DOCUMENT RESUME

ED 339 329 HE 025 114

AUTHOR Thomas, Alice M.
TITLE Standards for the Conduct of Quality Assessment in Higher Education. ASHE Annual Meeting Paper.
PUB DATE 3 Nov 91
NOTE 22p.; Paper presented at the Annual Meeting of the Association for the Study of Higher Education (Boston, MA, October 31-November 3, 1991).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Data Collection; Delphi Technique; *Educational Assessment; *Educational Quality; *Evaluation Criteria; Higher Education; Postsecondary Education; Program Evaluation; *Research Design; *Research Methodology; School Effectiveness; *Standards; Undergraduate Study
IDENTIFIERS *ASHE Annual Meeting

ABSTRACT

This paper reports on research that was designed to produce a limited number of assessment standards judged by experts as the most important standards in the practice of quality assessment in undergraduate higher education. The study collected the opinions of a group of national assessment experts and employed a modification of the Delphi technique in order to arrive at a group consensus. The research resulted in identification of 40 standards that were agreed upon by the expert group. The rationale for each of these standards is presented. The study's research phases included: (1) the clarification of the component of the institution/program to be assessed and its context; (2) the designing of the assessment process; (3) collecting and analyzing the data; (4) communicating the assessment's findings; and (5) using the findings to make recommendations, make decisions about improvements, and make judgements about quality. The appendix contains descriptions of the 40 standards. Contains 19 references. (GLR)

STANDARDS FOR THE CONDUCT OF QUALITY ASSESSMENT IN HIGHER EDUCATION

Alice M. Thomas
Visiting Assistant Professor
Department of Educational Leadership
University of Tennessee, Knoxville

Annual Meeting of the Association for the Study of Higher Education
October 31 - November 3, 1991
Boston, MA

Texas A&M University
Department of Educational
Administration
College Station, TX 77843
(409) 845-0393

ASSOCIATION FOR THE STUDY OF HIGHER EDUCATION

This paper was presented at the annual meeting of the Association for the Study of Higher Education held at the Park Plaza Hotel & Towers in Boston, Massachusetts, October 31-November 3, 1991. This paper was reviewed by ASHE and was judged to be of high quality and of interest to others concerned with the research of higher education. It has therefore been selected to be included in the ERIC collection of ASHE conference papers.

The current concept of assessment in higher education rose to the forefront in the mid-1980s through national attention on accountability and quality in higher education. The public expressed uncertainty about the worth of higher education and asked institutions, primarily public institutions, to demonstrate their educational effectiveness. These concerns appeared in several national reports, including those of the National Institute of Education [14], the Association of American Colleges [1], the Southern Region Education Board [16], and the Education Commission of the States [6].

Quality, accountability, and assessment are not new issues in higher education, although each has acquired a slightly different meaning in recent years. A traditional meaning of assessment focused on gathering information about an individual, usually through multiple tests and procedures, for purposes of placement or guidance [10]. In contrast, the term assessment currently has a broader meaning that generally applies to the gathering of information to indicate the extent to which an institution is achieving what it purports to do, that for which it is accountable. This research used the more inclusive definition below.

> Assessment: The use of various methods to gather both quantitative and qualitative information at the level of program, institution, and/or system, to describe and sometimes to make judgments about the inputs, resources, and/or outcomes of an undergraduate education for purposes of improvement and/or accountability (individual student diagnostic assessment has been omitted from the context) [modification of a definition provided by Boyer and Ewell [2]].

This usage reflects the expansion of the term assessment, one that has moved beyond gathering chiefly quantitative, testing information on the level of individual students for diagnostic purposes. Assessment as defined above extends the type of information to include gathering both qualitative and quantitative information, expands the level of analysis to several possible levels, and includes the conduct of assessment for improvement and/or accountability purposes.

Incorporated in the definition of assessment used in this research, was the use of the gathered information to make value judgments about quality. Whether or not to include the stage of making judgments in the definition of assessment is an issue of debate in the field [3]. When included, assessment becomes almost indistinguishable from evaluation in which values are placed on the object of the gathered information.

The above definition of assessment also illustrates the recent shift of the basis for accountability and judgments about quality from an emphasis on efficiency to one on effectiveness. Declining enrollments and resources, and increasing costs in the 1970s found researchers involved in assessment of efficiency, i.e., a form of accountability based on judgments of financial and other efficiency criteria, e.g., enrollment patterns, student-faculty ratios, credit hours, and cost per credit hour. In contrast, the current bases for accountability and quality judgments are related to effectiveness, i.e., the extent to which an institution, program, or student is reaching educational goals, goals for which the institution is accountable [9, 12].

BACKGROUND

National surveys have indicated that a majority of the states have formal initiatives labeled assessment, and that a wide variety of assessment programs are used by institutions [8]. Terenzini [17] provided a useful framework for classifying assessment programs based on the purpose and the unit of analysis: individual assessment for placement or guidance; individual assessment for gate keeping purposes (e.g., rising junior exams, comprehensive exams, and certification exams); group assessment for purposes of program enhancement; and group assessment for campus and program evaluation (e.g., alumni surveys and value added studies). This variety of assessment programs appropriately reflects the diversity among institutions in their mission, the type of entering student, the areas of expertise and priorities of the faculty, available resources, teaching and learning programs, the institutional reward system, and external and internal pressures for assessment. Personnel in individual institutions are faced with a wide variety of examples of assessment programs, and with what might seem a formidable task of deciding what to assess and how to assess at their unique institution.

NEED FOR RESEARCH

Much of the assessment literature addresses assessment from a very general perspective or a narrow approach such as institution-specific information that may not be highly transferable to other institutions. Missing from the assessment literature are general assessment principles that could be applied to any institution regardless of context [19]. Such tools are needed to guide practitioners in addressing the question of how to assess, and to encourage the conduct of sound assessment programs.

Evaluators working in higher education have expressed concern about the continued lack of specific methodology, models, or theoretical perspectives to inform the practice of assessment, despite the applicability of the theory and methodology of evaluation to the assessment enterprise [4, 15]. Although an adaptation of Stake's Countenance Evaluation Model was suggested as a framework for the assessment process [18], minimal use has been made in assessment of the developed tools used in the field of evaluation that are directly applicable, and potentially valuable in the conduct of assessment.

Evaluation standards are another potentially valuable evaluation resource for guiding the conduct of assessment. Within the field of evaluation, sets of standards from the Joint Committee [11] and the Evaluation Research Society [7] have provided guidelines for conducting evaluations. This research was based on the assumption that similar standards could provide guidance for the practice of sound assessment. The goal of this research was to promote the conduct of high quality assessment programs through identification of appropriate assessment standards.

Using the existing evaluation standards as a basis, the research was designed to produce a limited number of assessment standards that are judged by experts as the most important standards in the practice of quality assessment in undergraduate higher education. A limited number of standards for conducting an assessment might be more acceptable, and more readily used than the comprehensive sets of standards

for conducting general evaluations that are currently available. To improve the utility of the set of assessment standards, the research was also designed to generate an accompanying rationale for the importance of the standards to lend insight into the use of the standards and the general practice of quality assessment.

METHODOLOGY

A modification of the Delphi technique, a method for arriving at group consensus, was used in the study. National assessment experts served as the Delphi participants in this research designed to identify standards for the conduct of quality assessment programs.

Delphi Technique

Several factors contributed to the selection of the Delphi from among available group methods for reaching consensus. Because a set of assessment standards should be applicable in the many different contexts in which assessment takes place, obtaining a group of participants reflective of those contexts was particularly important to the research. The mail methodology of the Delphi facilitates diverse representation by allowing participation of experts who are widely geographically dispersed and whose schedules might not allow group meetings [5, 13]. The Delphi methodology also avoids face-to-face meetings that provide response anonymity for the participants. In this research, such anonymity prevented any possible effects on participant responses due to the professional prominence or assessment approach identified with particular experts or institutions.

Central to a Delphi is a series of succeeding questionnaires that ask for opinions or judgements, allow participants to review a summary of the group's responses from the previous round, and encourage revision of opinions or judgments. This process works toward the goal of a convergence of ideas, i.e., a consensus about an issue. The iterative process of questionnaire distribution, analysis, and feedback resulted in three rounds in this research.

First round. A typical Delphi first round for this research would have asked participants for a listing of standards they thought were important. A compiled list of these standards would have provided an inventory of potentially applicable standards. Developing such an inventory was not necessary in the first round in this research, however, because of existing sets of evaluation standards developed by groups of professional organizations, i.e., The Joint Committee [11] and the Evaluation Research Society [7]. An initial inventory of 38 standards was compiled by the researcher from the two existing sets of evaluation standards after eliminating duplicates. The wording of the standards was revised for the context of quality assessment in higher education. To assure that the inventory was comprehensive, participants were requested to add any standards that they deemed to be missing. Participants in two partial pilots added two additional standards to the initial compiled set of 38 standards producing the 40 assessment standards used in the final research.

In the first round, participants were asked to rate the importance of the standards using a rating scale of one to five (1=very important, 5=not important).

Participants rated the importance of each of the 40 standards in each of the five phases of an assessment as defined in the research (a total of 200 ratings). The five phases of assessment as outlined for participants included the following:

Clarifying the Assessment Object and Context: Describing the component of institution/program to be assessed and its context, clarifying the purpose of assessment, deciding whether to conduct assessment.

Designing the Assessment: Identifying the questions to be answered, developing overall assessment plan including approach to be used and tactics for gathering information, and developing a management plan for use of resources.

Collecting and Analyzing the Data: Collecting and preparing data, analyzing and interpreting data.

Communicating the Findings: Reporting and communicating the collected information, the findings.

Using the Findings: Using the findings in applying results, making decisions about improvements, making and reporting recommendations, and possibly making judgments about quality.

Second round. In the second-round questionnaire, participants were asked to complete two tasks: re-rating and providing rationale statements for their ratings. Participants re-rated the importance of the standards that meet the established criteria of importance in the first round ratings. Two criteria were used throughout the research to define the levels of importance of the standards. The criteria were the values of the first quartile (Q_1) and the third quartile (Q_3) in the distribution of responses. For purposes of this research, the standards were defined as having the following level of importance based on the value of the quartile ratings:

"very important standard"
Q_1=1.0, Q_3=1.0, i.e., at least 75% assigned the standard a "1" rating of very important,

"generally important standard"
Q_1=1.0, Q_3=2.0, i.e., at least 25% of the participants assigned the standard a "1" rating of very important and at least an additional 50% assigned the standard at least a "2" rating of generally important,

"somewhat important standard"
Q_1=1.0, Q_3=3.0, i.e., at least 25% of the participants assigned the standard a "1" rating of very important and at least an additional 50% assigned the standard at least a "3" rating of somewhat important.

In addition to completing importance ratings in the second round, participants also selected two standards in each phase for which to provide a short rationale of the importance of the standards.

Third round. In the third round participants were asked to reconsider their rationale for the importance of the standards. As the basis for that task, a summary of the rationale statements from the second round was compiled by the researcher. The compilation work was validated by a monitor team consisting of three members who had been active in assessment nationally: a faculty member, an administrator, and an executive staff member of a national educational association.

The third-round participants read the summarized rationale statements for each phase, indicated any gaps they found in the statements of rationale for the importance of the standards, and responded to any statements with which they disagreed by crossing out the statement and entering their own position. The researcher used the responses to revise the summary; revisions were subsequently validated by the monitors.

Expert Participants

A sample of 50 participants from each of two categories of experts constituted the initial total sample of 100 assessment experts: 50 individuals with major responsibilities in institutions with established, active assessment programs, and 50 authors of assessment conference papers or articles in journals and publications of six professional organizations that had addressed the topic of quality assessment.[1] No duplication of institutions existed among the participants. A total of 85 assessment experts (45 active in assessment and 40 authors) agreed to participate.

In the first-round, a return rate of 82% was achieved with 70 of the 85 participants returning the questionnaire that requested only ratings. Of the 70 participants who returned the first round ratings, 53 (76%) returned the second round re-ratings and written rationale statements, and 38 (54%) completed the third round revisions of their rationale statements.

RESULTS

The results of this research included both a set of standards agreed upon by a sample of assessment experts as being the most important in the conduct of quality assessment programs in each of the five phases of an assessment, and a rationale for the importance of each of those standards. The former, i.e., the set of assessment standards produced by the importance ratings, is the primary focus of this paper; only a few rationale comments will be reported to illuminate the ratings.

Important Standards by Phase

The number of assessment standards rated as important within each of the five phases of assessment used in this research is presented in Table 1.

Place Table 1 about here

The research identified 12 standards as being *very important*, the highest of the importance ratings, in at least one phase of an assessment (one standard was rated as *very important* in both of the first two phases). The number of standards identified as being *very important* differed greatly among the assessment phases I through IV. As indicated in Table 1, none of the standards were identified as *very important* in Phase V. The importance ratings given each standard by assessment phase is reported in Table 2.

Place Table 2 about here

Only those assessment standards (AS) meeting the criterion established for the level of *very important* (the highest of the three importance categories) are discussed below.

Phase I: Clarifying Object and Context. The first phase of the assessment as defined in this research included the following tasks: describing the component of the institution/program to be assessed and its context, clarifying the purpose of the assessment, and deciding whether to conduct the assessment. As indicated in Table 2, two standards met the criterion for the highest level of importance, *very important*. Those standards were as follows:

> **AS 2 Described Purposes and Procedures:** The purpose of the assessment should be clarified; and procedures selected should be appropriate for the purpose. Descriptions of the purpose, procedures, and range of activities should be in enough detail, so that they can be identified and assessed.
>
> **AS 25 Object Identification:** The purposes and characteristics of the object of the assessment to be addressed in the assessment should be specified as precisely as possible, e.g., program objectives or institutional mission statement.

Participants indicated in the rationale statements that they viewed the clarification of what will be assessed and why to be the major tasks of the first phase. Standards AS 2 Described Purposes and Procedures, and AS 25 Object Identification were perceived by participants to directly address those tasks and to essentially define the first phase.

Participants noted that an assessment program may be for the purpose of improvement in an area such as teaching or learning, or for the purpose of accountability to a group such as trustees or legislators, or some combination of both. Because assessment for one purpose may not fulfill the needs for another, clarifying the purpose was deemed essential in providing appropriate direction and guidance for subsequent decisions about the object of the assessment, the procedures, and the appropriate tools and strategies. A major reason cited in the rationale summary for the importance of attending to clarification of the purpose related to the consequences of an unclear purpose: an assessment program that can easily stray from the original intent and result in expenditure of funds but unusable information.

Standard AS 25 Object Identification was viewed by participants as important for reaching an agreement on what will not, as well as on what will be assessed, and for other necessary decisions about subsequent assessment activity. A possible outcome identified by participants of an unfocused assessment was a snowball growth of assessment activity through attempts to assess everything.

Phase II: Designing the Assessment. As defined in the research, Phase II included the tasks of identifying the questions to be answered, developing an overall assessment plan including an approach to be used and tactics for gathering information, and developing a management plan for the resources (see Table 2 for

standards considered important in Phase II). The standards rated by participants as *very important* in the second phase included the following:

> **AS 2 Described Purposes and Procedures:** The purpose of the assessment should be clarified; and procedures selected should be appropriate for the purpose. Descriptions of the purpose, procedures, and range of activities should be in enough detail, so that they can be identified and assessed.
>
> **AS 23 Rights of Human Subjects:** Assessments should be designed and conducted to protect and respect the rights and welfare of the human subjects.
>
> **AS 21 Reliable Measurement:** The information-gathering instruments and procedures should be chosen or developed and then implemented in ways that will assure that the information obtained is sufficiently reliable for the intended use.
>
> **AS 28 Valid Measurement:** The information-gathering instruments and procedures should be chosen or developed and then implemented in ways that will assure that the interpretation arrived at is valid for the given use.

The rationale given by participants for assigning high importance of standard AS 2 Described Purposes and Procedures in both the first two phases, related to the close tie between clarifying the purpose of an assessment in Phase I, and selecting an appropriate assessment design for that purpose in Phase II.

Standard AS 23 Rights of Human Subjects was viewed important in protecting the rights of assessment participants. Such protection was deemed essential from an ethical, professional, and legal standpoint. In addressing AS 21 Reliable Measurement, and AS 28 Valid Measurement in the rationale summary, participants acknowledged that reliability and validity of instruments and procedures to gather the information are basic in any field of inquiry, and must be considered in designing an assessment if the results are to be accepted and of utility.

Phase III: Collecting and Analyzing Data. This phase was defined for the participants as collecting and preparing data, and analyzing and interpreting the data (see Table 2 for important standards in Phase III). The standards identified by the participants as those that were *very important* in the phase of collecting and analyzing data included the following:

> **AS 9 Evaluator Credibility:** The persons conducting the assessment should be both trustworthy and competent to perform the assessment.
>
> **AS 24 Analysis of Quantitative Information:** Quantitative information in an assessment should be appropriately and systematically analyzed to ensure supportable interpretations. Report potential weaknesses in the data collection and describe possible influences on findings.

In the rationale statements, participants indicated that they considered assessment to be not only a scientific endeavor but also a social, political, and educational activity. Consequently, participants emphasized the importance of the

persons conducting the assessment being both competent for the technical endeavors, and trustworthy in performing a task so heavily social and political in nature. Participants viewed this standard as not only important for the quality of the information but also for the credibility of the program.

The standard AS 24 Analysis of Quantitative Information was the only other assessment standard identified as *very important* in this phase by the experts. Participants indicated in the rationale summary that such adherence to standards regarding the analysis of information was basic to good research as well as sound assessment.

Phase IV: Communicating the Findings. For the purposes of this research, Phase IV was defined as reporting and communicating the collected information, i.e., the findings (see Table 2 for standards identified as important in Phase IV). Participants identified the following five standards as *very important* in reporting the findings.

AS 20 Balanced Reporting of Findings: The assessment should be complete and fair in its presentation of the findings, e.g., balanced reporting of data that may lead to conclusions about strengths as well as weaknesses, and inclusion of intended as well a unintended outcomes.

AS 35 Objective Reporting of Results: The assessment procedure should provide safeguards to protect the assessment findings and reports against distortion by the personal feelings and biases of any party to the assessment. When appropriate, draft reports should be checked with representatives of audiences for clarity and accuracy.

AS 6 Full and Frank Disclosure: Oral and written material containing gathered information should be open, direct, and honest in the disclosure of pertinent findings, including the limitations of the assessment.

AS 34 Report Clarity: Reports of results should describe the object being assessed and its context, and the purposes, procedures, and findings of the assessment. Limitations caused by constraints on time, resources, and data availability should also be stated.

AS 7 Targeted Dissemination: Persons, groups, and organizations who have contributed to the assessment should receive feedback appropriate to their needs. When appropriate, check draft reports with representatives of audiences for clarity and accuracy; consider tailoring dissemination for specific audiences using a variety of formats and approaches.

Three of the standards identified by participants as *very important* in this phase addressed the content of the report and included AS 35 Objective Reporting of Results, AS 20 Balanced Reporting of Findings, and AS 34 Report Clarity. In their statements of rationale, participants indicated that objective and balanced reporting was critical in presenting a complete picture that can enhance a full, broad perspective of the object, and can improve the quality of subsequent conclusions and recommendations.

Assessment standards AS 6 Full and Frank Disclosure, and AS 7 Targeted Dissemination, which focused on the dissemination of the report, were also deemed *very important* by participants in this phase. Participants considered these standards as important to assure that those needing access to the information would obtain it in a form that would enable them to act with full knowledge. This rationale was closely linked by participants in the summary to that given for Standard AS 7 Targeted Dissemination in which participants acknowledged that different audiences have different needs for assessment, and that although some data may not be appropriate for one audience, the same data may be imperative for the work of others.

Phase V: Using the Findings. The last phase of an assessment was defined in this research to include using the findings in applying results, making decisions about improvements, making and reporting recommendations, and possibly making judgments about quality. As indicated in Table 2, however, none of the 40 standards listed in the original questionnaire inventory were rated by participants as *very important* (the highest category of importance) in the last phase of an assessment.

Standards Important Across Phases

An additional analysis was conducted to determine if any standards played a consistently important role in all or most phases without necessarily receiving the highest importance designation, *very important*, in any one phase. Seven of the total of 40 standards met the criterion for *somewhat*, *generally*, or *very important* in at least four of the five phases of assessment. The standards were AS 1 Human Interactions (important in all five phases), AS 2 Described Purposes and Procedures, AS 9 Evaluator Credibility, AS 23 Rights of Human Subjects, AS 30 Conflict of Interest, AS 31 Audience Identification, and AS 32 Unit of Analysis. Four of these standards (AS 1, AS 30, AS 31, and AS 32) had not been identified as *very important* (the highest importance category) in any phase.

Five of the seven standards that were consistently important across the assessment phases addressed dimensions with human components. These five standards (AS 1, AS 9, AS 23, AS 30, and AS 31) appeared to include all but two of the standards in the inventory (AS 12 Audience Cooperation and AS 26 Public's Right-to-Know) that addressed issues directly involving individuals.

Final Set of Assessment Standards

The research identified a set of standards that were judged by assessment experts to be the most important in the conduct of quality assessment programs in higher education. The set of 16 assessment standards, identified in Table 2, is composed of the 12 standards that received the highest ratings in at least one phase and the four standards that were consistently across the phases rated important to some degree.

CONCLUSIONS

The importance ratings and the rationale statements provided in this research by the assessment experts identified several areas that were deemed critical but which had been identified in an earlier review [19] as receiving sparse attention in the assessment literature. A gap between the importance and the available guidance

was identified for several assessment dimensions: human and political issues including political viability, the rights of human subjects, ethics, conflict of interest, and human interactions; collection of information involving the role and analysis of qualitative information, and the credibility and competency of persons conducting assessment; report writing and dissemination issues of full and frank disclosure, right-to-know audiences, and targeted reports; and utilization issues such as assessment side-effects, responsibility for use of findings, and promotion of utilization.

To improve the quality of the assessment information available to client practitioners, appropriate professional organizations should expand the scope of the higher education quality assessment issues addressed in their publications and conference presentations to include the areas identified in this research as important but sparsely represented in the literature. The importance of the human dimensions in assessment found in this research may warrant particular attention by those in a position to provide needed skills and understandings in this area. The need to attend to this aspect of assessment was largely ignored in the literature both as a general concern about practice, and as an area of necessary skills for those guiding or conducting assessment on the campuses.

The ratings of particular standards provided some insight into the current status of assessment, and assessment as a process. The standards identified as important in the initial planning focused on the need to address the questions of what and why to assess within the specific context of each assessment, e.g., consideration of the purpose of assessment, and clarification of the objectives or mission of the specific institution. This approach supports designing and adapting programs to the setting rather than adopting assessment programs from other institutions, as was common in early quality assessment activity.

The inclusion of quantitative information but the absence of qualitative information in the list of most important standards for the collection of the data may be related to the professional training, orientation, and practice of those conducting quality assessment in higher education. The educational background of one-third of the participants was in organization, finance, planning or institutional research, areas which may more likely include the collection and analysis of quantitative information, e.g., cost/graduate and F.T.E. Such quantifiable data are often the bases for efficiency questions which have dominated over questions of effectiveness in higher education in the past. Few of the seventy participants identified their educational background as evaluation (n=3) or an area within the social sciences (n=7), backgrounds that might include training in the collection and analysis of qualitative information.

The recency of quality assessment activity in higher education, and the consequential lack of experience with the utilization of information in the context of quality assessment may account for failure of the research to identify standards that are very important in the final phase of an assessment, i.e, using the findings. This explanation is consistent with an earlier review of literature which indicated that standards emphasized by various authors related almost exclusively to

the early phases of assessment [19]. The higher proportion of re-ratings found in the Phases IV and V may also indicate uncertainty and some of the lack of understanding and experience on the part of participants with the later phases. The lack of important standards might also indicate that the inventory was incomplete in this area, i.e., the two sets of standards from which the inventory was compiled may not be a valid representation of applicable standards for utilization in quality assessment programs in higher education. The participants were, however, encouraged to add standards they deemed as missing.

The importance across the assessment phases attributed to standards that directly involved people indicated that participants recognized the need to attend to the human dimensions of the assessment process. The on-going nature of assessment, the campus-wide responsibility and involvement, and the possible consequential effects of assessment on individuals may account for the consistently important role of the standards that focus on people.

This on-going nature of assessment and the campus-wide responsibility for its conduct may also account for the lack of importance attributed by participants to the standard regarding formal obligation in assessment. The continual process of assessment was also addressed by the two standards that were added to the assessment inventory by pilot participants: AS 39 Integration of Data Bases and AS 40 Longitudinal Reporting. The view of assessment as an on-going process may account for the more prominent role of these standards in assessment than in the conduct of evaluation in other contexts such as external summative evaluations.

The research produced a small, but useful subset of standards deemed most important by experts for the conduct of quality assessment in higher education. The assessment experts considered some standards to be more important than others in the conduct of quality assessment, and judged that the standards are more important during some phases than at other points in the assessment process. This information can be valuable as phase-by-phase guidance for those conducting assessments. Although the subset included standards that were identified as of most importance, the remaining standards should not be ignored in the conduct of assessment. Because of the general lack of experience with assessment, these results may to some extent represent uninformed rather than ideal practice. Nevertheless, the findings can also serve to inform and guide future discussions and research in the sound practice of quality assessment in higher education.

REFERENCES

1. Association of American Colleges. Integrity in the College Curriculum: A Report to Academic Community. Washington, D.C.: Association of American Colleges, 1985.

2. Boyer, C. M. & Ewell, P. T. State-based Approaches to Assessment in Undergraduate Education: A Glossary and Selected References. Denver, Colo.: Education Commission of the States, 1988.

3. Braskamp, L.A. "Discussion on Assessment in Higher Education." Evaluation Practice, 10 (November 1989), 51-56.

4. Davis, B. G. "Demystifying Assessment: Learning from the Field of Evaluation." In Achieving Assessment Goals using Evaluation Techniques, edited by P. J. Gray. New Directions for Higher Education, no. 67. San Francisco: Jossey-Bass, 1989, pp. 5-20.

5. Delbecq, A. L., Van de Ven, A. H. & Gustafson, D. H. Group Techniques for Program Planning: A Guide to Nominal Group and Delphi Processes. Middleton, Wisc.: Green Briar Press, 1986.

6. Education Commission of the States. Transforming the State Role in Undergraduate Education. Washington, D.C.: Education Commission of the States, 1986.

7. Evaluation Research Society Standards Committee. "Evaluation Research Society Standards for Program Evaluation." In Standards for Evaluation Practice, edited by P. Rossi. New Directions for Program Evaluation, no. 15. San Francisco: Jossey-Bass, 1982, pp. 7-19.

8. Ewell, P. T., Finney, J. and Lenth, C. "Filling in the Mosaic." AAHE Bulletin, 42 (April 1990), 3-5.

9. Folger, J. "Assessment of Quality for Accountability." In Financial Incentives for Academic Quality, edited by J. Folger. New Directions for Higher Education, no. 48. San Francisco: Jossey-Bass, 1984, pp. 75-85.

10. Hartle, T.W. "The Growing Interest in Measuring the Educational Achievement of College Students." In Assessment in Higher Education edited by C. Adelman, pp. 1-9. Washington, D.C.: The American Association for Higher Education, 1985.

11. Joint Committee on Standards for Educational Evaluation. Standards for Evaluations of Educational Programs, Projects, and Materials. New York: McGraw-Hill, 1981.

12. McCoy, M. & Smith, H. L. "Changing Academic Environments: Implications for Our Profession." General Session Presentations of the 26th Annual Forum. Tallahassee, Florida: The Association for Institutional Research, June 1986.

13. Moore, C. J. Group Techniques for Idea Building. Newbury Park, Cal.: Sage Publications, 1987.

14. National Institute of Education Study Group on the Conditions of Excellence in American Higher Education. Involvement in Learning: Realizing the Potential of American Higher Education. Washington, D.C.: National Institute of Education, 1984.

15. Sell, G. R. "Making Assessment Work: A Synthesis and Future Directions." In Achieving Assessment Goals using Evaluation Techniques, edited by P. J. Gray. New Directions for Higher Education, no. 67. San Francisco: Jossey-Bass, 1989, pp. 109-120.

16. Southern Regional Education Board, Commission for Educational Quality. Access to Quality Undergraduate Education. Atlanta, Ga.: Southern Regional Education Board, 1986.

17. Terenzini, P. T. "Assessment with Open Eyes: Pitfalls in Studying Student Outcomes." Journal of Higher Education, 60 (May 1989), 644-664.

18. Thomas, A. M. "Identification of the Standards for the Conduct of Quality Assessment in Higher Education using a Delphi Approach." Ph.D. dissertation, University of Minnesota, 1990.

19. ----------. "Using an Evaluation Model to Guide Development of a Quality Assessment Program." Paper presented at the Annual Forum of the Association for Institutional Research, Phoenix, Arizona, May 1988.

Appendix

Assessment Standards (AS)

AS 1 **Human Interactions:** Those conducting the assessment should respect human dignity and worth in their interactions with other persons associated with an assessment, and should show respect for cultural and social values of participants.

AS 2 **Described Purposes and Procedures:** The purpose of the assessment should be clarified; and procedures selected should be appropriate for the purpose. Descriptions of the purpose, procedures and range of activities should be in enough detail, so that they can be identified and assessed.

AS 3 **Fiscal Responsibility:** The allocation and expenditure of assessment resources should reflect sound accountability procedures and otherwise be prudent and ethically responsible.

AS 4 **Formal Obligation:** What is to be done, how, by whom, and when should be agreed to in writing by the formal parties to an assessment. This obligates the parties to adhere to all conditions of the agreement or formally renegotiate it (includes the release or information).

AS 5 **Assessment Side-effects:** Those conducting the assessment should bring to the attention of decision makers and other relevant audiences suspected side effects (positive or negative) of the assessment process itself.

AS 6 **Full and Frank Disclosure:** Oral and written material containing gathered information should be open, direct, and honest in the disclosure of pertinent findings, including the limitations of the assessment.

AS 7 **Targeted Dissemination:** Persons, groups, and organizations who have contributed to the assessment should receive feedback appropriate to their needs. When appropriate, check draft reports with representatives of audiences for clarity and accuracy; consider tailoring dissemination for specific audiences using a variety of formats and approaches.

AS 8 **Report Dissemination.** Assessment findings should be disseminated to clients and other right-to-know audiences, so that they can assess and use the findings.

AS 9 **Evaluator Credibility:** The persons conducting the assessment should be both trustworthy and competent to perform the assessment.

AS 10 **Context Analysis:** The context in which the assessment object exists should be examined in enough detail to identify the conditions that may affect its functioning.

AS 11 **Practical Procedures:** The assessment procedures should be practical and disruption kept to a minimum.

AS 12 **Audience Cooperation:** The necessary cooperation of those directly involved in the assessment as well as other audiences should be planned and assurances of cooperation obtained. Efforts should be made to assist audiences to develop realistic expectations for assessment and prioritize the information requests.

AS 13 **Cost Effectiveness:** An estimate of the cost of the proposed assessment and, where appropriate, of alternatives should be provided. Agreement should be reached at the outset that the assessment is likely to produce information of sufficient value, applicability, and potential use to justify its cost.

AS 14 **Recommendations:** Those responsible for making recommendations about courses of action, should consider and state the relative importance of the recommendations and the probable effectiveness and costs of the recommended courses of action. Recommendations should be clearly distinguished from the findings of the assessment on which recommendations were made.

AS 15 **Assessment Impact:** Attention should be given to actions that would encourage follow-through by members of the audiences.

AS 16 **Information Scope and Selection:** Information collected should be of such scope and selected in such ways as to address pertinent questions about the object of the assessment and be responsible to the needs and interest of specified audiences.

AS 17 **Report Timeliness:** Release of reports should be timely. Be responsible when possible, to audiences' changing timetables.

AS 18 **Defensible Information Sources:** The sources of information should be described in enough detail to assess the adequacy of the information. Recognize the complementariness of qualitative and quantitative data, and the importance of multiple measures.

AS 19 **Political Viability:** The assessment should be planned and conducted with anticipation of the different positions of various interest groups.

AS 20 **Balanced Reporting of Findings:** The assessment should be complete and fair in its presentation of the findings, e.g., balanced reporting of data that may lead to conclusions about strengths as well as weaknesses, and inclusion of intended as well as unintended outcomes.

AS 21 **Reliable Measurement:** The information-gathering instruments and procedures should be chosen or developed and then implemented in ways that will assure that the information obtained is sufficiently reliable for the intended use.

AS 22 **Justified Judgments and Recommendations:** Those responsible for making judgments and recommendations should explicitly justify such conclusions to enable audiences to assess them. Assumptions should be explicitly acknowledged. Any reports of such groups should include plausible alternative explanations and judgments using the findings, and distinguish among objective findings, opinions, judgments, hunches, and speculation.

AS 23 **Rights of Human Subjects:** Assessments should be designed and conducted to protect and respect the rights and welfare of the human subjects.

AS 24 **Analysis of Quantitative Information:** Quantitative information in an assessment should be appropriately and systematically analyzed to ensure supportable interpretations. Report potential weaknesses in the data collection and describe possible influences on findings.

AS 25 **Object Identification:** The purposes and characteristics of the object of the assessment to be addressed in the assessment should be specified as precisely as possible, e.g., program objectives or institutional mission statement.

AS 26 **Public's Right to Know:** The formal parties to an assessment should respect and assure the right to know of various audiences, within the limits of other related principles and statutes, such as those dealing with public safety and the right to privacy. Restrictions, if any, on access to the data and results from an assessment should be clearly established and agreed to at the outset.

AS 27 **Feasibility:** The feasibility of undertaking the assessment should be estimated either informally or through formal analysis and include such factors as needed cooperation, availability of time, money, and expertise, and administrative, fiscal, and legal constraints.

AS 28 **Valid Measurement:** The information-gathering instruments and procedures should be chosen or developed and then implemented in ways that will assure that the interpretation arrived at is valid for the given use.

AS 29 **Responsibility for Use:** Those responsible for the assessment should try to anticipate and prevent misinterpretations and misuses of the information, and when appropriate, make follow-up contacts with users, provide rebuttals of misinterpretation, and promote an open exchange of information. Guard against parties using collected data for purposes different from those understood by persons providing the information.

AS 30 **Conflict of Interest:** Conflict of interest, frequently unavoidable, should be dealt with openly and honestly.

AS 31 **Audience Identification:** Audiences involved in or affected by the assessment should be clearly identified.

AS 32 **Unit of Analysis:** The unit of analysis should be appropriate to the purpose of the assessment, the types of conclusions to be drawn and the way the data are collected.

AS 33 **Assessment Flexibility:** Flexibility should be maintained during the assessment to respond to changing timetables and other needs. However, provision should be made for the detection, reconciliation and documentation of any departures from the original assessment design, whether unplanned or a planned response to needs.

AS 34 **Report Clarity:** Reports of results should describe the object being assessed and its context, and the purposes, procedures, and findings of the assessment. Limitations caused by constraints on time, resources, and data availability should also be stated.

AS 35 **Objective Reporting of Results:** The assessment procedures should provide safeguards to protect the assessment findings and reports against distortion by the personal feelings and biases of any party to the assessment. When appropriate, draft reports should be checked with representatives of audiences for clarity and accuracy.

AS 36 **Valuational Interpretation:** Those responsible for making judgments about quality and recommendations should carefully describe the perspectives, procedures, and rationale used to interpret the findings to insure that the bases for value judgments are clear.

AS 37 **Data Quality Assurance and Control:** The data collected, processed, and reported in an assessment should be reviewed and corrected. Analysis of the sources of error should be undertaken.

AS 38 **Analysis of Qualitative Information:** Qualitative information in an assessment should be appropriately and systematically analyzed to ensure supportable interpretations. Report potential weaknesses in the data collection and describe possible influences on conclusions.

AS 39 **Integration of Data Bases:** Pertinent, previously collected data that may exist in various sources at the institution or that exist outside the institution such as available comparable data should be utilized and linked to any newly collected data.

AS 40 **Longitudinal Reporting:** Ongoing, update reports of findings should be used to keep audiences informed about assessment activity and results.

FOOTNOTES

[1] American Association of Higher Education (AAHE), American Evaluation Association (AEA), Association for Institutional Research (AIR), Association for the Study of Higher Education (ASHE), American Educational Research Association) AERA, and National Center for Higher Education Management Systems (NCHEMS).

TABLE 1

Number of Standards Rated at Each Level of Importance within Each Assessment Phase

	Assessment Phase				
Importance Category	I	II	III	IV	V
Very important	2	4	2	5	0
Generally important	3	16	8	18	13
Somewhat important	8	5	4	3	8

TABLE 2

Final Results of Importance Ratings by Phase

Assessment Standards in General Chronological Order	Importance by Assessment Phase I	II	III	IV	V
# AS 25 Object Identification	***	**	-	-	-
# AS 2 Described Purposes and Procedures	***	***	-	**	**
AS 27 Feasibility	**	**	-	-	-
AS 12 Audience Cooperation	*	**	-	-	-
AS 13 Cost Effectiveness	*	**	-	-	-
AS 16 Information Scope and Selection	*	**	*	-	-
AS 10 Context Analysis	**	**	-	-	*
# AS 23 Rights of Human Subjects	*	***	**	*	-
# AS 32 Unit of Analysis	*	**	**	*	-
# AS 1 Human Interactions	*	**	*	**	**
# AS 30 Conflict of Interest	*	**	-	**	**
# AS 31 Audience Identification	**	**	-	*	*
AS 26 Public's Right to Know	*	*	-	**	-
# AS 9 Evaluator Credibility	-	**	***	**	*
AS 3 Fiscal Responsibility	-	**	-	-	-
AS 19 Political Viability	-	**	-	-	-
# AS 21 Reliable Measurement	-	***	**	-	-
# AS 28 Valid Measurement	-	***	**	-	-
AS 11 Practical Procedures	-	**	**	-	-
AS 39 Integration of Data Bases	-	**	**	-	-
AS 33 Assessment Flexibility	-	*	*	-	-
AS 18 Defensible Information Sources	-	**	*	**	-
AS 29 Responsibility for Use	-	*	-	**	**
# AS 35 Objective Reporting of Results	-	*	-	***	*
AS 5 Assessment Unintentional Outcomes	-	*	-	**	**
AS 37 Data Quality Assurance and Control	-	-	**	**	-
AS 38 Analysis of Qualitative Information	-	-	**	**	-
# AS 24 Analysis of Quantitative Information	-	-	***	**	*
# AS 34 Report Clarity	-	-	-	***	*
# AS 6 Full and Frank Disclosure	-	-	-	***	**
# AS 7 Targeted Dissemination	-	-	-	***	**
# AS 20 Balanced Reporting of Findings	-	-	-	***	**
AS 8 Report Dissemination	-	-	-	**	**
AS 14 Recommendations	-	-	-	**	**
AS 15 Assessment Impact	-	-	-	**	**
AS 22 Justified Judgments & Recommendations	-	-	-	**	**
AS 36 Valuational Interpretation	-	-	-	**	**
AS 17 Report Timeliness	-	-	-	**	*
AS 40 Longitudinal Reporting	-	-	-	**	*
AS 4 Formal Obligation	-	-	-	-	-

*** = Very important
** = Generally important
* = Somewhat important
# = Included in final subset of standards