with others in the academic community. This is not done systematically, however. Teams have latitude in whom they speak with and the context of the conversation.

It is tempting to say that the procedures used by psychology and engineering, reviewed here, are superior. On the face, they do provide for more centralized control. In the case of the American Psychological Association, the activity of the site team is highly constrained. The team does not even make recommendations on accrediting, but rather attempts to verify the materials provided by the program itself. The engineering standards seem to have the advantage of focusing on curriculum, physical and faculty resources, and actual knowledge gain. Teams would seem not to have such great latitude in interpreting these standards. A clear disadvantage, however, is that such an approach would lead almost inevitably to relatively little creativity in curricular development, as the safe strategy would be to do what has been accredited elsewhere and in the past. Whether such an approach is appropriate to journalism education, with its focus on expressive rights, is questionable. Also questionable, however, is whether the current system does in fact lead to creative responses to challenges facing those programs.

The biggest problem in comparing journalism and mass communication accrediting, as it operates in fact, with psychology and engineering programs, as they operate in the abstract, is that the failings of the former are easy to see, while the failings of the latter are unspecified. What is clear, however, is that the current accrediting standards in journalism and mass communication education do little to provide specific guidance to site teams in terms of measurement and the minimum standard to be met and, for that reason, give great liberty of interpretation to the team, the Accrediting Committee and the Accrediting Council. That this happens in fact seems quite clear from the data assembled for the 1995-96 accrediting cycle.

**Implications of the U.S. Experience**

Program evaluation came into prominence in the United States in the 1960s, when the country experimented with a variety of initiatives as part of the War on Poverty and Great Society programs of presidents John Kennedy and Lyndon Johnson. Critics, particularly, wanted to know if the programs were effective, and by 1965, the Elementary and Secondary Education Act for the first time required federal funding recipients (in this case, educational institutions) to provide formal evaluations of the programs implemented.

Concern with program effectiveness, however, is by no means uniquely American. Much of the
academic literature on evaluation reviewed in this report comes from outside the United States. Much of it is comparative in nature.

Commercial and media evaluation of educational products also is common outside the United States. The cover story of the May 5 issues of the German newsweekly, Der Spiegel, for example, contained a ranking of the "best" universities in Europe by discipline. The report was based on a survey of more than 1,000 professors and 7,400 students from 15 European countries.

What is unique in the U.S. experience, at least to this point, is a national system for accrediting of journalism and mass communication programs. Is this experience so uniformly positive that it should be exported to other contexts?

The evidence provided in this report would suggest that great caution should be exercised in this regard. While the goals of accrediting as practiced in the U.S. may be laudatory, the practice has great limitations. Without a clearer understanding of the conceptual base for the evaluation, specification of measures, and consensus on actual standards for assessment of performance, such an evaluation exercise seems destined to failure.

In the end, it is clear, that program evaluation is an extremely complex process. A large number of different criteria can be applied, and probably should be applied. Not all universities or programs in a given country should be expected to excel on all criteria. It probably also is true that not all countries should be expected to have the same goals for their programs and, consequently, standards against which a program should be evaluated.

In its simplest form, program evaluation is determining the worth or merit of a program. The process of evaluation may be generalizable from one national setting to another, but "worth" and "merit" are clearly culturally bound.

1. The authors acknowledge the support of the John S. and James L. Knight Foundation of Miami, FL, USA, which provided a grant to the authors for review of evaluation as practiced in journalism education.


6. The coefficient used was the Pearson Product Moment Correlation coefficient.


10. We conduct the Annual Survey each year at the University of Georgia.


20. Technically, the coefficient should not be used with a rank. The Spearman's rho is a +.54, while the Kendall tau b is +.40, indicating that the coefficient makes little difference in terms of the inference.

21. This coefficient is a Kendall's tau b.

22. Ranking data are reported as well. In these analyses, the raw scores are used.


33. The data come for the *Annual Survey of Journalism & Mass Communication Enrollments* for academic year 1995-6.

34. The page citations are to the team reports.


38. The evidence required to meet the standard is now a written statement of goals for "diversity and inclusivity" and a description of "actions they have taken and the specific results achieved and progress made toward diversity and inclusivity goals." See "Standard 12: Diversity," *AEJMC News*, July 1997, p. 5.


