REPORT OF THE ACADEMIC SENATE STUDENT EVALUATION OF TEACHING (ASSET) COMMITTEE

Executive Summary

The ASSET Committee has spent approximately one year familiarizing itself with both the issues and current practices of the evaluation of teaching (ET) as found at Wayne State University and as reported in the relevant literature and practices at other universities.

As a result of its inquiries, the ASSET Committee has reached several conclusions about current ET practices at WSU as well as the place of student assessment of teaching (SET) in an overall assessment of teaching program. The committee has developed a modest set of nineteen (19) recommendations that are designed to support a broader definition and assessment of teaching than currently employed at Wayne State University, improve the quality of the evidence used in assessing teaching effectiveness (referred to in the literature as summative evaluation), strengthen the WSU SET form to take advantage of its heretofore largely unused capability to provide information useful to improve and enhance teaching skill (referred to in the literature as formative evaluation) and link the ET process to the proposed Office of Teaching and Learning. Ten of the recommendations are designed to expand our current teaching assessment program into a system for the evaluation and improvement of teaching (Recommendations 1-10). Five recommendations provide for enhancements of the WSU SET form (Recommendations 11-15) and four recommendations are proposed to improve the interpretability of SET data (16-19). Overall, the recommendations are designed to position the current WSU SET process within a broader conception of evaluation of teaching and to be responsive to and supportive of the Wayne State University Strategic Plan--Instruction Initiatives.

The thrust of the committee's recommendations is the establishment of a comprehensive system for the evaluation and improvement of teaching and learning. The feasibility of successfully implementing the proposal is judged to be high because the proposed plan is based largely upon resources or essential elements that are in place at WSU. The system would utilize the existing SET program with modest enhancements and would take advantage of the WSU promotion and tenure policy that supports a broadened definition of teaching. The plan would help to focus the mission of the Office of Teaching and Learning. The proposal is supportive of the institution's teaching mission and is consistent with the long range program goals for that mission.

May 17, 1995

The committee would like to thank Angela Wisniewski of the Academic Senate Office for the considerable administrative support she provided to the committee during its deliberations.
REPORT OF THE ACADEMIC SENATE STUDENT EVALUATION OF TEACHING (ASSET) COMMITTEE

TABLE OF CONTENTS

I. BACKGROUND

II. INTRODUCTION
   Committee Membership, Committee Charge

III. APPROACH TAKEN

IV. FINDINGS AND RECOMMENDATIONS
   A. CREATE A SYSTEM FOR THE EVALUATION OF TEACHING AND IMPROVEMENT OF TEACHING AND LEARNING
   B. CREATE A SYSTEM OF STUDENT EVALUATION OF TEACHING
      PART 1. MODIFY THE CURRENT SET INSTRUMENT
      PART 2. CREATE A SYSTEM FOR INTERPRETING SET DATA

V. APPENDICES
   A. BIBLIOGRAPHY OF CITED AND REVIEWED LITERATURE
   B. SURVEY LETTER TO INSTITUTIONS SURVEYED
   C. INSTITUTIONS HAVING EXEMPLARY STUDENT ASSESSMENT OF TEACHING PROGRAMS
   D. LIST OF EXAMPLE EVALUATION OF TEACHING HANDBOOKS
   E. COPIES OF CURRENT WSU SET AND PROPOSED MODIFICATIONS TO SET FORMS
I. Background

The information below is to inform the reader of the conditions under which the ASSET Committee of the Academic Senate deliberated and formulated its report.

On May 4, 1994, the Academic Senate adopted a resolution calling for the Senate to participate in the development of an evaluation instrument for student assessment of teaching (SET). The Academic Senate requested participation of Wayne State University (WSU) Administration in this work.

The position of the WSU Administration as of May, 1994, was that the student evaluation of teaching instrument was a matter of collective bargaining and outside the province of the Academic Senate. To clarify the role of governing bodies such as the Academic Senate in matters related to collective bargaining, Wayne State University and the American Association of University Professors (AAUP) signed a letter of agreement (August 1, 1994) wherein consultative bodies may deliberate and develop informational reports about matters relating to collective bargaining if the University and the AAUP notify the consultative body that such matters are subject to collective bargaining. Under this agreement, reports about collective bargaining matters developed by such bodies do not constitute advice to the University under University Statutes.

A letter in October, 1994, from the President of the AAUP to the President of the WSU Academic Senate stated the AAUP’s concurrence with the Administration that student evaluation of teaching is a matter of collective bargaining. This notification, along with President Adamany’s similar notification of May, 1994, put the deliberations of the ASSET committee under the August 1, 1994, agreement.

This arrangement was beneficial to the committee. The information subsequently provided to the committee by President Adamany in the form of numerous university reports and discussions with key university administrators was an important and substantive contribution to the committee’s understanding of how the current SET program was established and functions in the evaluation of teaching (ET) at WSU. These understandings, integrated with insights gained from literature reviews, discussions with evaluation experts and review of programs at other institutions were essential to the development of informed recommendations for the further development of the evaluation of teaching program at WSU.
II. Introduction

The ASSET Committee has spent approximately one year familiarizing itself with both the issues and current practices of the evaluation of teaching (ET) as found at Wayne State University and as reported in the relevant literature and practices at other universities. While the literature review was in depth, the examination of practices at other universities was limited to institutions having one or more of the following characteristics: 1) formal offices or centers supporting programs of teaching improvement; 2) standardized student evaluation of teaching (SET) instruments comparable to the Wayne State University assessment form or 3) national recognition for the quality of their SET instruments and systems.

As a result of its inquiries, the ASSET Committee has reached some conclusions about current ET practices at WSU as well as the place of SET in an overall assessment of teaching program. The committee has developed a modest set of recommendations which are designed to support a broader definition and assessment of teaching, improve the quality of the evidence used in assessing teaching effectiveness (referred to in the literature as summative evaluation; Bloom, Hastings and Madaus, 1971), strengthen the WSU SET form to take advantage of its heretofore largely unused capability to provide information useful to improve and enhance teaching skill (referred to in the literature as formative evaluation), and link the ET process to the proposed Office of Teaching and Learning. Finally, the recommendations are designed to position the current WSU SET within a broader conception of evaluation of teaching and to be responsive to and supportive of the Wayne State University Strategic Plan, Instruction Initiatives, Goal 1 (improve teaching and learning) and Goal 4 (incorporating richer, more varied and objective evidence in personnel decisions).

The literature citations supporting the committee recommendations below are not the totality of the literature reviewed by the committee. References cited were selected because of their compelling nature, currency or representativeness of a larger set of references relevant to a given recommendation. A bibliography of all literature cited and reviewed can be found in Appendix 5.

Committee Membership

- Gallagher, Richard
- Bledsoe, Timothy
- Lenarcic, Neal
- Lichtman, Cary
- Oermann, Marilyn
- Pope, Jennifer
- Richards, Doris

Family Medicine, Medicine, Chairperson
Political Science, Liberal Arts
Student Representative
Psychology, Science
Nursing
Student Representative
Music, Fine, Performing and Communication Arts

Two additional faculty members were appointed to the committee but were unable to serve due to conflicts.
Academic Senate Charge to Committee

Examine the current standardized student assessment of teaching program (SET) in place at Wayne State University within the broader context of the overall assessment of teaching; assess the extent to which the SET program is meeting its intended objectives; and make recommendations if and as appropriate.

III. Approach Taken

The committee examined the existing WSU SET in light of its intended purposes. The intended uses of SET data as described on the SET instrument are:

a. improvement in the quality of instruction
b. annual decisions made on salaries, and
c. faculty promotion and tenure decisions

The committee noted that students also use SET data for purposes of class selection.

The committee approached its task by undertaking a search of the relevant literature; by examining ET practices at other Universities noted for their exemplary programs; and by examining how well the WSU SET instrument and program was achieving its intended goals.


To gain an understanding of the "state of the art" of ET, current scholarly thinking and exemplary systems of teacher evaluation, the committee reviewed the existing relevant literature, conducted a survey of ET practices at other institutions, and examined in some detail the ET programs of a selected group of institutions known for their exemplary programs of using SET data in the ET process. Personal conversations were held with three nationally recognized experts in ET, professors Larry A. Braskamp, John A. Centra and David M. Irby. The institutions surveyed are the universities and colleges holding membership in the Professional and Organizational Development Network. Members of this organization all have programs of faculty development focused on the improvement of teaching and learning. Members were queried about the ET and SET programs at their institutions as well as the policies governing such programs. Of the 80 institutions contacted, 40 responded. (See letter of inquiry in Appendix 2. The membership of this organization and the ET and SET material accompanying the survey replies can be found in the Academic Senate ASSET files, file NO.2).

A more detailed understanding of the programs, instruments and policies of a selected number of institutions identified in the literature as having exemplary ET and SET programs was achieved through review of program documents and personal conversations with administrators of these programs. A list of these institutions can be found in Appendix 3. Associated ET documents can be found in the Academic Senate ASSET files, file NO. 3).
The committee's understanding of the history of the current SET program came from an examination of the 1986 report of the Student Classroom Teaching Evaluation Committee Status Report, a joint faculty/administrative committee, whose recommendations led to the present program. A more detailed understanding of both the development as well as the current functioning of the WSU SET was acquired from useful discussions between the committee and Provost Marilyn Williamson and Mr. John A. Crusoe, Executive Director and Dr. Thomas J. Wilhelm, Associate Director, Counseling Services. Additional information was obtained with a limited number of telephone interviews to members of the WSU academic community.

Other documents that proved to be of great benefit to the committee were thoughtfully and generously provided by President Adamany. These documents included:


2). Student Evaluation of Teaching Report 1.0 (June 3, 1991), A Factor Analysis of 1989 SET data

3). Report 8.0, Winter 1993 Faculty Survey Results. An Administration survey of faculty views of the SET program.

4). WSU Promotion and Tenure Procedures and Factors for Faculty (January, 1994).

A copy of each of the above documents can be found in the Academic Senate Asset files, file NO. 1).

IV. Findings and Recommendations

The committee found both strengths and weaknesses in WSU current ET and SET practices and programs.

On the positive side the committees' findings clearly show that WSU has or will shortly have in place the essential elements to create an ET program of national distinction. These elements include University promotion and tenure criteria and standards that support the development and use of a wide base of evidence on teaching effectiveness, a proposed Office of Teaching and Learning to work with faculty in the enhancement and documentation of teaching skill, and an in-place SET program that can be strengthened to better support the faculty development process of improving teaching.

On the other hand, several serious deficiencies were noted. The current day-to-day ET practices at WSU are based on a narrow and now outdated definition of teaching, are too dependent on SET data, are not linked to programs of faculty development, and fail, in some instances, to meet recognized standards of assessment practices.
The benefits of making the changes proposed are numerous. A broader definition of teaching and its associated measures will assist the institution in refining its expectations, standards, and means of assessment for itself and its teaching faculty; it also will more effectively document the full range of the faculty teaching contribution. Modifications of the existing WSU SET, as recommended below, will provide a substantially strengthened diagnostic tool for use by faculty and the Office of Teaching and Learning for teaching improvement; it can, in fact, be the beginning of an institution-wide program of faculty development. Consultants in the Office of Teaching and Learning will be able to offer judgements about an instructor's teaching based, in part, on the same kind of data used by colleagues and academic supervisors. Personnel-decisions will also be strengthened by a better and more easily interpreted measurement of global teaching, along with the evidence provided from a broader range of measures employed in addition to SET data.

A. Create a System for the Evaluation of Teaching and the Improvement of Teaching and Learning

1. Recommendation: Broaden the definition of the teaching role at WSU and its assessment to consider the teaching aspects of advising, supervising, guidance and mentoring; developing learning activities; actively pursuing a program of faculty development in teaching; as well as the delivery aspects of the instructional role which predominates at present. Measures of productivity exist or can be developed for each aspect of the broadened role.

Rationale: While this report recommends that WSU develop an institutional definition of teaching that goes beyond the delivery act of the teaching role, it is not recommending a specific definition. The committee believes that this topic is being addressed, at least in part, by Provost's task forces addressing issues of evaluating teaching and the establishment of an Office on Teaching and Learning. It is anticipated that both of these task forces will address the need and rationale for a broadened definition of teaching. This topic can also be pursued by the ET joint oversight committee proposed below. Expanded definitions and conceptions of teaching can be found in the literature (Cashin, 1993; Boyer 1990; Centra, 1993; Irby, 1993; Braskamp, 1994). For purposes of illustration, a table depicting common activities included in a broader definition of teaching is shown below (Braskamp, 1994, p. 41).
[Table 1.] The Work of Teaching

Instructing

Instructing students in courses, laboratories, clinics, studio classes
Instructing participants in workshops, retreats, seminars
Managing a course (grading, maintaining student records, planning learning experiences).

Advising, Supervising, Guiding, and Mentoring Students

Supervising students in laboratories, fieldwork
Advising and mentoring students (career, academic, personal counseling referral)
Supervising teaching assistants
Supervising students with internships and clinical experiences
Supervising students in independent study
Advising students in their senior research projects, theses and dissertations.

Developing Learning Activities

Developing, reviewing, and redesigning courses
Developing and revising curriculum
Developing teaching materials, manuals, software
Developing and managing [distance learning courses]
Developing computer courses
Conducting study-abroad programs

Developing as a Teacher

Evaluating teaching of colleagues
Conducting instructional and classroom research
Attending professional development activities.

2. Recommendation: SET data and other evaluation of teaching data should be linked to the concept and process of faculty development.

Rationale: In order to succeed, programs of faculty development for teaching skill will need to be supported by institutional goals, policy and resources.

The proposed Office of Teaching and Learning is one such resource. The Office has the potential to become a source of faculty and departmental consultation on the improvement of instruction, identification and development of other measures of effective teaching, the interpretation of data on teaching effectiveness and the effective presentation of the teaching record where required. Effective consultation on these matters will be dependent on the expertise of its staff, the philosophy of the office and the willingness of faculty and departments to utilize its services.
Faculty development of teaching entails serious self-assessment. Braskamp (1994), in commenting on this, has stated that in its best role, faculty assessment is a strategy for thorough self-examination, reflection, discussion and building. Such self-reflection, according to Stake (1967), an insightful scholar in the field of evaluation, requires some privacy. He has stated that "evaluation is everybody's business, but not everybody else's business". Commenting on this, Braskamp has stated that "...faculty need to balance individual and institutional purposes. If the institutional purpose dominates, a climate of control rather than commitment may be created." This, in part, is why the committee is recommending (see recommendation 6 below) that the Instructor Feedback Diagnostic items be reserved under relatively private conditions for the personal, possibly threatening task of self-assessment. The productive use of the services of the proposed Office of Teaching and Learning is much more likely under supportive as contrasted to threatening conditions.

3. **Recommendation**: The SET program must require and employ SET instruments that are appropriate and administratively feasible for the type of teaching/learning being assessed.

**Rationale**: The proposed revised WSU SET form, like its predecessor, was designed for the most prevalent form of instruction at WSU, classroom teaching. Classroom teaching is characterized by a unit of instruction that is largely taught by a single teacher who has the prime or major responsibility for delivering the instruction to a group of learners over a university semester. This form of teaching is largely didactic. This pattern does not fit all teaching situations in the university. The committee is aware that there are courses being taught at WSU based on team-teaching, problem-based-teaching and forms of clinical teaching where the current and proposed revised SET form cannot justifiably or administratively be used. One danger in requiring the use of one instrument designed for a particular form of teaching to other teaching methods (aside from the absence of any methodological justification) is the stifling of teaching and curriculum creativity (Wilson, 1988). This poses a serious problem when schools are faced with curriculum accrediting agencies requiring innovative teaching. In those instances where a form of teaching does not lend itself to evaluation with the classroom SET form, WSU policy and the ET system should encourage faculty and administration to develop instruments and procedures appropriate for the teaching mode in question.

4. **Recommendation**: The oversight committee should conduct a campus wide survey of the extent and type of non-classroom teaching in place or under contemplation by the various teaching units at WSU.

**Rationale**: This information will be useful in the planned development and possible sharing of appropriate SET forms. The information gathered from this survey will also be of use to the proposed Office for Teaching and Learning.

The literature on student assessment of teaching provides minimal help or guidance on this matter. Our survey of other institutions with exemplary SET programs as well as our discussions with SET authorities indicate that most institutions are either ignoring the assessment of these forms of instruction or are in the early stages of finding appropriate assessment techniques. The committee does recommend that in teaching situations where traditional SET forms are not feasible, special consideration be given to the use of other measures addressing a broadened definition of teaching.
5. Recommendation: Develop a faculty-administrative-student handbook explaining the institutional purpose of ET in the institution, the policies governing the assessment, the methods to be employed and the responsibilities of individual faculty, departments and schools.

Rationale: The handbook would provide guidance to faculty members in methods of documenting their teaching practices; identify institutional resources available to assist faculty, departments and administrators in this process; and provide guidance in the interpretation of ET data, describing both the usefulness as well as the limits of such data in decision making. Some institutions have tied the information in such guides to the promotion and tenure policies of the institutions. A list of sample handbooks from other institutions can be found in Appendix 4. Copies of the handbooks on this list can be found in the Academic Senate Office Asset files, file NO. 4.

6. Recommendation: The results of the student evaluation of faculty teaching effectiveness should be distributed as follows:

DISTRIBUTION OF RESULTS

<table>
<thead>
<tr>
<th></th>
<th>Faculty</th>
<th>Admin</th>
<th>Student</th>
</tr>
</thead>
<tbody>
<tr>
<td>STUDENT SELF EVALUATION</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>COURSE EVALUATION I &amp; II</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>INSTRUCTOR FEEDBACK DIAGNOSTICS &amp; INSTRUCTOR SUPPLIED ITEMS</td>
<td>X</td>
<td>Dept. Chair for Fac. Dev. only*</td>
<td></td>
</tr>
<tr>
<td>DEPARTMENT &amp; SCHOOL SCHOOL SUPPLIED ITEMS</td>
<td>X</td>
<td>X*</td>
<td></td>
</tr>
<tr>
<td>STUDENT OPEN-ENDED COMMENTS</td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Rationale: The above recommendation reserves diagnostic information about teaching and the course for the faculty. This feedback to the teacher is intended for improvement of instruction, not for personnel decisions. Research has established the multidimensional nature of student ratings (Cashin, 1990 and Kulik and McKeachie, 1975). Further, research has established that the best SET measurement of overall teaching effectiveness is derived from the global type question; and the most useful SET questions for improving teaching effectiveness comes from specific diagnostic type questions. As noted above, the specific diagnostic items are most useful for teaching improvement because they attempt to measure specific teaching behaviors or course characteristics (Braskamp, 1994).
7. **Recommendation:** The committee endorses the elements of WSU policy with respect to how the SET form is to be administered as described in the WSU document entitled "Policy on Student Evaluation of Teaching, Fall, 1993" unless there are specific alternative recommendations in this report.

**Rationale:** The procedures mandated for administration of the SET form follow closely the research-based recommendations in the literature. Examples of current WSU policy supported by this literature include providing for anonymous ratings, providing class time in the last two weeks of instruction and excusing the instructor from the room while students fill out the forms.

8. **Recommendation:** Achieve a goal of 90% of all classes. This means the timely administration, scoring and return to the instructional units and teaching faculty within the intervals established by WSU policy.

**Rationale:** The committee, in its informal survey, found apparently wide variability with respect to the percentage of classes in compliance with current SET policies of form administration and feedback to the instructor. Our informal estimates put compliance as varying between 60% to 90%.

9. **Recommendation:** The proposed revised SET program must be open to orderly, informed, planned and tested modifications. This means that the proposed revision and future subsequent revisions can not be arbitrarily introduced but must be potentially related to improvement in the quality of SET and Evaluation of Teaching (ET) data, decision making or the improvement of instruction. Recommendations should be treated as hypotheses with potential benefit and should be tested before they are adopted as a functioning part of the SET program.

**Rationale:** The system must have sufficient stability to assure orderly system development and assessment. The SET program must be seen as a program in development.

10. **Recommendation:** The creation of a joint faculty-administrative-student oversight committee. Functions of the committee include but are not limited to the following:

a. Determine the resources needed to implement a revised SET program.

b. Oversee implementation of the program beginning in Fall, 1995.

c. Plan and conduct formative evaluation(s) of the effectiveness of the implementation of the revised SET.

d. Plan and conduct summative evaluation(s) of the planned and unanticipated results of implementing the revised SET.

e. Plan and conduct the recommended institution-wide survey of existing and planned non-classroom-type teaching and associated SET plans and instruments.
f. Receive, evaluate and where justified make recommendations for changes in the SET forms, administration, interpretation and data use.

g. Undertake the development of an expanded Instructor-Feedback Diagnostic section of SET.

h. Plan and oversee the development of policy, methods handbook.

B. Create a System of Student Evaluation of Teaching

General Recommendation: We propose the building of a comprehensive system of student evaluation of teaching that will serve to create more accurate measures of student perceptions, reduce or eliminate known biases to student ratings, and provide for a more understandable and coherent mechanism of judging teaching performance.


Overview: The current SET instrument at Wayne State is reasonably well-suited for the summative evaluation of teaching. Indeed, four of the 15 substantive survey items could be classified as "global" in nature; that is, they tap an overall student perception of the classroom experience. On the other hand, the current section entitled "Student Profile" seems less useful and the number of items on the instrument that might serve to assist in faculty development appear too limited.

Recommendation 11: Reduce the number of student profile items.

Rationale: The section on the instrument titled "Student Profile" contains items that might be useful in conducting analyses of SET results or for norming SET scores in some way. However, it seems that limited use is currently made of most items in this section. While we are not opposed to gathering information that may prove useful at some future time, the committee believes that for the sake of efficiency items that go unused or for which there are no expressed intentions of future use should be dropped from the questionnaire. Subject to the consent of all interested parties, we propose deleting items A, B, C, E, and G under the Student Profile heading. For a class-level (as opposed to an individual-level) analysis, items A and B are probably better measured by the level of the course (100, 300, 500, 700, etc.) for which students are completing questionnaires. Item C is better assessed through the initial interest items described below. Items E and G may offer little promise for subsequent analysis of SET results (Fieldman, 1978 and Marsh, 1984). The new section tentatively labeled "Self Evaluation" will contain items D and F as well as two new indicators described below.

Recommendation 12: Re-categorize substantive items on the questionnaire.
Rationale: The current categories employed on the questionnaire appear to result from a misinterpretation of a factor analysis conducted several years ago. At that time, the two-factor solution was interpreted as representing an "Instructor" dimension and a "Course" dimension. A careful review of that analysis, along with a new analysis conducted on 1994 data (see Academic Senate Asset files, file No.5) and a thorough review of the literature on student assessment of teaching, indicate a reinterpretation of the initial findings. The literature identifies a "summative" dimension of student evaluations that seems to describe more accurately the current "Course Evaluation Items", and a "diagnostic" dimension that we currently label "Instructor Evaluation Items." Hence, on the proposed new form we propose distinguishing these two dimensions, summative (or Course Evaluation) and diagnostic.

In comparing the earlier to the more recent factor analysis, we observed that one global item experienced a modest change in the way it fit with other summative items. Specifically, in the earlier analysis the current item #15 (then item #3: "Overall, the instructor did a good job teaching this course.") loaded strongly on the summative dimension as might be expected. However, the new factor analysis shows this item more distant from the other summative items, though closer to other summative items than were most of the diagnostic items. We suspect that moving what face validity would suggest to be a robust summative measure from the beginning of the questionnaire to the end caused that item to become influenced or contaminated by the diagnostic items. In other words, we suspect that as students work their way through the diagnostic items they become enlightened about aspects of teaching they should consider in arriving at a summary evaluation, and this tends to slightly color their impressions of the instructor. In a sense, this may provide a more informed or thoughtful indication of student perceptions.

As a result of this, we determined that two of the four summative evaluation items should appear before the diagnostic section, where they would be free from the influence of the diagnostic items and hence immune from any change in the content of the diagnostic section. Two summative items will appear after the diagnostics and therefore hold the potential for more informed assessments of teaching performance.

Recommendation 13: Replace one substantive item on the instrument.

Rationale: Of the 15 substantive items on the existing SET instrument, we propose removing and replacing only one, current item #1 ("The course content met or exceeded my expectations for the class"). First, responses to this question may be as much a function of student "expectations" as instructor performance. More importantly, in the initial analysis of SET data this item was so highly correlated with another summative item as to offer little unique information about student opinions. We propose replacing this item with one that would allow students to assess how much they have learned in the course ("How much have you learned from this course?"); an item that may be more meaningful for students from an assessment standpoint and that may offer a more distinct bit of new information than the current course content item.

Recommendation 14: Revise the response categories on the questionnaire.
Rationale: We propose two changes in the response categories: (A) reducing the number of categories from seven to five; and (B) replacing the Likert-style summative questions with ones employing polar-anchored responses. Regarding the former, we are adopting multiple-item scales for critical measures (see below) and thus there is less need for a large number of response categories for individual items. In addition, and particularly with the Likert-style responses that will remain on the diagnostic items, a five category measure may offer better potential to achieve equal conceptual distance among categories. In other words, we typically treat these as interval-level data, and this assumes equidistance between the categories... So, the conceptual distance from "Neutral" to "Disagree Somewhat" should be the same as the distance from "Disagree Somewhat" to "Disagree" or "Disagree" to "Strongly Disagree". Whether this is the case seems debatable, but a five-category scale is undoubtedly more defensible. Finally, five categories (the norm among the other SET instruments we reviewed) probably offer a level of precision in the measures appropriate to the precision of the attitudes being assessed, and therefore offer more reliable measures of student opinion.

For the four summative items (two in each of the two Course Evaluation sections), we opt for a return to anchored-response categories similar to those used on the 1989 SET form at Wayne State. We prefer the old system to the new one based on Likert-style agree-disagree questions for a variety of reasons. The old questions (roughly the ones we propose) are less wordy than the ones currently in use and they seek a more direct assessment of objects themselves rather than statements about objects. More significantly, there is more potential for ambiguity in responding in an agree-disagree format to statements containing specific normative terms such as "good" or "excellent". For example, we currently offer students the statement, "Overall I would rate this course as excellent." A student might disagree that the course was "excellent" but agree that it was "good". This may be unlikely, but most survey researchers would prefer an approach that minimizes ambiguity. Finally, and perhaps most importantly, the old system employed on the SET at Wayne in 1989 produced better variation in responses. In general, it appears that there are about ten percent more responses in the highest two categories now than there were then. Of course, this may result from campus-wide improvement in instruction, but we suspect that changes in the response categories are partly responsible. Anchored-response scales may simply produce better discrimination in student perceptions.

Recommendation 15: Study the eventual development of an expanded set of diagnostic items.

Rationale: An expanded set of diagnostic items can increase the utility of the information gained for instructional development. Many of the SET forms we reviewed provide a wider array of diagnostic indicators than are employed on Wayne State's SET form. Whereas we have 11 diagnostic measures, Syracuse University uses 22, Scranton University has 28, and Kansas State has 30. Some examples of diagnostic indicators not currently on our form include: the value of the class in stimulating intellectual interest beyond the course; the utility of textbooks in the course; the clarity with which the instructor presented material.
These additional pieces of information may prove useful in helping faculty identify their strengths and weaknesses and in revising their courses in the future. However, we think it would be wise to generate direct faculty involvement in supplementing the current selection of diagnostic indicators and therefore we propose no changes in the diagnostics at this time. A future committee should make this a priority concern.

Part 2. Create a System for Interpreting SET data.

Overview: The current procedures for incorporating SET data into personnel decisions suffer from an "overload" of possibly conflicting indicators. These indicators may be viewed differently by various decision makers as well as by instructors themselves. In addition, the literature recognizes one important bias that should be taken into account in comparing SET results across instructors who are inevitably faced with different teaching assignments. Finally, any potential indicators are prone to both measurement error and sampling error (see below), and a fair system of student evaluation of teaching should take these limitations into account.

Recommendation 16: Develop a single summary measure of teaching performance in a class, a Summary Evaluation Scale (SES). The SES will consist of the sum of the four summative (or "global") items on the questionnaire and will range from four to 20.

Rationale: As the system currently stands, data from the SET are evaluated in an almost impressionistic way by personnel decision makers. We are told that the Testing Office compiles multiple-item scales from among select SET items, but several decision makers with whom we spoke seem unaware of these scales. Otherwise, it is not clear which of several potential indicators decision makers actually use in making personnel decisions, and undoubtedly there is considerable variation among decision makers in their choice of indicators.

Adopting a single Summary Evaluation Scale would serve two important purposes. First, it would eliminate ambiguity about the bottom-line indicator of student perceptions. The SES would offer, in a single number, the mean or average student rating for an instructor in a given class. This would eliminate uncertainty and confusion about how decisions regarding teaching performance are arrived at.

Second, a single SES would provide a superior indicator of student perceptions as compared to any single indicator. Multiple-item scales are preferred to single-item measures because they reduce measurement error. By measurement error, we mean that any single question will elicit a response that is no better than an approximation of the person's true attitude. If we have three or four indicators of the same general attitude, as we clearly do in the cases of the summary indicators, we can reduce this error by combining the several indicators into a single summary indicator, thus allowing random errors to offset one another.

Recommendation 17: Categorize the Summary Evaluation Scale.

Rationale: This proposal provides a solution to likely sampling bias and over-interpretation of statistical data. It must be remembered that, on average, SET forms are completed by 70 percent of students enrolled in a course. Statistics resulting from these forms should therefore be regarded as sample statistics,
subject to potential sampling error. It is therefore unreasonable and unjustifiable to draw meaningful inferences from relatively small differences in sample means (Centra, 1993).

We propose the development of three broad categories of SES scores and the primary use of these categories for personnel decision making. Mean SES scores in the bottom 10 percent of all classes of similar initial student interest should be judged "low" in performance, the next 80 percent should be rated "moderate" in performance, and the top ten percent should be rated "high" in performance.

To allow for widespread improvement in instruction over time, the initial campus-wide norms should be used to establish a criterion-based measure for assignment into the highest category. The 90th percentile score for each initial-interest category would thus provide a yardstick for assignment in the "high" category in the future. For example, our Summary Evaluation Scale will range from four to 20. The first semester we use this system we might discover that the 90th percentile for high interest classes starts with a mean score of 17.5; for the middle-interest category the 90th percentile starts at 17.0; and for the low-interest category the 90th percentile begins at 16.5. In subsequent semesters a class in the middle-interest category would be classified "high" in performance if its mean rating was in the 90th percentile of classes of middle-interest or if the mean score for the class exceeded 17.0. It is essential to supplement a norm-based measure of teaching performance with this criterion-based measure if we are to recognize, and reward, campus-wide improvement in the quality of teaching.

**Recommendation 18: Control for initial student interest.**

**Rationale:** The single source of evaluation bias that is most widely recognized and addressed is that for initial student motivation (Cashin, 1988). Several other SET systems structure their evaluations so that in assigning ratings for Summary evaluation and for diagnostic considerations, classes are compared only to other classes with similar mean student interest. Initial student interest is measured by such indicators as, "I really wanted to take this course REGARDLESS OF WHO TAUGHT IT," and "Before enrolling, I had a strong interest in the subject matter of this course." Classes are then grouped according to their mean level of initial student interest, and comparisons are made across classes of the same interest level.

For example, Scranton University groups classes into three categories, those whose mean student interest is in the highest 25 percent of all classes, those whose mean initial student interest is in the middle 50 percent of all classes, and those classes students would like to avoid where the mean initial interest is in the bottom 25 percent of all classes. Also like the system we propose, Scranton groups classes into three categories of teaching performance, low, middle, and high. For someone teaching a high-initial-interest class to be rated "high" on performance, they have to score at least 4.64 on a summative indicator, while those teaching "middle-interest" classes need score only 4.36 and "low-interest" classes need only score above 4.1. The range is greater at the low end of teaching performance in classifying courses in the bottom category. For someone in a high-initial-interest class to score in the bottom category their mean rating on a summative measure must be below 3.47, but those in the middle-interest group must be below 3.08 and in the low-interest group below 2.64.
Ultimately, if we are going to make use of SET data in personnel decisions we must make the playing field as level as possible, insuring everyone an equal chance to score high (or low) in student ratings. Obviously we cannot structure the system so that everyone is teaching the same courses, but we can structure our evaluation system so that classes of differing initial student enthusiasm are treated differently in the ratings.

Recommendation 19: Use data from classes over two years and seek to identify patterns in SET results.

Rationale: Individual Instructors often vary in their performance over time. Personal misfortune, ill health, a particularly rambunctious class, or an initial course offering may precipitate unusually low ratings for a course. Summative evaluations should be based on typical or representative performance of the teacher. An important advantage of reducing each class taught to a single SES score is that with these more manageable data, decision makers can easily check for patterns or trends in the data over two or more years. Centra (1993) recommends the use of at least five classes for personnel decision making, and up to ten classes if class size is typically fewer than 15. A two-year window on teaching performance should be sufficient to produce five to 12 classes on which an instructor is rated. If, as an example, a two-year period yields eight evaluated classes, and six or seven of the eight receive the same mid-level rating, one might reasonably conclude that the instructor is performing at an acceptable or typical overall level compared to other teachers on campus. Similarly, an instructor with seven of eight classes rated in the high range could safely be regarded as a top-tier teacher. The critical point is to watch for a pattern to SET results across courses, and be particularly alert for signs of a trend in ratings during the early or formative (and pre-tenure) years. Be prepared to entirely dismiss the odd class rating when it conflicts with a persuasive pattern.

The goal here is to help personnel decision makers use data wisely and avoid drawing improper inferences from statistical results. Decision makers must be educated to the fact that to be rated in the middle range according to this system means that the instructor is performing at a level consistent with other instructors on campus. This is fundamentally different from being assigned an average "middle" rating by student evaluators. As one member of the committee discovered when she contacted the University of Scranton to inquire about their system, a middle-tier rating is the expected (and perfectly acceptable) level of performance under such a system.

V. Summary

The thrust of the committee recommendations is the establishment of a comprehensive system for the evaluation and improvement of teaching and learning. The feasibility of successfully implementing the proposal is judged to be high because the proposed plan is based largely upon resources or essential elements that are in place at WSU. The system would utilize the existing SET program with modest enhancements and would take advantage of the WSU promotion and tenure policy that supports a broader definition of teaching. The plan would help to integrate and focus the mission of the Office of Teaching and Learning. The proposal is supportive of the institution’s teaching mission and is consistent with the long range program goals for that mission.
The plan does not incorporate every refinement the committee found in the literature or in existing ET programs of distinction, rather, the committee made a careful and judicious selection of recommendations designed to build on the existing strengths at WSU that would yield the greatest immediate benefits.

The technical knowledge and know-how to create a nationally recognized program of ET are well known and available to the WSU community. Whether or not WSU achieves such a program, depends in large measure, on the cooperation, the will and goodwill of the faculty, the Administration and the AAUP.
APPENDIX A

REFERENCES CITED AND REVIEWED
REFERENCES


BOOKS AND MONOGRAPHS


UNPUBLISHED REPORTS


Policy on Student Evaluation of Teaching, Fall, 1993. Wayne State University, Detroit, MI.

Report 8.0, Winter 1993 Faculty Survey Results. An Administration survey of Faculty views of the SET program. Wayne State University, Detroit, MI.

Student Classroom Teaching Evaluation Committee Status Report. 1986. Wayne State University, Detroit, MI.

Student Evaluation of Teaching Report 1.0 (June 3, 1991), A Factor Analysis of 1989 SET data, Wayne State University, Detroit, MI.


Wayne State University Promotion and Tenure Procedures and Factors for Faculty. (January, 1994), Wayne State University, Detroit, MI.
APPENDIX B

SURVEY LETTER MAILED TO MEMBER INSTITUTIONS OF THE PROFESSIONAL AND ORGANIZATION DEVELOPMENT NETWORK
July 13, 1994

Dear

The faculty of Wayne State University, like most colleges and universities today, is faced with the necessity of formally evaluating the effectiveness of our faculty teaching. Our current method is limited almost entirely to collecting data reflecting student perception of faculty teaching using an instrument of unknown reliability or validity. Since we are not satisfied with the adequacy of either the process or the instrument, we are in the process of rethinking the entire question. Your institution has been identified to us as one that has given serious consideration to this question. We are particularly interested in any information that you could provide us that addresses the following concerns:

1. Your institutional policy on evaluation of faculty teaching.

2. Copies of any instruments you utilize to gather data on student satisfaction with instruction or effectiveness of teaching.

3. Did your evaluation instruments undergo formal development and testing to assess reliability and validity?

4. Can you identify other colleges or universities that have developed exemplary formal systems and instruments for assessing effectiveness of teaching?

The material should be sent to ASSET Committee, c/o Academic Senate Office, 1279 Faculty/Administration Building, Wayne State University, Detroit, MI 48202.

We appreciate your willingness to share this information with a sister institution. We would appreciate receiving your
July 13, 1994
Page 2

materials by August 30 if at all possible. We will be willing to provide you with a copy of our report once it has been completed.

Sincerely yours,

[Signature]

Richard E. Gallagher, Ph.D.
Professor
Chairperson, Academic Senate Student Evaluation of Teaching Committee
Wayne State University

REG/aw
APPENDIX C

INSTITUTIONS HAVING EXEMPLARY STUDENT ASSESSMENT OF TEACHING PROGRAMS

1. Kansas State University, Manhattan, Kansas. Instructional Development and Effectiveness Assessment System (IDEA).


3. Syracuse University, Syracuse, New York.

APPENDIX D

EXAMPLE OF UNIVERSITY HANDBOOKS PROVIDING FACULTY-ADMINISTRATIVE GUIDANCE TO EVALUATION ON TEACHING PROGRAMS

A Guide To Evaluating Teaching for Promotion and Tenure. Center for Instructional Development, Syracuse University.

Effective Faculty Evaluation: Annual Salary Adjustments, Tenure, and Promotion: A Resource for Faculty and Administrators, Kansas State University.

Guide to the University of Scranton Course Survey. Prepared by the Course Survey Subcommittee of the Academic Policy Committee of the Faculty Senate, Spring, 1993.


APPENDIX E

CURRENT AND PROPOSED ENHANCED, WAYNE STATE UNIVERSITY
STUDENT ASSESSMENT OF TEACHING FORMS

SETCOMB2.DOC
05/04/95
Please complete the following form. Use the first five evaluation items to give your reactions to the course you are now taking. Then use items 6 through 15 to rate the instruction you have had during this past semester. Your responses are very important to your instructor and to the University. This information is used to contribute to:

1. improvements in the quality of instruction at Wayne State,

2. the annual decisions made on salaries, and

3. faculty promotion and tenure decisions.

Please read each item carefully. Fill in the circle with the number that best represents your response to each item. You may add a brief narrative comment to the right of the numbered items; in addition, your written comments regarding course improvement are solicited on the reverse side of the comment page.
Student Profile

Please provide the following information about yourself and your participation in this class.

A. I am currently enrolled for the following degree program:
   1. Bachelor's
   2. Master's
   3. Ed.D or Ph.D
   4. J.D., M.D., or Ph.D.
   5. Other

B. I have completed the following number of college level courses:
   (Fill in only number 6 below if you are a graduate student.)
   1. 0
   2. 1-5
   3. 6-10
   4. 11-15
   5. 16 or more
   6. Not applicable, graduate student

C. The course for me is:
   1. Required
   2. Elective

D. I expect a final grade in this course of:
   1. A
   2. B
   3. C
   4. D or E
   5. Other

E. My grades in college (or this graduate/professional school) have been:
   1. Almost all A's
   2. Mostly A's and B's
   3. Mostly B's and C's
   4. Mostly C's and lower
   5. No previous experience in college (or this graduate/professional school)

F. I attended approximately this percentage of class meetings in this course:
   1. Less than 60%
   2. 60 - 69%
   3. 70 - 79%
   4. 80 - 89%
   5. 90 - 100%

G. I have completed approximately this percentage of the assigned work for this class:
   1. Less than 60%
   2. 60 - 69%
   3. 70 - 79%
   4. 80 - 89%
   5. 90 - 100%

Course # ______________________ Section # __________________________

Instructor’s Name

Course Evaluation Items

1. The course content met or exceeded my expectations for the class.
2. I would recommend this course to a friend with similar academic interests.
3. This course was well organized.
4. I will be able to use what I learned in this course.
5. Overall, I would rate this course as excellent.

Instructor Evaluation Items

6. At the beginning of the course, the overall class plan was clearly presented.
7. The class plan was followed reasonably well.
8. At the beginning of the course, my responsibilities as a student were made clear.
9. All things considered, the instructor was available to me.
10. The instructor treated all students in the class, including me, fairly and with respect.
11. The instructor provided prompt feedback on my performance on assigned activities.
12. The grading procedures were clearly explained at the start of the course.
13. The material presented in class and in assignments was fairly represented on examinations.
14. The instructor's English speaking skills were sufficient to present the course material clearly to me.
15. Overall, the instructor did a good job teaching this course.

If you responded to any of the items above with "Disagree" or "Strongly Disagree," please write the next page to explain your answer. (Remember we are asking you to tell the instructor how to improve the course or his or her instruction)

Instructor Supplied Items

16. __________________________
17. __________________________
18. __________________________
19. __________________________
20. __________________________
21. __________________________
22. __________________________
23. __________________________
24. __________________________
25. __________________________
26. __________________________
27. __________________________
28. __________________________
29. __________________________
30. __________________________
31. __________________________
Your comments regarding course improvements are solicited.
Please use the space below:

Would you like to see something added to this course; if so, please indicate below:

Would you like to see something deleted from this course; if so, please indicate below:

General Comments:
Proposed Student Evaluation of Teaching Instrument

Self Evaluation
1. I attended approximately this percentage of class meetings in this course:
   1. less than 50%  2. between 50% and 75%  3. 75% to 90%  4. More than 90%

2. I expect a final grade in this course of:

3. I really wanted to take this class regardless of who taught it:

4. Before enrolling, I had a strong interest in the subject matter of this course:

Course Evaluation, Part I
1. How much have you learned from this course?
   Nothing------------------------A great deal
   (1) (2) (3) (4) (5) (NA)

2. Would you recommend this course to a friend?
   Never------------------------Absolutely
   (1) (2) (3) (4) (5) (NA)

Instructor Feedback -- Diagnostics
Strongly Disagree(1) Disagree(2) Neutral(3) Agree(4) Strongly Agree(5) Don't know(DK)

1. This course was well organized.

2. I will be able to use what I learned in this course.

3. At the beginning of the course the overall class plan was clearly presented.

4. The class plan was followed reasonably well.

5. At the beginning of the course my responsibilities as a student were made clear.

6. All things considered, the instructor was available to me.

7. The instructor treated all students in the class, including me, fairly and with respect.

8. The instructor provided prompt feedback on my performance and assigned activities.

9. The grading procedures were clearly explained at the start of the course.

10. The material presented in class and in assignments was fairly represented on examinations.

11. The instructor's English speaking skills were sufficient.

Course Evaluation, Part II
1. Rate the course: Poor------------------------Excellent
   (1) (2) (3) (4) (5) (NA)

2. Rate the instructor's teaching in this course: (1) (2) (3) (4) (5) (NA)

Instructor Supplied Items (if present)
1.------------------
2.------------------
3.------------------
4.------------------