

Effective Faculty Evaluation: Annual Salary Adjustments, Tenure and Promotion Contents

Preface

Introduction

Chapter 1: Purpose of Evaluation

Chapter 2: General Considerations

- Areas Evaluated
- Need for Departmental Latitude
- Rationale for Selected Data Sources
- Lists of Data Sources
- Documenting and Packaging Summative Evaluation Data
- Criteria and Standards
- Bases of Evaluation

Chapter 3: Evaluation for Annual Salary Adjustment

- Creating or Revising an Evaluation System
- Implementing the System

Chapter 4: Evaluation for Tenure

- Bases for Evaluation
- Relationship of Annual Evaluation to Evaluation for Tenure
- Relationship of Tenure to Annual Reappointment of Probationary Faculty
- Nid-Probationary Review
- Versatility and Specialization
- Prudence Suggests Caution
- Assessment Should Reflect Responsibilities
- Faculty Participation
- Format of Recommendations
- External Peer Review
- Duration of the Probationary Period
- Assistant Professors Are Not Tenured

Chapter 5: Evaluation for Promotion

- Relationship of Tenure to Promotion
- Kinds of Errors
- External Peer Reviews

Appendices

- Appendix A: Selected Excerpts from *Scholarship Reconsidered*
- Appendix B: Possible Forms for Faculty Input
- Appendix C: Sources for Assistance

Effective Faculty Evaluation: Annual Salary Adjustments, Tenure and Promotion

Preface

In 1989, Provost James Coffman commissioned a task force to develop procedures and guidelines for the **annual evaluation** of unclassified personnel. Based on the resulting recommendations, new procedures for annual evaluation and distribution of merit salary increases were approved by Faculty Senate in December, 1990, and were included in the 1991 *Faculty Handbook*. A second document compiled by that task force, *Handbook for Annual Evaluation of Unclassified Personnel*, was distributed by the Provost to all departments and other academic units in August of 1990 to provide guidance in developing a sound annual evaluation system for each unit.

In the fall of 1990, Provost Coffman commissioned another task force to examine the impact of **tenure** and **promotion** practices upon institutional excellence and to make recommendations for changes that would promote excellence. This task force included four subcommittees: Teaching and Advising; Research and Creative Endeavors; Service; and Extension. The Executive Committee of the task force used the results of a faculty survey, a survey of deans, and the four subcommittee reports to develop a comprehensive set of recommendations for the *Faculty Handbook*. That portion of its final report was submitted to Provost Coffman in January, 1992, and, after significant excision by the Senate Faculty Affairs Committee, was approved without opposition by Faculty Senate in May, 1992.

This second task force also developed resource materials to serve as aids to academic units' work in developing sound criteria, standards, and procedures for promotion and tenure. Rather than producing a separate publication which would duplicate much of the earlier one, the second task force enlarged the earlier report. This publication is therefore an enlargement and revision of the earlier *Handbook for Annual Evaluation of Unclassified Personnel*. It is intended to be an aid to departments and similar academic units in developing evaluation criteria, standards, and guidelines for annual salary adjustments, for tenure, and for promotion.

[Contact us](#) [Emergency](#) [Statements and disclosures](#)

Kansas State University Manhattan, KS 66506 785-532-6011 © Kansas State University Updated: 7/22/12

Effective Faculty Evaluation: Annual Salary Adjustments, Tenure and Promotion

Introduction

Kansas State University is dedicated to contributing to the general welfare of the residents of our state, our nation, and our world through its work in adding to and passing on the fund of knowledge that is the common possession of humankind. This work with information and ideas includes imparting, expanding, storing and retrieving, integrating, interpreting, and applying them. To achieve these ends, the faculty pursues a wide range of activities in teaching, research and other creative endeavors, service and Extension. The University has a responsibility to employ the best faculty it can secure and then to provide a rich environment in which the members of this faculty can develop their full productivity and excellence, the University must also periodically assess the performance of its members and use those assessments as the bases for decisions regarding salary increases, reappointment, promotion and tenure.

The significance of these reviews requires that the criteria and standards upon which they are based, as well as the evaluation procedures used, be founded upon broad agreement among the people affected. The faculty must play a central role in developing the criteria and standards for the University as a whole and in establishing the goals, objectives and expectations of their respective administrative units. Likewise, the individuals charged with implementing these policies must have a say in their formulation. In short, the creation of effective systems of evaluation must be the mutual responsibility of both faculty and administrators.

Effective Faculty Evaluation: Annual Salary Adjustments, Tenure and Promotion

Chapter 1. Purposes of Evaluation

A primary function of formal assessment within the University is to produce what are known as summative evaluations. As the term suggests, summative evaluations come at the conclusion of an activity (e.g. a faculty member's evaluation year), and they are intended to produce judgments on the adequacy or effectiveness of the activity. Summative evaluations thus lend themselves to providing a basis for personnel decisions such as merit salary raises, promotion and tenure. The evaluations help to assure that the personnel decisions are reasonable and defensible and that they foster excellence. Summative evaluation is most effective when it is conducted with the cooperation and participation of those being evaluated.

A second type of assessment is formative evaluation, which is intended to provide feedback for changing the activity being evaluated while it is still in progress. Often less formal in design, this type of evaluation serves the vital purpose of faculty development or professional improvement. This too is critical to the pursuit of institutional excellence, so formative evaluation should be a major concern of unit faculty and heads. Formative evaluation can never be successful without the cooperation and participation of the faculty member being evaluated.

Both kinds of evaluation are vitally important, but the mandate to provide sound and defensible tenure, promotion, and salary recommendations clearly dictates that the systems devised by departments for this purpose shall be summative in focus. Faculty development and improvement are also desirable ends, and the report a faculty member receives for one year's summative evaluation can also be formative with respect to later years or to the pursuit of promotion and tenure. Nevertheless, all parties concerned should understand that the system that provides the basis for personnel decisions must be developed primarily to serve the institutional need to make personnel decisions. Therefore, no unit should sacrifice effectiveness or accuracy in summative evaluations to achieve formative goals.

The desire to develop a system that will simultaneously serve both summative and formative ends is understandable, but results of attempts along these lines are generally less than satisfactory for each purpose. Professional evaluators strongly advise that formative and summative evaluations be conducted separately because their focus, purpose, and timing differ as do the role of the evaluator and the kinds of information needed. For example, an evaluation system designed to provide a basis for personnel decisions cannot offer the unthreatening context in which faculty members can get help toward professional improvement. Similarly, a system intended to create a safe situation in which people can reveal their weaknesses in order to receive assistance is incompatible with the perceived threat of decisions regarding salaries, promotion, and tenure.

Conventional wisdom therefore holds that formative and summative evaluations should be distinctly separated. One way to achieve this is to divide the responsibility for them. The unit head is required to conduct summative assessments, but in some cases an external agency can contribute formative assistance. In particular, the Office of Educational Improvement can provide independent and confidential help to faculty members who wish to strengthen their instructional abilities.

Sometimes confusion arises concerning the data useful for summative and formative evaluation. Some kinds of data serve one purpose better than the other. Thus, observation of teaching is better for formative evaluation, while student ratings on TEVAL are better for summative purposes. However, some kinds of data (e.g., publications, IDEA ratings, assessment of tests, and evaluation of syllabi) can very effectively serve both purposes.

Evaluators must act ethically. In particular, faculty members should show in advance how the data they submit will be used. They must provide any information required by the summative evaluation system, and those data may be considered in the evaluation. However, a person who reveals evidence beyond what is required in the hope of obtaining assistance in improving performance has a need and a right to know whether that information might also be used to make or to support personnel decisions.

Effective Faculty Evaluation: Annual Salary Adjustments, Tenure and Promotion

Chapter 2. General Considerations

Each primary administrative unit (hereafter referred to as department) that includes unclassified employees must develop a system of evaluation for annual salary adjustment. In addition, those units in which promotion and tenure are possible must develop a system of evaluation to drive promotion and tenure decisions. These systems must be able to operate within the procedural context and the constraints imposed by University policies. The University's evaluation procedures establish certain general guidelines for the development of systems, especially with regard to participation by various parties. The same considerations apply whether the system is to be developed initially or simply revised. In the first place, each department's system should be the product of a cooperative effort on the part of the faculty and the responsible administrators. In most cases, these include the faculty, department head, dean, and provost. Effective systems require broadly based support among the people whom they affect; therefore, open and active participation by the faculty in the creation of the system is essential to its success.

Each system must reflect the responsibilities and goals of the department for which it is designed. These responsibilities are partially determined by external circumstances. Some departments have doctoral or master's programs, while others offer only undergraduate instruction. Some units have major commitments in Extension, while others have none. Some unclassified employees spend all of their time providing student services, while others devote all their efforts to research. Thus, each unit's evaluation system must reflect the kinds of tasks required of the people being evaluated. Those departments having considerable diversity among faculty assignments must provide for this specialization of labor in their evaluation systems. If a unit's responsibilities change over time, the evaluation system should be revised to reflect the changes.

At the same time, departments have a good deal of independence in establishing their goals. These objectives might include--among many possibilities--augmenting enrollment, achieving instructional excellence, developing new departmental foci, promoting the supervision of graduate students, increasing grant income, building a national reputation through scholarship, or providing new services to an off-campus clientele.

Whatever the unit's goals may be, an effective system of evaluation must encourage and reward activities that contribute to their achievement.

This need for departmental systems to foster the pursuit of institutional excellence requires that systems reflect at least (a) quality and quantity of the faculty members' work and (b) the extent to which faculty members' work match the needs of the department in the pursuit of its missions.

1. Areas Evaluated

One of the first tasks in creating a system of evaluation is to determine the general domains of faculty effort to be covered and the activities appropriate to each. The areas to be considered in faculty evaluation are largely determined by the responsibilities of the unit creating it.

Faculty activities can be categorized in a number of ways. One method worth considering was proposed by Boyer (1990) in *Scholarship Reconsidered: Principles of the Professoriate*. A brief summary of the Boyer categories is included in Appendix A and might be useful reading in preparation for creating a system of evaluation.

A more traditional approach is currently used in the K-State Faculty Handbook to identify and classify kinds of faculty work performed in a broad range of units across the University. The domains are: teaching and advising, research and other creative endeavors, directed service, non-directed service, and Extension. Each of these five domains includes numerous possible subordinate components.

a. Teaching

Teaching includes communicating knowledge to students and developing the intellectual foundation necessary to prepare students to continue learning for themselves. Teaching also involves preparing students for entry into the professional and scholarly disciplines. Effective teaching is based upon

Excellence in teaching is a primary criterion by which some important constituents (e.g., students and parents) judge the stature of a University.M

b. Research and Other Creative Endeavors

Research and other creative endeavors encompass a broad spectrum of scholarship and other activities that require critical analysis, investigation, or experimentation. These endeavors are directed toward discovery, interpretation, or application of knowledge and ideas. Creative activities also include innovative works in the fine arts, performing arts and design professions. The results of research, scholarship and other creative activity should be shared with others through publication, performance, or other media appropriate to the discipline.

Excellence in research and other creative endeavors is a primary criterion by which some important constituents (e.g., the national and international scholarly community) judge the stature of a University.

c. Directed Service

This type of service, whether performed for the department, the University, or the public at large, is explicitly delineated in a faculty member's position description, requires academic credentials and/or skills, and is often routinely and explicitly scheduled in terms of time and place. This service furthers the mission and is central to the goals and objectives of the unit. Typical examples are the responsibilities assigned to librarians, clinician/diagnosticians, and academic program directors in departments. The nature and time commitment for directed service becomes part of a faculty member's annual plan of work and performance appraisal.

d. Non-Directed Service

Non-directed service is often referenced by the generic term "service." It is non-directed in the sense that specific expectations are not usually delineated in job descriptions and much latitude exists for faculty members to choose how they will fill some obligation for non-directed service. It contributes substantially less to personnel decisions than do the major dimensions of teaching and research and other creative endeavors (among most faculty members) and directed service and Extension (among those who have such assignments). Non-directed service includes three subcategories:

i. Non-Directed Service to the Institution

Institutional service includes contributions to the department, to the college, and to the University as a whole. Many of these activities are related in some way to University governance, and they derive from the tradition that the faculty should establish and enforce standards both for itself and for its students. University service, therefore, embraces the broad range of activities involved in establishing and implementing policies at every level of the institution.

ii. Non-Directed Service to the Profession

Professional service encompasses contributions to the academic profession beyond the campus. These include holding office in professional societies or membership on their committees, performing editorial functions for professional publications, or organizing professional meetings.

iii. Non-Directed Service to the Public

Public service includes the application of knowledge gained through scholarship for the benefit of a non-academic audience, provided that it is not directed service or Extension (described below). Public service may be rendered to individuals, communities, organizations, and public agencies. It encompasses both the sharing of information and knowledge and the application of knowledge to solving problems. Faculty members involved in outreach activities have direct and often sustained contact with the general public, officials, and leaders. They perform assessments, develop programs, and provide training, consultation, and technical assistance. Effectiveness in public service requires expertise in appropriate subject matter, depth and breadth of knowledge, organizational capability, and excellent written and oral communication skills. Departmental criteria, standards, and guidelines should establish the parameters of non-directed service to the public in ways that distinguishes it, insofar as possible, from civic and personal service.¹

¹ Civic and personal service is not applicable to evaluation. It is generally viewed as a person's participation as a citizen and may be directed toward government, religious, fraternal, interest group, or philanthropic endeavors. It flows from personal skills and individual choice in use of private time.

Fairly clear examples include participation in workshops that are not job-related, holding positions in community or religious organizations, election or appointment to government offices. For a professor of English, clothing, or engineering, it is also quite clear that holding office in a local PTA, participating in activities of the Audubon Society, performing musical solos in church, or taking part in local dramatic productions should not count for summative evaluation.

Less obvious cases would involve a professor of education holding office in a local PTA, a wildlife

biologist participating in activities of the Audubon Society, a professor of music performing solos in church, and a professor of drama working in community theatre. Departments may differ in whether they consider such activities to be personal service which does not count for evaluation) or to be non-directed service to the public (which does count). Moreover in some cases reasonable professionals within a department will hold differing opinions. It is to be expected, however, that most such issues will be resolved in the departmental statement of criteria, standards, and guidelines.

e. **Extension**

Extension, an area of endeavor especially characteristic of Kansas State University's land-grant mission, integrates elements of teaching and public service to provide practical, research-based assistance to clients such as individuals, families, farms, businesses, and communities. Extension programs are based on scientific knowledge, applied principles, and recommended practices. Extension is pledged to meeting the state's needs for knowledge and research-based educational programs that will capitalize on clients' "teachable moments" in order to enable them to make practical decisions.

Excellence in Extension and certain directed services is a primary criterion by which some important constituents (e.g., agricultural producers and business people) judge the stature of a land-grant University.

f. **Academic Citizenship**

The University needs collegiality to function effectively, and units may wish to consider it in evaluation, either as a part of the more traditional areas or as a separate domain of achievement. Some faculty members foster goodwill and harmony within a department, mentor colleagues, and generally contribute to the pursuit of common goals. Other individuals may display behavior that is highly disruptive to the department; as a result, collegiality and morale suffer. A system might include a statement that behavior affecting, whether positively or negatively, the ability of others to carry out their assignments in the department may be considered in the total evaluation. Such behavior should be documented in the narrative portion of the evaluation.

2. **Need for Departmental Latitude**

Units should have a good deal of latitude in deciding which activities are appropriate to each of the general categories of evaluation. The purpose of such determinations should not be simply to impose conformity within a department but to facilitate the process of comparing the individual accomplishments of different members of the unit.

For example, functioning as the Graduate School's representative on a doctoral committee could be viewed either as an instructional activity or as an institutional non-directed service activity. Similarly, some departments might regard the scholarship involved in publishing a book review to be most related to research or other creative endeavor, while others would treat it as professional service. And some units may treat the supervision of dissertation research as teaching, while others will elect to regard it as research.

Academic advising is another activity that departments may classify differently. Some, especially those whose academic advisement is relatively standardized and routine, may wish to treat advisement as institutional service. Other departments, especially those whose programs require more individualization, may consider academic advising to be more akin to teaching.

3. **Rationale for Selection of Data Sources**

It is a fundamental principle of evaluation that anything as complex as professional performance cannot adequately be captured by a single source of information. **Faculty evaluation should be based on multiple sources of data. And these data sources should be chosen to yield disparate kinds of information.**

One purpose of using multiple and disparate data sources is to provide different perspectives on the performance of interest. Just as one picture of a house cannot show it from every angle, one source of information cannot reveal all that is important about any major domain of faculty activity. Moreover, isolated bits of data, even if true, can be unintentionally--or, indeed, intentionally--misleading. Single data sources are best conceptualized as yielding pieces of circumstantial evidence, no one of which is persuasive by itself. However, when multiple data sources are consistent in what they indicate, they can be taken more seriously. This is especially true if the sources of evidence are dissimilar in nature.

Excessive reliance on a single source of information can have quite another effect as well. When one source is given more evaluative emphasis than it merits, it comes also to receive more attention than it deserves from the people being evaluated. This creates the risk that they will narrow their objectives to achieving a good showing on that indicator rather than striving for general excellence in performance. **Evaluation has impact. Wise evaluators anticipate the consequences of evaluation.**

Those who develop evaluation systems and those who execute the task must consider the impact of evaluation in every area reviewed, but examples of three potential sources of data from the domain of

teaching will suffice to illustrate the point. Each will be considered with regard to its validity and with regard to its impact on instruction. The discussions focus on the positive and negative consequences of using various sources of information. Similar analysis could be made of any data source used to evaluate any aspect of professional accomplishment, and departments are urged to do so in developing their lists of criteria and standards.

a. **Example 1: Teaching Materials**

For reasons unrelated to evaluation, instructors prepare and distribute important written material, including syllabi, reference lists, assignments with directions, and tests. The use of such materials for faculty evaluation provides a rich source of information at very little extra trouble or expense. To be sure, there is much more to teaching than producing instructional materials, but those materials cast light on important aspects of the communication between teacher and students.

The time, trouble, and expense of reliably rating instructional material is also very minor compared with that needed for reliably assessing the adequacy of teaching by means of direct observation by evaluators. As long as the importance of such materials is kept in perspective, their inclusion as one of several data sources in an evaluation system adds a perspective that is otherwise likely to be neglected. Furthermore, the review of these materials by an evaluator would likely prompt some faculty to do a better job in preparing them, a result that would be generally desirable. Moreover, constructive feedback would address such aspects as the clarity of directions on assignments, the extent to which tests go beyond mere factual recall, the extent to which sexist language and gender-role stereotyping is avoided, and the accuracy and currency of the information. Surely such feedback would enable most of us to improve our teaching. Therefore, the inclusion of informed evaluation of teaching materials as one of several data sources for purposes of accountability and/or improvement would likely have a positive effect upon teaching.

b. **Example 2: Student Achievement Test Scores**

From time to time the idea arises that measures of student achievement offer a sensible way by which to evaluate instructional effectiveness. On the surface, this seems to be a reasonable, direct, reliable, and valid data source. But without assessments of students' prior knowledge and skills, the use of student achievement measures for faculty evaluation may work a grave injustice on faculty who are assigned to teach less capable students.

Use of this data source may also have an undesirable effect on teaching. If faculty members are being judged on the basis of student test scores, then the faculty goal may adaptively become high student test scores (in contrast to high student learning). This may cause a serious curricular overemphasis on those outcomes that are most easily tested. Too, it can cause teaching to the test (i.e., focusing on material because it happens to be sampled in a test rather than because of its importance to command of the subject.). Thus, this approach to evaluating instructional effectiveness could actually subvert the enterprise of providing well rounded education.

This serves as a sobering example of the need to consider carefully the likely consequences of evaluative procedures before they are implemented. If the consequences are highly undesirable, it may be better to neglect otherwise useful evaluative data sources than to use them at the risk of undermining teaching effectiveness.

c. **Example 3: Student Ratings**

University students are able to provide useful information about teaching effectiveness. For example, who could better rate the clarity of a professor's speech, how meaningful a teacher's explanations are to students, whether a professor has distracting or annoying mannerisms, or how well an instructor understands student questions? By the end of a course, students possess substantial amounts of information about the teaching that can be tapped at very little extra trouble or cost.

There are, of course, other kinds of information that students are not qualified to provide. For example, they are not likely to be adequate judges of how well an instructor knows the subject matter or keeps up with the field, the appropriateness of content balance, or a course's adequacy of preparation for subsequent courses.

Assuming, then, that student feedback is employed to secure only those kinds of information that students are well qualified to provide, questionnaires provide a useful source of information concerning teaching effectiveness. Used either for instructional improvement or for making personnel decisions (if appropriate adjustments are made for differences in student motivation), student feedback from expertly developed instruments offers reliable, valid, and credible information that would likely prompt desirable consequences in teaching behaviors.

4. **Lists of Data Sources**

The lists that follow are organized around broad areas of endeavor appropriate to faculty evaluation. They are intended to give an idea of the variety of activities encompassed by each and the many forms of documentation available to evaluators. Earlier versions of these lists were adapted from *The University of Georgia Guidelines for Appointment, Promotion, and Tenure* (2nd ed., 1981), pp. 6-16. **The following lists are not intended to be exhaustive.** Departments will find it helpful to use what is applicable from the lists and to augment them where desirable. **The listing of an activity under one area here should**

not preclude a department from deciding to consider it under another category.

a. Mission Relevance

Before listing of some of the ways by which performance in the various domains can be assessed, mission relevance of work--an issue that, although rather obvious, is of paramount importance--must be addressed. The lists that follow illustrate data sources by which the quality of professional work can be evaluated. The kinds of information illustrated in the lists are useful in evaluating the quantity and quality of faculty members' work. If, as is usually the case, a person's work is relevant to departmental mission(s), then it can be considered in faculty evaluation. On the other hand, if portions of a faculty member's productivity were in areas having nothing to do with departmental mission(s), then the irrelevant work should not be considered. Between the extremes of centrally relevant work and wholly irrelevant work are various shades of relevance to missions that may have differing priorities. **Departmental policies on faculty evaluation must provide for consideration of the mission relevance of faculty members' work.**

b. Teaching

As a department selects indicators of instructional effectiveness, several rules of thumb should be observed. First, an adequate system of summative evaluation of teaching will include at least three kinds of information: (a) classroom effectiveness, (b) preparation of instructional materials and syllabi, and (c) student assessment practices. Second, no one such kind of information should contribute half or more of the weight for the assessment of teaching effectiveness. Third, data should be gathered from at least two relatively independent sources including (a) systematically collected, anonymous student ratings on professionally designed instruments for every regular class and (b) ratings of peers or supervisors. The student ratings, of course, are best used to assess classroom performance while peer and supervisor ratings are superior for assessing instructional materials and student evaluation practices.

Some departments underemphasize teaching in personnel evaluations in the belief that it cannot be evaluated as precisely as research. Persons holding this view probably (a) are unfamiliar with some of the many valid ways by which instructional effectiveness can be documented and/or (b) do not fully understand the difficulties involved in obtaining valid evaluation of research effectiveness. Both can be assessed, but doing so requires knowledge, thought, and effort in the balanced use of multiple indicators. While there are many useful data sources, many authorities on the evaluation of teaching consider student ratings on well designed instruments to provide the single best source of data concerning teaching effectiveness. However, the development of adequate instruments is a specialized professional endeavor. A committee of people lacking specialized expertise in this field is not likely to be any more effective in designing an adequate student rating instrument than it would be in designing a new building.

Student rating instruments used for summative evaluation should have a number of technical features. They should (a) be norm referenced, (b) provide statistical adjustments for demonstrated sources of bias (especially initial student motivation, but also preferably for class size), (c) be administered under clearly specified, standardized conditions (including provision for instructors to be out of the room while the rating instruments are completed), (d) provide anonymity to student raters, and (e) focus on aspects of instruction that students are qualified to assess.

In order to achieve comparability among persons within departments, each department should mandate the use of a particular instrument for summative evaluation and rigorously adhere to standard procedures for administering it. For example, a trained person may be provided to administer the mandated form under standard conditions.

Faculty members should, of course, be free to supplement the mandated summative instrument(s) and entirely free to choose those instruments, if any, they use for formative purposes.

The indicators listed below encompass a wide spectrum of teaching activities assessed by students, peers, supervisors, and other appropriate judges. These are some of the indicators of teaching effectiveness that departments may consider.

- i. Student ratings from norm-referenced instruments that assess teaching effectiveness rather than popularity and that adjust for such known sources of bias as student motivation and class size.
- ii. Materials produced for individual courses such as reading lists, syllabi, and other instructional materials.
- iii. Tests and other materials and methods used to assess student achievement.
- iv. Depth, breadth, and currency of subject matter mastery.
- v. Appropriateness of course content.
- vi. Effective course administration, e.g., maintaining office hours and punctuality in performing teaching-related paper work, such as turning in textbook orders, reporting grades, and filing syllabi.
- vii. Development of effective courses, preparation of innovative teaching materials or instructional techniques, or creative contributions to a department's instructional program.
- viii. Assessment by faculty colleagues who are familiar with the teacher's performance or have taught that person's students in subsequent courses.

- ix. Successful direction of individual student work of high quality, e.g., independent studies, theses or dissertations, and special student projects.
- x. Effective and diligent advisement of students in pursuing their academic programs.
- xi. Successful performance of teaching responsibilities that are unusually demanding or require special expertise or preparation.
- xii. Versatility in contributing to the department's teaching mission, e.g., effective performance at all levels of instruction appropriate to the department, including membership on the Graduate Faculty and certification to direct dissertations.
- xiii. Special contributions to effective teaching for diverse student populations.
- xiv. Compiled student comments (such as those obtained from program assessments or exit interviews) that address a teacher's abilities to arouse student interest and to stimulate work and achievement by students.
- xv. Letters of evaluation from former students.
- xvi. Accomplishments of the teacher's present and former students; i.e., information showing the students' success in learning the subject matter of the discipline and in pursuing it to a point of intellectual significance.
- xvii. Students coming from other schools especially to study with the teacher.
- xviii. Professional publications on the topic of teaching or materials prepared for use in teaching such as textbooks, published lectures, and audio-visual or computerized instructional materials.
- xix. Presentation of papers on teaching before learned societies.
- xx. Adoptions of a faculty member's textbooks or other instructional materials, especially repeated adoptions, by reputable institutions.
- xxi. Honors or special recognition for teaching accomplishments.
- xxii. Selection for special teaching activities out-side of the University, especially in international assignments, e.g., Fulbright awards, special lectureships, panel presentations, seminar participation, and international study and development projects .
- xxiii. Membership on special bodies concerned with teaching, e.g., accreditation teams and special commissions.
- xxiv. Receipt of competitive grants or contracts to fund innovative teaching activities or investigations into effective teaching, especially for a diverse student population.
- xxv. Membership on panels to judge proposals for teaching grants or contracts.
- xxvi. Selection for teaching in special honors courses and programs.
- xxvii. Special invitations to testify before governmental groups concerned with educational programs.
- xxviii. Evidence of excellence in supervision of students being trained in clinical activities and practica: this includes, but is not limited to, work on campus in the Veterinary Hospital and Veterinary Diagnostic Laboratory, the Speech and Hearing Center, the Family Center, and similar training/service units; and off campus in student teaching and other approved educational programs such as practica, internships, and preceptorships.

c. Research and Other Creative Activities

This evidence encompasses evaluations of performance by peers and other important judges. In joint endeavors, the degree of each person's contribution should be identified.

- i. Books, reviews, monographs, bulletins, articles, and other scholarly works published by reputable journals, scholarly presses, and publishing houses that accept works only after rigorous review and approval by professional peers.
- ii. Exhibitions of art or design works at important galleries, selection for these exhibitions being based on rigorous review and approval by juries of recognized artists or critics.
- iii. Performances in prestigious recitals or productions, selection for these performances being based on stringent auditions and approval by appropriate judges.
- iv. Presentation of research papers before one's peers at scholarly meetings and learned societies.
- v. Scholarly reviews of the faculty member's publications or critical reviews of art works and performances.
- vi. Citations of research in scholarly publication.
- vii. Reprinting or quoting of publications, reproductions of art or design works, and descriptions of interpretations in the performing arts appearing in reputable works in the discipline.
- viii. Accomplishments of the faculty member's present and former graduate students.
- ix. Competitive grants and contracts to finance the development of ideas or performance, these grants and contracts being subject to rigorous peer review and approval.
- x. Prizes and awards for excellence of work done.
- xi. Development of, and where appropriate obtaining patents or copyrights for, processes or instruments useful in solving important problems.
- xii. Membership on important scholarly expeditions or explorations.
- xiii. Awards of special fellowships for research or artistic activities or selection for assignment at special institutes for advanced study.
- xiv. Invitations to testify before governmental groups concerned with research or other creative activities.

d. Directed Service

Some faculty members have directed service responsibilities, which may constitute a great part of their work assignment. The following data sources may be relevant to various kinds of directed-service assignments.

- i. Ratings by clients of the quality of service.
- ii. Peer or supervisor assessment of instructional or service materials developed.
- iii. Assessment by practicing professionals who come into contact with the faculty member.
- iv. Ratings by students of the supervisor's delivery of clinical services.
- v. Ratings by peers or supervisors who observe and are qualified to rate the delivery of professional services.
- vi. Evaluation by peers who receive the professional services.

e. **Non-Directed Service**

Non-directed service, or simply "service" as it is understood by most faculty members, is usually broken into three components:

i. **Non-Directed Service to the Institution**

This evidence encompasses evaluations of the performance of such activities by administrators, committee heads, and co-workers in the groups.

- i. Chairing of, membership on, and contribution to standing or ad hoc committees of the University or any of its subordinate units.
- ii. Chairing of, membership on, and contribution to bodies participating in faculty governance, such as the Faculty Senate and its committees, the Graduate Council, and the several College Committees on Planning.
- iii. Performance of unbudgeted administrative responsibilities at the departmental level.
- iv. Special assignments such as representing the University at national and international meetings.
- v. Honors or special recognition for contributions to the department, college or University or to faculty governance.

ii. **Non-Directed Service to the Profession**

This evidence encompasses evaluations of the performance of such activities by other members of, and leaders in, the organizations to which the service is rendered.

- i. Holding office in professional associations and learned societies.
- ii. Service on state, national, and international committees in professional organizations.
- iii. General presentations or addresses at conventions and other professional meetings.
- iv. Organizing or chairing sessions at professional meetings or organizing the meeting itself.
- v. Reviewing or editing for professional journals, e.g., writing book reviews for publication and service as editor, associate editor, book review editor, or member of an editorial board.
- vi. Membership on panels judging grant/contract proposals, juries judging art works, or juries auditioning performing artists.
- vii. Service as a consultant on problems appropriate to the discipline.
- viii. Honors or special recognition for contributions to an organization, discipline, or profession.

iii. **Non-Directed Service to the Public**

Non-directed public service involves the application of a faculty member's professional time and expertise for the benefit of non-academic audiences. This category does not include all activities a faculty member might perform for the public good, but only those that are job related. (See page 4 for a discussion of the distinction between personal service and public service.) This evidence encompasses evaluations of the performance of activities by members and leaders of the groups served.

- i. Written dissemination of professional knowledge or information to non-academic audiences through general interest publications.
- ii. Oral dissemination of professional knowledge or information to civic, religious, or private groups.
- iii. Providing expert testimony to courts or legislative bodies.
- iv. Consulting for state, national, and international public and private groups engaged in educational, scholarly, and artistic endeavors.
- v. Consulting for individuals or corporations engaged in business or industry.
- vi. Providing technical consultation to professional or non-academic groups.
- vii. Engaging in the delivery of technology through involvement in development projects-- especially in international assignments.

f. **Extension**

This evidence encompasses evaluations of the performance of activities by participants, peers, supervisors, and other important judges. In joint endeavors, the degree of each person's contribution should be identified.

- i. Extension program development, implementation and evaluation.
- ii. Extension instruction.
- iii. Instructional materials developed in Extension, including the incorporation of new knowledge and educational techniques into Extension materials and delivery methods.
- iv. Extension consultation and technical assistance.
- v. Development and maintenance of contact with clientele groups, advisory committees, and industry and with research, teaching, and Extension personnel in the area of program responsibility.
- vi. Dissemination of applied research through Extension.
- vii. Publication of research results in Extension publications and use of other methods of communicating information including both new materials and revisions of existing material.
- viii. Development and application of effective ways to identify problems and assess needs.
- ix. Adoption and use of the Extension specialist's program and activities in other state, national, and international programs.
- x. Reviews in appropriate media of the Extension specialist's work and innovations.
- xi. Development of, and where appropriate obtaining patents or copyrights for, instruments, processes, and programs useful in solving persistent problems encountered in Extension.
- xii. Honors or special recognition for contributions to Extension, e.g., Distinguished Service Award.
- xiii. Interdisciplinary program development.
- xiv. Membership on special task forces concerned with Extension programs and issues, e.g., youth at risk, water quality, food safety and quality, or waste management.
- xv. Receipt of competitive grants and contracts to finance development and delivery of innovative programs.
- xvi. Selection for membership on panels judging award, grant, or contract proposals for Extension programs.
- xvii. Invitation to testify before governmental groups about Extension programs.

5. **Documenting and Packaging Summative Evaluation Data**

The question of documentation of activities is closely related to the issue of "packaging." Some people with truly significant accomplishments may injure themselves by presenting their materials for evaluation ineptly or too modestly, while others with less impressive records may improve their images through effective packaging or exaggeration. In addition, some people have good memories or keep good records of their accomplishments and are thus well prepared to present them, while others are less systematic. Adoption of an evaluation form to be filled out by the persons evaluated or simply an outline indicating the activities to be reported and the format to be followed can eliminate some variations in presentation and thus reduce the influence of packaging. These devices can also serve to remind people of things they have done that deserve consideration in the evaluation, helping to make sure that they receive credit for what they have accomplished.

An excellent technique is for the faculty member to provide a succinct summary of evidence followed by well organized, detailed documentation. This enables those who wish to examine the accomplishments with reasonable speed to do so without danger of overlooking significant achievement. At the same time it accommodates the legitimate needs of an evaluator who might desire either documentation or detail. Those responsible for personnel decisions ought never to be expected to accept on faith or at face value all claims of accomplishments; documentation is necessary.

For example, half a single sheet of paper might be adequate to summarize a year's student ratings of instructional effectiveness. Such a summary is very helpful. Yet those responsible for personnel decisions should verify the accuracy of the summaries. Moreover, they may want to scrutinize the computer summaries. Thus the full reports should accompany the summary.

6. **Criteria and Standards**

Those charged with developing an evaluation system must attend not only to the information the evaluator needs but also to the standards applied in assessing the data. The problem is to achieve as much clarity and objectivity as possible without oversimplifying the task. For instance, it is appropriate to have a general understanding that senior or sole authorship in national refereed journals generally merits more credit than junior authorship, but it would not be appropriate to specify just how much more or to dictate that all articles of a given category merit equal credit.

Similarly, it is appropriate to specify that supervision of a master's thesis typically merits less credit than

guidance of a doctoral dissertation, but room must be left for thoughtful consideration of the differences within each category and for the overlap between them. Likewise, it is useful to specify that some kinds of consulting activities tend to merit more credit than others, yet it is not feasible to prescribe in detail exactly how much credit might be awarded for every possible consultation.

It is also entirely appropriate to distinguish various levels of responsibility among the faculty persons in the same unit. This might be done for evaluation for annual salary adjustment and it surely would be done for purposes of promotion to different ranks. In the typical academic department, for example, it is reasonable to establish different expectations for people in different ranks. One way to accomplish this is to express qualitative or quantitative differences in the activities expected. For example in instruction, a higher-level expectation might be for teaching courses at all levels and supervising dissertations while a lower-level expectation might require only a narrower range of teaching. In research, the higher level might require publications in more prestigious media than the lower level. Similarly, higher-level expectations might specify the winning of extramural funding, while lower-level demands could be met by submitting grant proposals. In service the difference might be expressed in chairing as opposed simply to serving on committees and in more significant contributions to professional service from established individuals than from newcomers.

Along somewhat different lines, it is necessary for evaluation systems to distinguish between criteria that relate to the **quality of a faculty member's work** and the vital criterion of the **relevance of this work to the departmental mission(s)**. The quality of one's work is, of course, an attribute of the individual, whereas the mission-relevance contribution of one's work is an interaction between the quality of work and its importance to the department, college, and University missions.

a. **Evaluation Requires Judgment**

It is important to recognize that effective completion of tasks assigned in the broad areas of effort is not by itself a sufficient measure of an individual's contribution to the unit. Some institutions have developed evaluation systems that provide detailed descriptions of responsibilities, criteria, and standards and lengthy lists of professional accomplishments in which particular activities are assigned predetermined numerical values. Such reductionistic systems which attempt to transform evaluation into mere point counting inadequately assess the individual's overall impact. Professional performance is simply too complex to lend itself to full pre-specification. Its adequate evaluation demands professional judgment.

b. **Relative Importance of Activities**

One of the most difficult tasks of the evaluator is judging the relative importance of the activities submitted for evaluation. Relative importance depends upon a number of factors. Ordinarily, the more time an activity requires, the greater its relative importance. For example, a standing committee meeting once a week would probably weigh more heavily than a short-term, ad hoc committee, and a four-hour course more heavily than a two-hour one. Nevertheless, the time or effort expended on a particular activity is not by itself an indication of significance, and an effective evaluation system or evaluator will not encourage activities that are merely time consuming. Relative importance can also be inferred from the actual or anticipated consequences of an activity. Those having or expected to have more profound effects--on individuals, knowledge, resources, etc.--should be judged more important than those whose effects are expected to be more superficial. Those enhancing the University's or department's reputation or image with important external constituents are generally more important than those that are of only internal interest.

c. **Group Activities**

Besides determining the significance of an activity, the evaluator must apportion the responsibility for tasks to which more than one person has been assigned. This is especially true in committee work, but is also apparent in "team teaching" and "team research." How should each member of the team be "credited" in terms of the team's performance? There is no obvious answer to this question. Clearly, some evidence of the team's productivity should be collected--a committee report and its impact on policy or operation, evaluation of team taught courses, publications or grants resulting from team research. The team leader (chair, course leader, principal investigator, etc.) should ordinarily be given more credit than other members of the team. In major team efforts, the leader may be asked to rate the contribution of each member of the team. In team teaching situations, the class may be asked to rate each of the instructors. In brief, some effort must be made to acknowledge the individual's involvement in the team effort.

d. **Remunerative Activities**

Questions sometimes arise in the evaluation of professional activities that provide the individual with private income or with compensation in addition to the annual salary. These include such undertakings as outside consulting, contract research, and the production of salable works such as textbooks. So long as such work is performed in compliance with University regulations, the question of whether it should be regarded positively or negatively in the evaluation depends on the quality of the work. Activities that contribute to the professional reputation of the individual and to the mission, goals, or reputation of the University are generally positive unless they interfere with the performance of regularly assigned duties. Also remunerative activities that are innovative or path-breaking could reasonably contribute to a positive evaluation, whereas routine work undertaken

for pay need not be judged particularly meritorious, even when it involves the use of professional expertise.

7. Bases of Evaluation

There are two broad factors that departments should take pains to accommodate in developing their standards for evaluation--performance of the individual and mission relevance of the person's work. In particular cases, these bases, if applied independently would not necessarily lead to the same decisions. Furthermore, the relevance of each of these bases of evaluation differs among kinds of personnel decisions (e.g., initial employment, annual salary adjustment, tenure and promotion).

Effective Faculty Evaluation: Annual Salary Adjustments, Tenure and Promotion

Chapter 3. Evaluation for Annual Salary Adjustment

General guidelines for the evaluation system at Kansas State University were laid down in 1974, when the Faculty Senate approved a policy statement regarding the annual evaluation of unclassified personnel for merit salary increases. The policy mandated that each department should create a system including the three specific features listed below. Policies and procedures have been elaborated over time, but these three points remain fundamental:

1. Criteria and procedures are to be developed jointly by faculty, department heads, directors, and/or deans.
2. Unclassified personnel will provide an update of relevant data on a yearly basis pertaining to whatever merit salary criteria are established within their unit.
3. Unclassified personnel will be provided the opportunity to review the final written evaluation being used as the department head's recommendation for merit salary increases before it is submitted to the dean.

a. Creating Or Revising an Evaluation System

i. Bases of Evaluation

Evaluation for annual salary adjustment is based jointly on the excellence and productivity of work and upon the extent to which work contributes to unit and University missions. Therefore within departments having strong priorities among programs, annual salary adjustments may and should reflect institutional needs as well as the excellence and amount of work.

Evaluation for annual salary adjustments may also be based secondarily (and separately) on considerations of supply and demand. Thus in a department having significantly different supply, demand, and cost factors operating among the specialities within its discipline(s), annual salary adjustments may reflect these differences as well as individual excellence of work and its relevance to the missions of the institution. This does not mean, however, that a person who could command a larger salary elsewhere should necessarily be favored over other members of the department in the annual salary recommendation. Supply and demand factors properly enter into consideration for salary adjustments only if the person (a) could command a significantly larger salary elsewhere, (b) could not likely be replaced at the level of the present salary, (c) is distinctive within the department in these respects and (d) demonstrates individual excellence in endeavors that contribute to the unit's missions.

Fortunately, the similarities within many departments' programs render relatively minor both differences in mission relevance to the institution and supply and demand factors. In the common case where these factors are essentially equal, annual salary adjustments can be based solely on individual accomplishments. This is both simpler and better for faculty morale.

ii. Context and Constraints

The context of annual evaluation is established by the University's guidelines for conducting this process as contained in the Faculty Handbook and by the institution's goals and standards. These provide guidance on procedures to be followed in evaluations and broad definitions of the goals toward which every unit should direct its efforts.

The constraints are more varied. Some have their origin in policies established by the Legislature, the Board of Regents, or the University, while others are simply "realities" that may be founded solely on circumstance. Nevertheless, whatever their origin, these constraints impose certain parameters within which any evaluation system must be developed and implemented.

An example of a policy constraint is the legislative requirement that all salary adjustments be

Another element influencing the functioning of any system is the University's practice of assigning salary adjustments as a percentage of an employee's base salary rather than as a dollar increment. One result of this practice is that two people having the same rating in an evaluation will receive the same percentage raise, but the one with the higher salary will receive more dollars.

An example of the "non-policy" realities that affect the functioning of evaluation systems is the often sizable differences in the amount of money provided by the Legislature for salary adjustment from one year to the next. Given that the State's budget represents a balance of competing demands for resources, and that the resources themselves differ significantly from year to year, these variations are unavoidable. Yet they create problems for evaluators charged with the equitable distribution of rewards within their units.

Another constraint is the "fixed-sum" system by which salary raises are allocated. Evaluations are relative because they are based on a comparison of the performance of the unclassified persons within an administrative unit. The results of the evaluation do not indicate an individual's absolute status or even status with respect to a large external reference group, but only his or her standing within the department. Therefore half of all the members of the unit must be below the median. This situation is difficult for many faculty members to accept because they are accustomed to excelling in most endeavors. Indeed, if they had not been among the top performers at every level in their education, they would not have attained the positions they hold today. This makes it extremely difficult for them to contemplate being "average" (let alone "below average") regardless of the peer group with which they are compared.

These problems associated with annual evaluation are exacerbated by the assignment of raises. The unit head is allotted a fixed sum of money for salary increases, usually expressed as a percentage of the unit's present salary pool. That money must be distributed to the members of the department in accordance with their individual evaluations. As a result, about half of the members of the unit will receive salary increases below the average percentage increase approved for the University by the Legislature, and this is true even for people whose evaluations indicate that they are meeting or even exceeding the unit's expectations.

iii. **Weighting the Elements of Evaluation**

The department must also consider the contribution of each of the areas of evaluation toward the determination of the final overall assessment. As in defining the areas themselves, the weight assigned to each area will depend upon the responsibilities and goals of the unit. Thus, for example, a department with a doctoral program might reasonably demand greater scholarly accomplishments from its faculty than a department that offers only baccalaureate degrees, and this difference could be expressed by the weight assigned to research in the overall evaluation. Similarly, a unit might examine the various activities that contribute to each broad area of evaluation to determine whether further specification of weighting is appropriate. Thus, a department with a graduate program should not evaluate teaching in a way that ignores the contributions of those faculty who supervise graduate students.

Some units adopt a single scale of weighting that applies to all faculty members. For example, a department might evaluate its members on their contributions to teaching, research, and non-directed service, giving respective weight of 65%, 30%, and 5% to the areas. Another might assign a value of 40% to teaching, 45% to creative endeavors, and 15% to service.

A system based on a single scale is workable in a unit where all faculty have similar assignments, but it is not appropriate for a department having differentiated assignments. Where the responsibilities and goals of the unit dictate specialization of effort among the faculty, the weighting of various areas of endeavor must be sufficiently flexible to accommodate the individual assignments. This is necessary both to permit the unit to make the best use of its resources in performing its tasks and to offer its members an equal opportunity to receive recognition and reward for their accomplishments.

Units having substantial differentiation of faculty assignments often set the relative weight of the areas of professional activity separately for each faculty member. This should be agreed to by the head and each faculty person at the beginning of each evaluation year. Should the head and faculty member be unable to reach agreement, then the head, after consulting with the dean, should prevail. In either case, it is important for the weights to be clearly established before the evaluation period begins. Of course the weights should be subject to re-negotiation in the event of significant changes in assignment during the year.

Other units employ a variable scale of weights in which the individual employee can select, within a predetermined range, the percentage that each area will contribute to the overall evaluation. This practice allows for some variation in individual assignments. Whether the unit specifies that

the percentages should be selected at the beginning of the evaluation year as a part of a planning process or at the end, this approach permits individuals to emphasize the areas in which they expect to accomplish or have accomplished the most in a given year. When adopting a system of variable weighting, a unit must establish maximal and minimal values for each area to assure that the individuals' activities accord with the unit's responsibilities, goals, and expectations.

iv. **The Evaluation Period**

Each unit must also determine the period to be considered in each evaluation. Each unclassified employee must be evaluated every year for the purpose of salary adjustment, and the results of the evaluation must be available in the spring when the University determines its budget for the coming year. The University's budgetary process establishes certain deadlines by which evaluations and the accompanying salary recommendations must be completed or reviewed at various administrative levels. Therefore, many units adopt the calendar year as the period covered in each evaluation. This has the effect of minimizing the time between the end of the evaluation period and the beginning of the fiscal year in which the salary adjustment based on that evaluation takes effect. Other units use the fiscal year as their evaluation period, but since this begins and ends in the middle of the summer, it may not be well suited to departments having many nine-month employees.

Units may also consider including in the annual review accomplishments from more than one year. This is particularly relevant in departments where responsibilities or goals make it desirable to encourage faculty to undertake long-term projects that do not lend themselves to an "in progress" evaluation. In some disciplines, for example, the most respected form of scholarly communication is the book, and it is in the interest of these departments to encourage their faculty to write books. Yet between the inception of a book and its publication (a period that includes research, writing, refereeing, editing, printing, and reviewing) a number of years may pass. Authors who receive full credit for such publications in only one annual evaluation may suffer in salary raises in comparison to colleagues who produce smaller and less significant studies more frequently. Moreover, evaluations of book authors would reveal peaks and valleys in spite of uniform productivity. The inequities in this situation may be exacerbated by variation in the range of salary increases from one year to the next. Therefore, a department that wished to promote book production among its faculty might well establish a system of annual evaluation that included consideration of publications not only from the immediately preceding year but from several earlier years as well. The same logic could apply to a unit wishing to encourage long-term projects in any area.

v. **The Rolling Average**

Departments may also wish to adopt a mechanism to hinder the development of salary inequities deriving from external circumstances. In the simplest systems for assigning salary adjustments, there is a direct relationship between the annual evaluation and the adjustment for the next fiscal year. Under such a system, the unit head might assign salaries for fiscal year 1995 (beginning July 1, 1994) based entirely on the evaluation of individuals' performance in the calendar year 1993. The simplicity of this approach is certainly a virtue, but the system is not without problems. The amount of money available for salary increases depends on legislative appropriations that vary greatly from year to year. Evaluators tend to assign a wider range of raises in years when the average percentage increase is high than when it is low. Performance evaluation is also an annual process, and for most people the evaluation differs somewhat from one year to the next. Thus inequities can easily arise between the person who receives a high evaluation in a year when the range of salary increases is large and the one whose high evaluation comes in a year when the range is small.

One way to mitigate the effects of such variations is to use a rolling average as the basis for assigning salary increases. This method of averaging does not affect the process of evaluation. Rather it influences the way in which the results of the evaluation are used to assign salary increases. The simple system described above employs a single year's evaluation, which is, in effect, weighted at 100%. A rolling average is based on the use of several years' evaluations with a weight assigned to each of them. Thus, for example, the weighting might be 50-25-25. In this case, salaries would be distributed to individuals based on an average of the last three years' evaluations, to which the immediately preceding year (e.g. 1993) would contribute 50% and each of the two years before that (1992, 1991), 25%. A wide range of schemes is possible (e.g., 60-40, 50-35-15, 40-30-20-10), but the unit must agree on a distribution consonant with its goals and other criteria. A unit moving from a single-year system to a rolling average might also wish to consider phasing in the new system over several years.

vi. **Varying Expectations**

Some departments establish different expectations for faculty at different stages of their careers. For example, greater versatility might reasonably be expected of full professors than of assistant professors. Or senior faculty might be expected to provide more broadly based institutional service (e.g., serving on University committees rather than mainly on departmental ones). Or beginning faculty might not be expected to have fully-developed, ongoing research programs. Thus units that vary their expectations might evaluate the same performance at different levels

for persons at different stages of their careers. For example, if an assistant professor and a professor published articles of equal quality in the same journal, the assistant professor's accomplishment might be judged the more meritorious.

vii. **Pitfalls in Developing Systems for Annual Evaluation**

The goals of summative evaluation are positive, but the practices have certain dangers. Many of these can be avoided by sensitive implementation of the system, but some of them are best circumvented in the creation of the system itself.

One very serious threat is that the evaluation system may encourage undesirable behavior. It seems tautological that evaluation should influence behavior, and because the goal of evaluation is to recognize and reward excellence, it would appear that its influence would be salutary. But such is not always the case. Either way, **evaluation has impact**. Some evaluation practices can serve to channel faculty improvement efforts into highly constructive activities. Requiring appropriate sources of data can contribute to positive changes as in the use of student ratings of their learning or the use of peer ratings of instructional materials or tests. Creating appropriate procedures can also contribute to achievement as in devising a means to grant a fair amount of credit to long-term undertakings.

Other evaluation processes can divert efforts from fruitful pursuits into counterproductive ones. Consider some examples concerning the assessment of instruction. A plan for evaluating instruction that placed too much emphasis on the number of students taught could encourage a dilution of standards to increase enrollments. An overemphasis on student ratings of how well they like the content of the course (which is ordinarily desirable) could result in instructors trying to "sell" students in survey courses on inappropriate majors. (The purpose of some exploratory courses is to see if one likes the field; deciding that one does not like a field can reflect attainment of this goal.) A system that judged quality of instruction by student performance on specified achievement tests could promote a narrowing of the curriculum to that which is easily assessed with paper and pencil measures. In addition, focusing can prompt conscious or unconscious teaching to the tests. An evaluation that was narrowly limited to classroom performance would discourage faculty service on graduate committees. It would also fail to take account of such important activities as maintaining office hours and student advisement.

In evaluating research, narrow attention to quantity of publications (which is tempting to naive evaluators because it is both easy and objective) would encourage people to produce numerous "quick-and-dirty" works rather than fewer significant ones. Exclusive attention to works published in a single year would discourage long-term research projects unless the results would lend themselves to piecemeal publication. Overreliance on citation counts of published works could fail to account for the nature of the citations; some works are often cited because of their grave limitations! **Evaluation requires professional judgment**; counting alone does not suffice.

In service, looking primarily at the number of assignments people accepted rather than the quality of their contributions would place a premium on purely nominal participation. For example, it is far easier to count committee memberships than to assess effectiveness (or even counterproductivity) on the committees. But to limit the evaluation to mere counting is to abdicate a mandated administrative responsibility.

Clearly, all those taking part in the creation of evaluation systems should be mindful of the impact that evaluation can have upon the behavior of those evaluated. In devising effective systems, both faculty and administrators must carefully identify the activities and achievements they wish to encourage and then make sure that the system promotes and rewards those contributions. Likewise, all interested parties should identify possible outcomes they wish to avoid and then devise the system to avoid rewarding such behavior.

In addition, those who draft evaluation systems must consider the virtue of simplicity and the danger of "overkill." Because evaluation is an important and ego-involving activity, some departments develop systems that are extremely elaborate and detailed. As a consequence, the process becomes so demanding that its participants lose patience, its potential benefits are threatened, and the normal work of the unit suffers. For example, procedures which require colleagues to rate the quality of each individual publication are probably not worth the required investment of time. A global rating of each person's publications would probably suffice. In any event, the evaluation process should not be so elaborate and time consuming that its participants develop needlessly negative reactions to it.

Yet a certain amount of complexity is necessary to achieve the desired ends. For example, it is necessary to stipulate the relative importance of the major areas of professional activity. Likewise, it is necessary to use computational methods that ensure that the desired weights are, in fact, achieved.

The goal of achieving objectivity in evaluation can also lead to the establishment of systems characterized by excessive precision. Professional activities are too diverse to be accommodated by totally objective means of assessment. Thus it is not ordinarily appropriate to establish a hard and fast system of points for sole authorship, senior authorship, or junior authorship in national

refereed journals, regional refereed journals, unrefereed journals, etc. There is simply too much diversity within any category of activities to make a rigid, predetermined point system defensible. For example, single-authored articles within a given journal often differ greatly in significance and importance so that some would typically merit several times as much credit as others. Effective faculty evaluation must be based on sound professional judgment, and the system of evaluation must allow for the exercise of judgment.

The necessity for professional judgment introduces a degree of subjectivity into evaluation that makes many people uncomfortable. The subjectivity can, however, be reduced. One way to limit the effects of subjectivity on the process is to make use of multiple judgments. Competent persons will ordinarily arrive at similar, although not identical, judgments regarding the merit of professional activities, and the pooled judgment of several competent professionals tends to be more reliable than the sole judgment of any one. Use of multiple raters clearly enhances the reliability with which such things as professional publications, instructional materials, student ratings of teaching effectiveness, and various service activities are evaluated. But there is a tradeoff; use of multiple judges increases both the cost and complexity of the enterprise. Used in moderation, this is likely to be a desirable tradeoff.

A department can organize its evaluation system to provide for multiple raters either by having selected individual faculty members offer advisory judgments to the head or by having a committee perform that task. The use of individual raters has the advantage of offering several independent judgments for comparison, while the committee provides the benefit of shared insights.

Employing multiple judgments does enhance the reliability of the evaluation, but it also involves costs. These include the professional time used and the reduction of confidentiality. When, as is normally the case, the additional raters come from within the department, the possible influence of personal friendships and animosities on the evaluation is a factor to be considered. These problems are exacerbated when the use of multiple raters is carried to an extreme, as in the occasional proposal to have every member of a department rate every other member. Such a process could consume significant amounts of time without any proportional improvement in the result.

b. Implementing the System

Evaluation for salary adjustment is an annual procedure, and University policies specify the most important responsibilities of the people to be evaluated and the initial evaluator for the unit, usually the department head, as well as those who review the evaluations at higher administrative levels. The purpose of this section is to provide some practical advice for department heads and faculties on implementing evaluation at the initial and most important level.

Unit heads must be aware that the University regards effective personnel evaluation as one of the most important administrative responsibilities. It requires informed, thoughtful judgment and should never be done superficially, carelessly, or with undue haste. At the same time, it is a task that often must be performed relatively quickly. The University's budgetary calendar imposes limits on the time available, but equally important is the strain that the process places on working relationships within a unit. Despite all efforts to emphasize its positive aspects, evaluation is a stressful experience for the people subject to it. It is, therefore, important that the evaluator carry out the process with all deliberate speed in order to conclude it in the shortest time consonant with accuracy and fairness.

Also related to timing is the need to have clear and reasonable deadlines for the submission of materials for evaluation. Because the process is an annual one, all but first-year faculty members should have a general idea of when the materials are due. Nevertheless, both professional courtesy and practicality dictate that the exact deadline for a given year should be announced sufficiently early to give people ample opportunity to assemble and submit the required materials. If any person, in spite of reasonable notice, fails to provide the necessary information, the head should send a written reminder. Nevertheless, if after being informed of the possible consequences the person still does not make the materials available, the head may reasonably assign that individual an unsatisfactory evaluation. Since annual evaluation provides the basis for salary recommendations, any faculty member who fails to submit materials provides the head with no justification for recommending a salary increase.

i. Communication

One of the most fundamental elements of effective evaluation is maintaining clear communication. There are two major communication requirements associated with the evaluation process. The first is to create a mutual understanding as to what the individual will be held accountable for in the coming year, the relative importance of each assignment, and the specific methods that will be used to assess performance. The second is to communicate the evaluation clearly and constructively.

Near the beginning of the evaluation period, each unclassified person will meet with the unit head to establish personal goals and objectives for the new evaluation period and to discuss their relative importance within the context of the unit's goals. The head should incorporate the results

of this discussion into a statement of expectations for the individual for the coming year. This statement is intended to guard against misunderstandings regarding work assignments and expectations. Of course, it may be necessary to modify the statement as the year unfolds, since it is impossible to anticipate all contingencies, and modifications should be communicated in writing, again to avoid misunderstandings. The individual should also be reminded to consult the department's written evaluation document regarding the unit's common policies and procedures.

The written evaluation itself must be carefully prepared. Many department heads may find it helpful to compose a draft for each individual early in the process and then to edit these rigorously at a later date. To insure that each evaluation receives adequate care and attention, it is best not to attempt more than two or three written evaluations in one sitting.

The written evaluation should contain four distinct parts, (1) a review of the individual's assignment and the weight attached to each type of responsibility, (2) a summary of the substantive evidence which was used to arrive at evaluative judgments, (3) succinct assessments of effectiveness in performing each responsibility and a statement of the overall evaluation, which must be consistent with the weights assigned to the individual ratings, and (4) where appropriate, formative suggestions for improvement.

After the evaluations have been drafted and edited, it is desirable to arrange for interviews with each unclassified person for purposes of reviewing the draft. The faculty member should be invited to correct any errors of fact and to supply additional documentation to correct possible errors of judgment. The purpose of the interview is to insure that the "final evaluation," prepared after the interview, represents the most valid, fair statement of professional achievement possible.

Once the final evaluation has been prepared, the department head will recommend a salary adjustment for each person evaluated. The recommended percentage increases based on the annual evaluations for persons with higher levels of accomplishment shall exceed those for persons with lower levels of accomplishment. If merit salary categories are used, then the percentage increases recommended for persons in the first category will be higher than those for the second category, which in turn shall exceed those for the third category, etc.

Each unclassified person will review and must have the opportunity to discuss his or her final written evaluation with the individual who prepared it. Before the unit head submits it to the next administrative level, each unclassified person must sign a statement acknowledging the opportunity to review and to discuss the evaluation and his or her relative position within the unit. Because the amount of funds available for merit increases is generally not known at this time, specific percentage increases should not normally be discussed at this stage.

ii. Temporary Assignments

A problem may arise when faculty members are asked to assume an unusually difficult assignment on a temporary basis. Examples include directing an important search, accepting administrative responsibilities on a temporary basis, assuming the duties of a colleague on a temporary basis, chairing a key committee, directing an accreditation study, etc. It is natural to want to express appreciation for performing these "extra duties" by recommending a larger merit salary increase, and this would often be appropriate. However, this may not always be the most appropriate method of providing reward because of the one-time nature of such duties. It may be preferable in certain cases to find some other way to reward performance of temporary duties on a temporary basis (e. g., payment to a Developmental Reserve Account, a summer appointment, support of travel, purchase of equipment, provision of part-time student employees for a year, etc.).

iii. Potential Errors

i. The Error of Unintended Weighting

This problem arises at the stage of integrating the evaluations of individual areas of effort into an equitable overall assessment. There are two major steps in combining separate ratings of component parts (e.g., teaching, research, service) into total ratings. First, a decision or judgment must be made as to how much weight each part is to bear, and these weights should be specified either in the unit's written guidelines or in the head's assignments of individual responsibilities. Second, the evaluator must apply the agreed upon weights in a competent fashion. The second step is purely mechanical, but it is not as simple as it may appear.

Although it seems counterintuitive to some people, the weight of each part of a total score is not a function of possible points. Nor is it a function of average scores awarded for the respective parts. The weight that the respective parts contribute to the total is approximately proportional to their standard deviations.

An elementary example will illustrate the problem and a procedure for dealing with it. This simple procedure approximates the results that could be achieved with more sophisticated statistical tools, and for the purpose of annual evaluations, it will generally yield quite

reasonable results. Suppose that an evaluation system includes only two component areas, research and teaching. Further suppose that the department wishes them to contribute equally to the overall evaluation. Finally, suppose that the department head rates each component on a seven-point scale, with a score of 7 representing the highest possible ranking. On completing the evaluations of the component areas, the head discovers that the ratings assigned in research have a much wider range than those in teaching:

Component	Range
Research	1-7
Teaching	4-6

Each of two faculty members, Dr. Brown and Dr. Green, is rated at the top of the department on one of these components and at the bottom on the other. Since the department intended the two elements to count equally, the two professors ought to receive an equal number of points, indicating that they have been evaluated at the same level. But this is not the case. Dr. Brown is at the top of the department in research, with a score of 7 points, and at the bottom in teaching, with a rating of 4, yielding a total of 11 points. Dr. Green, on the other hand, is at the bottom in research (1) and at the top in teaching (6), giving him a total of only 7 points.

	Research	Teaching	Total
Dr. Brown	7	4	11
Dr. Green	1	6	7

Contrary to intentions, the two components did not count the same. Instead, the component with the larger range of ratings has clearly weighed more heavily in determining the final result.

In order to give equal weight to each component, the ranges need to be made equal. The ratings for the research component have a range from 1 to 7, so the difference between the top and bottom ratings is 6. In teaching, the range is from 4 to 6, and the difference is only 2. In other words, the range of ratings in research is three times the range of ratings in teaching. Thus, if each person's rating in teaching is multiplied by 3, the range will be 12 to 18, and the difference will be 6, i.e., equal to that of the unadjusted scores in research. A recomputation of the rankings using the new scores for teaching gives each professor a total of 19.

	Research	Teaching	Total
Dr. Brown	7	12	19
Dr. Green	1	18	19

By rendering the range of the two distributions equal, their relative contribution to the total has been made equal as well.

Many elementary texts in statistics or educational measurement address this topic and offer solutions. Persons responsible for implementing a department's evaluation system have a responsibility to become familiar with this topic. Assistance can also be obtained through the Office of Planning and Evaluation services.

ii. **The Error of the Permanent Doghouse**

Most individuals display some level of inconsistency in their performance, and sometimes a faculty member has an exceptionally bad year or performs a particular assignment very poorly. If, in evaluating performance in subsequent years, the department head keeps referring to (or silently keeps recalling) these disappointing performances, she or he is committing the **error of the permanent doghouse**. The merit policy focuses on performance in the specified evaluation period. It intentionally provides individuals with the opportunity to "turn a new page." At the same time, it discourages individuals from resting on their laurels. The effective evaluator is able to set aside memories of the past and focus exclusively on performance during the preceding year. It is, of course, equally important to avoid the inverse **error of the permanent halo** in which an evaluator's rating of current productivity is favorably biased by the faculty member's past achievements.

iii. **The Error of Disproportionality**

Evaluative evidence needs to be placed in perspective, and no single finding should sway the conclusion suggested by the bulk of the data. The error of disproportionality occurs when one aspect of the evaluation is given excessive weight in arriving at the overall evaluation. This would be the case if a faculty member were rated as "needs improvement" in teaching after receiving low assessment on one of the six courses taught during the year, even though the evaluations on the other five were all above average. The error of disproportionality would also be present if a head granted a high overall rating based primarily on an individual's exceptional contributions in non-directed service to the department, when that had been designated as a minor responsibility.

iv. **The Error of the Unsupported Judgment**

This error occurs when evaluations are made without reference to supporting data. Simple decrees that an individual is in the "top category" or "needs improvement" do not insure that the evaluation is fair and valid. Because evaluations require human judgment, it is doubtful that any two observers will agree precisely. Nevertheless, if this error is to be avoided, enough relevant evidence should be presented that any two credible observers would arrive at similar evaluative conclusions.

v. **The Error of Non-evaluation**

It sometimes happens that those responsible for preparing evaluations fail to offer judgments of effectiveness or of the "value" of a given performance. Instead, the written "evaluation" consists exclusively of a thorough description of professional responsibilities. Details of what activities the person engaged in do not describe how well those activities were performed. When description is confused with evaluation, the error of non-evaluation occurs.

vi. **The Error of Bashing the Beginner**

New faculty members, particularly those at the beginning of their careers, cannot reasonably be expected to be as productive as they will be later. If a department makes no provision for the need for startup time in evaluating the work of beginners, these persons are placed at unfair and potentially demoralizing disadvantages. One way or another, adequate systems of annual evaluation provide time for newcomers to "get up to speed." For example, a department could have less ambitious publication expectations for people in their first two or three years out of graduate school. Or a department that employs a rolling average could specify that average "ratings" shall be entered into the data for the years preceding the faculty member's employment by Kansas State University (and possibly also, where it would be advantageous, for the first year of employment). Or a department could routinely move any below-average ratings of beginners some specified amount toward the mean, e.g., two-thirds of the way for the first evaluation year and one-third of the way for the second year.

iv. **Persistent Inequities**

Systems of annual evaluation are intended to result in equitable salaries, but inequities do sometimes appear. Therefore, each year salaries should be reviewed in light of the cumulative relative contributions of the faculty. If such a review reveals significant inequities that cannot reasonably be rectified within the limits of a single year's allocation, these inequities should be brought to the attention of the dean and handled within the guidelines for justifying and making salary adjustments on bases outside the annual evaluation.

Effective Faculty Evaluation: Annual Salary Adjustments, Tenure and Promotion

Chapter 4. Evaluation for Tenure

1. Bases of Evaluation

Those making tenure recommendations have an obligation of stewardship to students, consumers of research, the existing community of scholars, and other University constituents to secure the best faculty possible. The faithful exercise of this stewardship requires that a hard question be honestly addressed: "Would the University likely do better if it denied tenure to this person and tried to get a better person for the job?" The "cold heartedness" of this question troubles caring people. But not addressing it would reflect "cold heartedness² to the interests and needs of students, faculty, other University clientele, and the citizens of Kansas.

Wise tenure decisions (be they positive or negative) are never made solely on the basis of individual excellence. Tenure should be granted only to those who have demonstrated individual excellence and whose expertise corresponds to the missions of the University. Therefore, probationary faculty members should be regularly informed of the evolution of institutional missions just as they must be notified of evaluations of their performance.

Tenure should be granted only to those who have demonstrated individual excellence and whose expertise corresponds to the present and anticipated continuing needs of the University. **Thus tenure decisions are based mainly on candidates' contribution** to institutional mission.

2. Relationship of Annual Evaluation to Evaluation for Tenure

Tenure evaluation is not merely the sum of the annual merit evaluations. In practice, the factors of mission relevance of work and supply and demand should receive greater weight in tenure recommendations than in evaluation for annual salary adjustment. Too, tenure decisions are focused on the anticipated future responsibilities of the University while annual salary evaluations are focused on the recent past. Nonetheless, well prepared annual evaluations should, in general, give the probationary faculty member an awareness of his or her comparative performance within the department as well as suggestions for improvement. As such, annual evaluations provide relevant, but not sufficient, information to predict tenure decisions. The relevance of the annual evaluation data to tenure decisions resides in the fact that excellent annual evaluations are a necessary, albeit not sufficient, condition for tenure.

3. Relationship of Tenure to Annual Reappointment of Probationary Faculty

Like other faculty, those on probationary status must be evaluated annually for salary adjustment. In addition, University policy requires that probationary faculty be evaluated annually for reappointment decision. Depending on when units conduct annual evaluation for salary adjustment, departments may or may not combine these two processes.

Even where evaluation for the two purposes is conducted simultaneously, the bases for the evaluations may differ. To illustrate, an excellent probationary faculty member might merit an above average raise in a department, yet be given a terminal contract because of declining enrollment in the program. Or a probationary faculty member in an unusually weak department could conceivably receive the top rating in the department for the purpose of salary adjustment, yet be judged to lack sufficient excellence to merit awarding another regular probationary annual contract. Or a probationary faculty member in a very strong unit might receive below-average ratings for salary adjustment, yet be sufficiently excellent in areas matching unit mission to merit reappointment.

University policy provides maximum durations for probationary status. These are maximum periods--not guaranteed periods. If a department judges that a probationary faculty member's achievement and match to its mission(s) are unlikely to improve sufficiently to warrant tenure, then it is appropriate for the person not to be reappointed for the full probationary period. On the other hand, probationary faculty

4. Mid-Probationary Review

A more formal review of probationary faculty members is ordinarily conducted during their third year of employment. Even more than annual evaluation for reappointment, this review is designed to provide tenure-track faculty members with helpful substantive feedback from faculty colleagues and administrators regarding their accomplishments relative to the institution's missions, objectives, and obligations to its constituents.

The mid-probationary review offers an opportunity for the candidate, the tenured faculty of the department, the department head, and the dean to assess the candidate's performance while there may still be time for mid-course corrections. It also allows for changes in work assignment if there has been inadequate opportunity to develop the candidate's skills or to assess the candidate's performance in some of the areas evaluated. It provides a preliminary test of the match between the candidate's performance and the institution's missions. At times mid-probationary review recommendations may serve to enhance a faculty member's chances of subsequently being tenured. At other times, mid-probationary review may help to dispel a faculty member's unrealistic expectations concerning tenure. The faculty member must realize that a positive review does not insure that tenure will be granted, nor does a negative review necessarily mean that it will be denied.

Departmental procedures for the mid-probationary review should be similar to the department's procedures for the tenure review and are established by the departmental faculty in consultation with the department head and dean. The department head is responsible for making the faculty member's file available to tenured faculty members in the department and is advised by them regarding the person's progress and the departmental needs. Input may also be sought from students, from other faculty and department heads in the college or University, and from outside reviewers. The department head discusses the review with the dean and then provides to the faculty member a letter of assessment, including a summary of faculty comments and suggestions. The department head discusses the review and assessment with the candidate.

5. Versatility and Specialization

Institutional excellence is enhanced by faculty specialization not only in different fields of knowledge, but also in the kinds of professional activities performed. All faculty are occupied with knowledge, but, whereas some emphasize generating or creating knowledge, others focus on integrating, interpreting, transmitting, or applying knowledge. All members of a department need not, and often should not, have the same profile of assigned responsibilities. Considerable specialization is useful.

Nevertheless, specialization of labor carried to excess could seriously limit the extent to which faculty would be able to meet changing conditions or accommodate temporary needs in their departments. Thus institutional excellence is enhanced by faculty versatility. For example, if several people had equal competence in their area(s) of specialization, one who could also perform outside the speciality would be of greater value to the institution. **A major purpose of the probationary period is to provide opportunity to assess a candidate's versatility.** Department heads should therefore plan to provide some variety in assignment in order to give probationary faculty opportunity to develop and to document competence in a variety of kinds of work.

Versatility may be exhibited in numerous of ways. The first dimension that comes to mind is the ability to function well across major areas of work, i.e., teaching, research and other creative endeavors, directed service, non-directed service, and Extension. Thus one might demonstrate versatility by functioning competently in undergraduate teaching, institutional service, and library research. Another might exhibit it with excellent performance in theoretical work, applied research, and undergraduate teaching. Another by superior performance in directed service and laboratory research. Another by excellence in graduate instruction, professional service, and artistic productions.

In addition to versatility across areas of professional activity, versatility within areas also greatly enhances an individual's contribution to institutional missions. Within teaching and advising, one may be able to perform well in various modes of instruction such as undergraduate classroom teaching, undergraduate laboratory instruction, undergraduate clinical supervision, graduate classroom teaching, graduate laboratory instruction, graduate seminar instruction, and graduate clinical supervision. In addition, one may exhibit excellence in undergraduate and/or graduate advising. Faculty may also be proficient in teaching across areas. Thus a history professor who can only teach under-graduate and graduate courses in Latin American history is less valuable as a teacher than one with equal competency in that area who can also offer introductory courses in western civilization and United States history.

Within research and other creative endeavor, one may function competently in a variety of activities. Thus an education professor who is proficient in conducting literature reviews, syntheses, and analyses; developing sound theory; conducting laboratory y-based experimental research; orchestrating school-based experimental research; and performing program evaluations is relatively versatile. So is a physics professor who is very expert in a highly specialized, narrow research area; provides consultation in a

variety of settings for applications of this specialized research; functions well in team research efforts that address somewhat broader content; and contributes productively and insightfully to interdisciplinary research ventures.

Within non-directed service, there are many areas in which one may contribute. These include technical consulting to government, industry, or education; inservice instruction to personnel in government, business, or education; service on departmental, college, or University committees or governance bodies; holding office in professional organizations; and providing professional service for professional organizations. Within directed service or Extension, versatility involves being effective in providing a variety of professional services in diverse settings for different kinds of audiences.

Versatility is an important criterion to be considered in reaching tenure recommendations. Nonetheless, it should not be considered in isolation. We do not seek to have departments staffed by a large number of "jacks of all trades who are masters of none." Rather we seek specialists who retain sufficient flexibility to address a variety of institutional needs--present and future. Moreover a department might be able and willing to tolerate very narrow, non-versatile specialization in a few of its faculty who bring great recognition to the institution by virtue of their eminence.

6. Prudence Suggests Caution

Tenure is a unilateral guarantee of long-term employment by the University. For most practical purposes, the decision is irreversible. Tenuring a person who does not merit it can work a hardship on a whole generation of students and be a source of embarrassment, extra work, and depressed morale to colleagues. Before recommending tenure, prudence and good stewardship requires decision makers to be satisfied beyond reasonable doubt that our constituents, including students, faculty colleagues, and Kansas taxpayers, will be well served by the decision.

If our evaluation procedures were perfect, only correct decisions would be made. However, decision-making processes are never perfectly valid, and, even after our most conscientious efforts and judgments have been made, errors will inevitably occur. Therefore, we should attend to the issue of which kind of error is more serious to the university community--incorrectly making a decision to tenure a person who should not be tenured or incorrectly making a decision not to tenure a person who should be tenured. The more serious kind of error can then be contained, provided that we are willing to trade it off against more errors of the other type. Given the gravity of incorrect decisions to offer tenure, the operational rule should therefore be "When in doubt, don't."

A factor that has, on occasion, influenced tenure decisions is the fear that if an incumbent probationary faculty member is not tenured, the position will be lost to the department. Administrators should make every effort to ensure that departments are not penalized when they responsibly recommend against tenuring questionable candidates.

7. Assessment Should Reflect Responsibilities

In the discussion of annual evaluation there was a generally reasonable presumption that a direct relationship exists between the responsibilities assigned to an individual and the weight allocated to those responsibilities in the evaluation for merit salary increase. In evaluation for tenure, this same presumption is appropriate, although it requires considerably more qualification.

On the one hand, there is a strong case for a correspondence between job assignment and tenure evaluation. Thus if a person is assigned a heavy teaching load, there should be a corresponding lower expectation for productivity in other areas. If another person is assigned little or no teaching, directed service, or Extension responsibilities, then it would be appropriate to expect much greater productivity in research or other creative endeavors.

It is particularly important for those who participate in tenure recommendations at the college and University level to keep in mind that faculty assignments differ. Assignments reflect differences in departmental missions. Some probationary faculty members are assigned much, much heavier teaching, directed service, or Extension responsibilities than others whose departments have strong research missions. It would be unreasonable to expect as much research productivity in the former as in the latter .

Vital as this point is, it needs qualification. For one thing, it could, if carried to the extreme, interfere with evaluation for versatility. To illustrate, suppose a person had a relatively heavy teaching assignment. Suppose the person performed the instructional duties with excellence but totally failed to fulfill the relatively modest expectations in research and other creative endeavors. The person should not normally be awarded tenure. Similarly, consider a person with, say 80% of his or her time assigned to research who is expected to teach only an occasional class. Further suppose this person's research is excellent but the teaching is unsatisfactory. The individual should not normally be tenured.

Qualification is also needed to cover unprofessional or incompetent behavior. Suppose a probationary faculty member performs well on most indicants concerning instruction, has a fine record of research or other creative endeavor, and a solid performance in non-directed service. Yet there are in the records several, independent, credible complaints by students of research exploitation with regard to credit for publications, sexual harassment, or violation of the rights of human subjects. Although a narrow numeric

calculation of such a person's performance might yield an acceptable "score," such a person should not be tenured because tenure should be awarded only to those who are excellent overall and who are at least adequate in every significant aspect of job performance. Similarly, behaviors that adversely affect collegiality or are chronically disruptive would properly influence tenure decisions.

8. Faculty Participation

The Provost prepares the University's final recommendations regarding the granting of tenure and the awarding of promotions for submission to the Board of Regents for final approval. The Provost's action culminates a process that includes recommendations from the departmental faculty, the department head, a college tenure and promotion committee in those colleges having one, and the dean of the college.

a. Departmental Faculty Input

Departments have considerable latitude concerning how faculty input will be obtained. Regardless of the system adopted, each faculty member's recommendation shall be transmitted to the department head. In most departments the tenured faculty members all participate directly and fully in the process. In some larger departments, the faculty elects a tenure and promotion committee. Such a departmental committee can (a) obtain input from the departmental faculty eligible to recommend and/or (b) synthesize the recommendations of individual faculty members into summary statements.

The departmental faculty eligible to recommend may participate in any of several ways, each having advantages and disadvantages. One mode is for the group to meet after each member has had opportunity to study the candidates' materials, to discuss the candidates, and to take a ballot of recommendations that will be forwarded to the department head. This has the benefit of providing for the sharing of information, but it is attended by vulnerability to being swayed (or even coerced) by a strong personality.

At the opposite extreme, the tenured faculty might never meet but have each member individually provide recommendations to the department head. This offers protection against any one member having undue influence, but the cost is its lack of provision for sharing of information, insights, and perspective.

Some arrangements seek to capture the advantages of both of the above methods and simultaneously to minimize the disadvantages of each. In one such method, in its simplest form, each member of the tenured faculty first reviews the candidates' materials and reaches tentative recommendations. Then the tenured faculty meet to share, discuss, and even debate, judgments and conclusions. Finally, the group adjourns and each member separately communicates to the department head the recommendation that she or he believes to be appropriate.

Some variations on this arrangement are possible. For example, in the event that additional information is felt to be needed, the tenured faculty should seek it and reconvene after it is available. Or in cases in which unfavorable information surfaces from discussion that the candidate reasonably would not have anticipated, the candidate might be invited to respond in writing or in person at a future meeting. Indeed, a department's tenured faculty might routinely permit candidates to make oral presentations to the body if they so choose.

b. Relationship of Department Head and Dean with Advisory Committees

The administrator is sometimes in an awkward position during the faculty input process. One issue concerns whether the administrator should be present and provide input to the committee. There are pros and cons, and units will differ in the ways in which they wish to function.

Another issue concerns certain kinds of damaging information about a candidate that may be known only to the administrator. Favorable information concerning a candidate is, of course, usually known to the faculty, but the administrator may possess damaging information that has ethical considerations of confidentiality attached. Faced with this situation, what is the administrator to do? Sharing it with the senior faculty is unacceptable when doing so would be unethical. Explaining to the senior faculty that damaging information exists but not sharing it is another unacceptable alternative; people should not be asked to make recommendations on the basis of what they do not know.

The best option would be for the administrator, in reaching his or her own recommendation to take account of the information and to note the fact that the faculty recommendations may not reflect it. The provision for separate recommendations by administrators exists, in part, to provide for such situations. There are times when an administrator definitely should disagree with the majority faculty recommendation. Recalling the "When in doubt, don't" adage, the direction of this disagreement should most often be for the administrator to recommend against tenure or promotion when the faculty or faculty committee have recommended in favor of it.

c. The Option of College Tenure and Promotion Committees

For a variety of sound reasons, some colleges have college committees to advise on tenure and promotion decisions. Each college has great latitude in deciding whether to have such a committee and, if it elects to have one, in such matters as determining how the members are chosen and what the responsibilities and procedures of the committee shall be. Following are some considerations for colleges electing to have tenure and promotion committees.

Since the function of a tenure and promotion committee is to provide faculty input, most colleges' faculties will probably wish to provide for faculty election of all or most of the committee members.

Members may be selected from the faculty at-large from the college or from departmental constituencies. People often feel some loyalty to the constituency that selected them. If all members are selected at-large, the members may be more free to "vote their consciences." At-large membership also avoids issues of proportional representation among departments of unequal size.

Concern for both continuity and experience argue for multi-year terms. At the same time it may be desirable to prevent certain people from becoming permanently entrenched in the college promotion and tenure committees. Something along the lines of staggered three-year terms might be appropriate. A college's bylaws could also stipulate that a person may not be elected to two (or three) consecutive full terms. Staggered terms also enable a faculty--and especially a nominating committee--to seek balance of perspectives with regard to disciplines, age, gender, ethnicity, etc., among the members of the committee.

Obviously, only tenured persons are eligible for selection. College rules should provide either that (a) those selected for the committee may participate in the process only for persons seeking tenure or seeking promotion to a rank equal to or below the rank of the person making recommendations or (b) those selected for the committee may participate in all of the recommendations that come before the committee.

9. **Format of Recommendations**

At the departmental level, means are needed for securing recommendations from the eligible faculty as well as summarizing and forwarding the comments and recommendations. Similar issues apply to college tenure and promotion committees in those colleges having them. Units have great latitude in choosing how to perform these functions.

At one extreme, departmental faculty or college committee recommendations could be conveyed as a simple vote (e.g., 8 yes, 1 no, 1 abstain). This, of course, has the disadvantage of not revealing the basis for the recommendations. At the other extreme, each person making a recommendation might write a letter. This has the problem of being difficult to summarize and interpret. It is best if units forward both their vote and their bases for reaching the recommendations. Indeed, university policy requires that all written comments by tenured faculty of the department be sent forward at every level and good practice dictates that college tenure and promotion committees do the same.

Appendix B provides an example of a form that a unit might consider using. It incorporates four useful elements that a departmental faculty or college tenure and promotion committee should consider, even if it elects to adopt a different form.

The first element is to explicitly render legitimate the consideration of credible information that is not in the candidate's materials. Candidates prepare their files to cast themselves in a favorable light; they cannot realistically be expected to reveal all of their limitations. Yet it is important for the limitations to be considered.

The second element is to prompt faculty to supply bases for their recommendations.

The third element is to prompt communication of the strengths of one's recommendation, thereby providing subsequent considerers with more information. E.g., a 5 to 3 vote might have a rather different meaning if it were 5 "definitely yes" against 3 "probably no" than if it were 5 "probably yes" against 3 "definitely no." The form contains an even number of categories because tenure and promotion are inherently dichotomous decisions. Of course a person always can abstain, but it does not seem helpful to invite abstention.

Finally, the signature of the person making the recommendations is required. This enables adequate records to be maintained. The copy of the form that is forwarded to the next level can, if the unit so elects, have the signatures removed. (The signed original must be retained in the files.)

10. **External Peer Reviews**

Outside reviews of credentials by recognized and respected professionals in the candidate's field are often used as a source of information considered in arriving at tenure decisions. (See the discussion of the topic on pages 31-32 in the section on promotion.)

Outside reviewers cannot, of course, ordinarily speak to institutional needs. This limitation of their perspective should be kept in mind when their opinions are used in making tenure recommendations. They can, however, provide valuable, expert, objective information concerning individual merit in the context of prevailing supply and demand conditions.

11. **Duration of the Probationary Period**

No probationary time is transferred into Kansas State University. Candidates who prove themselves ready in less than maximum time (and for whom there exists sufficient program need, etc.) may apply early for

tenure and be favorably considered. Candidates who have not demonstrated their merit beyond reasonable doubt in less than the maximum time have the benefit of the full probationary period (if the University chooses to reemploy them each year for this period).

In the past, probationary faculty who had prior experience at tenure-earning ranks could, by mutual agreement, transfer in some time, thereby shortening their maximum probationary period at Kansas State University. This sometimes led to problems if the abbreviated period was not long enough to enable the candidate to demonstrate sufficient accomplishments at Kansas State University for a favorable tenure recommendation. More time would often have been beneficial to the candidate as well as to the University.

12. Assistant Professors Are Not Tenured

The granting of tenure should be based primarily on candidates' contribution to the institution's missions; this contribution is an interaction between the excellence of the individual and the responsibilities of the institution. Candidates lacking sufficient individual excellence to justify the rank of associate professor are not deemed to have contributed sufficiently to institutional missions to justify tenure. Therefore, tenure is not granted below the rank of associate professor except in special circumstances approved by the provost.

Effective Faculty Evaluation: Annual Salary Adjustments, Tenure and Promotion

Chapter 5. Evaluation for Promotion

1. Relationship of Tenure and Promotion

Two special cases concerning the relationship between of promotion to the rank of associate professor and tenure should be mentioned. First, as noted in the previous section, a negative decision concerning promotion to the rank of associate professor during the final year of probation ordinarily constitutes a decision that the candidate's work lacks sufficient excellence to justify tenure.

Second, a probationary faculty member might have truly exceptional performance but be employed in a program that was declining or facing elimination. It would be possible in this case for the person to be promoted in a shorter-than-average time, yet be denied tenure because of factors of institutional need that were beyond his or her control. Thus tenure is neither adequately predicted by annual salary adjustment recommendations nor even by promotion recommendations.

2. Kinds of Errors

Although it does not have the gravity of a tenure decision, a promotion decision for a tenured person also involves a very long-term commitment by the University. Here, too, prudence suggests caution in making a favorable decision. On balance, the lesser harm to the greatest number of people is caused by an error in making an incorrect negative promotion recommendation than in making an incorrect positive one.

An incorrect decision to promote a person can create long-term problems. The University is committed to providing the person the benefits of the rank for decades to come even if productivity is seriously deficient. This can cause embarrassment and create morale problems for colleagues.

An incorrect unfavorable decision is likely to be much less serious because it can be righted in any subsequent year. Although it will probably cause a temporary morale problem for the person not promoted and could trigger a resignation, it does not typically create any irreversible long-term problems. Here, too, however, ethical and humanitarian concerns dictate that the University should minimize the number of such errors to the extent that this is possible without increasing the number of the other kind.

3. External Peer Reviews

For promotion to the rank of full professor, reviews by outside peers should be a standard part of the materials assembled. By the time a person merits consideration for promotion to full professor, the record of achievements should be extensive and the reputation with appropriate populations should be well established. Hence external peer reviews have their greatest utility in reaching decisions regarding promotion to the rank of full professor. Their limitation of not addressing institutional missions and needs is not a disadvantage in promotion considerations because promotion is based mainly on individual excellence, which peer reviews are particularly well suited to address.

As mentioned above, outside peer reviews are often used in tenure deliberations and in decisions concerning promotion to the rank of associate professor. These two uses could well be considered options available to the department faculty, the department head, the dean, or the candidate.

Selection of outside reviewers should include some named by the candidate as well as some named by the department. It would seem reasonable for equal numbers to be chosen by each party. Four or six outside reviewers might be a reasonable number.

Reviewers should be selected on the basis of their qualifications to assess the work of the candidate. This does not require personal acquaintance with the candidate. Care should be taken to avoid the candidate's co-authors, dissertation committee members, or other persons who may be perceived to have potential biases. At the same time, personal acquaintance and even some degree of professional association with the candidate should not necessarily disqualify a reviewer. Faculty members make contact with

The reviews should be invited by the head and sent directly to the head with assurance that, to the extent allowable by law, the candidate will not have access to them.

Once the reviews have been obtained, it is appropriate for the department head to provide a brief cover statement to be included in the candidate's file for the edification of those who will review it. The cover statement should, as a minimum, indicate (a) who selected each reviewer, (b) the qualifications of each reviewer, and (c) the aspect(s) of the candidate's performance that each reviewer is well qualified to assess.

Outside review is more useful for evaluating research and professional service than other forms of service or teaching. It is important also to recognize that it can have a built-in bias favoring those with heavy research assignments. Research publications are easily sent off for review. Teaching is less easily "packaged." For this reason, the outside reviewers should be provided with a clear sense of the candidate's assignments in each area of professional activity.

Another way to avoid bias is to secure reviews from peers who are experts in what the candidate has been expected to do most. Thus a person with heavy research responsibilities might very appropriately be evaluated by eminent researchers.

A person with heavy teaching responsibilities would appropriately be assessed by outside reviewers known for their excellent instruction or for their scholarship in the evaluation of instruction in the field of interest. In identifying such persons, one might consider persons who (a) have received awards for outstanding teaching that are based on rigorous criteria, (b) are recognized for scholarship in the evaluation of college teaching, (c) are widely recognized by peers and students as outstanding teachers, (d) provide editorial services to professional journals concerned with teaching in the candidate's field, (e) publish research in the teaching of the candidate's discipline, and (f) publish informed, thoughtful commentaries concerning instruction in the discipline. Such reviewers can not, of course, observe the candidate's teaching, but this is not a problem because direct observation is not a very good basis for summative evaluation of teaching. Outside reviewers of instructional effectiveness can examine such packageable indicators of teaching effectiveness as syllabi, tests, instructional materials, and bibliographies.

Effective Faculty Evaluation: Annual Salary Adjustments, Tenure and Promotion

Appendix A

Selected Excerpts from *Scholarship Reconsidered*

Scholarship Reconsidered: Priorities of the Professoriate. A Special Report by Ernest L. Boyer. The Carnegie Foundation for the Advancement of Teaching. Princeton University Press, 1990.

Scholarship Reconsidered is a welcome, new perspective in the debate about the work and responsibilities of higher education, presented in eighty pages of thought-provoking reading. The following excerpts are offered only as an "appetizer," with the hope that faculty members will choose to read the complete report before they endeavor to establish evaluation standards, criteria, and guidelines for their departments.

"What we urgently need today is a more inclusive view of what it means to be a scholar--a recognition that knowledge is acquired through research, through synthesis, through practice and through teaching. We acknowledge that these four categories--the scholarship of discovery, of integration, of application, and of teaching--divide intellectual functions that are tied inseparably to each other. Still, there is value, we believe, in analyzing the various kinds of academic work, while also acknowledging that they dynamically interact, forming an interdependent whole. Such a vision of scholarship, one that recognizes the great diversity of talent within the professoriate, also may prove especially useful to faculty as they reflect on the meaning and direction of their professional lives." (p. 24)

"What we propose, in short, is that faculty expectations and related evaluation not only be *broadened* but that they be *individualized* and *continuous* as well. If faculty are to build on their strengths and contribute constructively to the institutions where they work, evaluation criteria must be tailored to personal talents, as well as campus needs. And it is especially important, we believe, that the criteria used reflect changing patterns of personal and professional growth across a lifetime. Once again, *diversity*, not uniformity, is the key." (p. 50)

"Broadening scholarship has implications not only for individuals but for institutions, too. Today's higher education leaders speak with pride about the distinctive missions of their campuses. But such talk often masks a pattern of conformity. Too many campuses are inclined to seek status by imitating what they perceive to be more prestigious institutions. We are persuaded that if scholarship is to be enriched, every college and university must clarify its own goals and seek to relate its own purposes more directly to the reward system for professors." (p. 53)

In order to return the term *scholarship* to the broader meaning, Boyer proposes that the work of the professoriate be thought of as having four separate, yet overlapping, functions; the scholarship of *discovery*; the scholarship of *integration*; the scholarship of *application*; and the scholarship of *teaching*.

The Scholarship of Discovery (pp. 17-18)

"Scholarly investigation, in all the disciplines, is at the very heart of academic life, and the pursuit of knowledge must be assiduously cultivated and defended. The intellectual excitement fueled by this quest enlivens faculty and invigorates higher learning institutions, and in our complicated, vulnerable world, the discovery of new knowledge is absolutely crucial."

"The scholarship of discovery, at its best, contributes to the stock of human knowledge but also to the intellectual climate of a college or university. Not just the outcomes, but the process, and especially the passion, give meaning to the effort."

The Scholarship of Integration (pp 18-21)

"By integration, we mean making connections across the disciplines, placing the specialities in larger context, illuminating data in a revealing way, often educating nonspecialists, too. . . . what we mean is serious, disciplined work that seeks to interpret, draw together, and bring new insight to bear on original research."²

The Scholarship of Application (pp. 21-23)

"The scholarship of application, as we define it here, is not a one-way street. Indeed, the term itself may be misleading if it suggests that knowledge is first 'discovered' and then 'applied'. The process we have in mind is far more dynamic. New intellectual understandings can arise out of the very act of application--whether in medical diagnosis, serving clients in psychotherapy, shaping public policy, creating an architectural design, or working with the public schools. In activities such as these, theory and practice interact, and one renews the other."

"Clearly, a sharp distinction must be drawn between citizenship activities and projects that relate to scholarship itself. . . . [A]ll too frequently, service means not doing scholarship but doing good. To be considered scholarship, service activities must be tied directly to one's special field of knowledge and relate to, and flow directly out of, this professional activity. Such service is serious, demanding work, requiring the rigor--and the accountability--traditionally associated with research activities."

The Scholarship of Teaching (pp. 23-24)

"In the end, inspired teaching keeps the flame of scholarship alive. Almost all successful academics give credit to creative teachers--those mentors who defined their work so compellingly that it became, for them, a lifetime challenge. Without the teaching function, the continuity of knowledge will be broken and the store of human knowledge dangerously limited."

"As a scholarly enterprise, teaching begins with what the teacher knows. Those who teach must, above all, be well informed, and steeped in the knowledge of their fields. Teaching can be well regarded only as professors are widely read and intellectually engaged."

"Teaching is also a dynamic endeavor involving all the analogies, metaphors, and images that build bridges between the teacher's understanding and the student's learning. Pedagogical procedures must be carefully planned, continuously examined, and relate directly to the subject taught. . . . [G]reat teachers create a common ground of intellectual commitment. They stimulate active, not passive, learning and encourage students to be critical, creative thinkers, with the capacity to go on learning after their college days are over."

Appendix C

Sources of Assistance

A major purpose of this document is to identify and discuss several issues concerning summative faculty evaluation systems that relate to institutional excellence. In this appendix, the focus shifts to assistance for persons charged with the responsibilities of evaluating faculty performance. This discussion is divided into three subsections: administrative commitment to providing systematic orientation for heads of units, general sources of help regarding summative faculty evaluation, and sources of help regarding the specific domains of faculty responsibility.

Administrative Commitment

Faculty evaluations are essential for personnel decisions and for academic productivity. Excellence in the conduct of these evaluations requires a commitment of support from University administrators at the highest levels. Department heads and deans need to receive systematic orientation to the ideas, content, and processes of faculty evaluation, as well as opportunities for developing, renewing, and enhancing their competence in faculty evaluation. Follow-up to seminars and workshops on faculty evaluation has been demonstrated to be critical for retention and continued quality performance (Centra, 1979; Miller, 1987).

General Sources of Help on Faculty Evaluation

Following are lists of three kinds of help that may be considered for those charged with developing or implementing summative faculty evaluation.

Published Works

Some excellent printed materials are available on the overall topic of faculty evaluation. Some of the more highly regarded materials are cited below.

Books

Centra, J. A. (1979). *Determining faculty effectiveness: Assessing teaching, research, and service for personnel decisions and improvement*. San Francisco: Jossey-Bass.

This book is regarded by some authorities to be the best single source on faculty evaluation. Centra relies heavily on research to support his assertions. He emphasizes the need for local involvement in developing evaluation systems that are pertinent to the local conditions. The book covers goals and procedures of faculty evaluation, uses and limitations of student ratings, benefits of self-assessment and self-analysis, evaluation by

colleagues, measures of student learning, assessment of research, advising, and public service, and assembling data for making decisions.

Chait, R. P., & Ford, A. T. (1982). *Beyond traditional tenure: A guide to sound policies and practices*. San Francisco: Jossey-Bass.

The authors provide a thoughtful and balanced treatment of the pros and cons of tenure, its alternatives, and its major variants. The book, and particularly Chapter 7, contains well reasoned, research-based recommendations for sound personnel policies and practices.

Menges, R. J., & Mathis, B C. (1988). *Key resources on teaching, learning, curriculum, and faculty development: A guide to the higher education literature*. San Francisco: Jossey Bass.

This broad-scope book reviews a rich variety of published materials. Its seven chapters address literature on teaching, learning, curriculum, and faculty development; teaching and teaching effectiveness; learners and the learning process; college and University curricula; faculty and staff development; periodicals and reference tools; and key trends and issues for research and practice.

Miller, R. I. (1987). *Evaluating faculty for promotion and tenure*. San Francisco: Jossey-Bass.

Coverage includes guidelines for improving evaluation systems, evaluating teaching, evaluating scholarship and service, assessing other aspects of professional effectiveness, understanding purposes and procedures of academic promotion and tenure decisions, monitoring compliance with the law, administrative roles in promotion and tenure processes, and preparing for future evaluation and personnel needs. Three appendices present a variety of resource materials.

Seldin, P. (1980). *Successful faculty evaluation programs: A practical guide to improve faculty performance and promotion/tenure decisions*. Crugers, NY: Coventry Press.

The author's purpose was to help faculty and administrators avoid the pitfalls and develop the skills and sensitivity needed for successful evaluation programs. Ten chapters cover what is involved in evaluating faculty performance, current evaluation practices, student evaluations, peer evaluations, self-evaluation, other ways of evaluating teaching, institutional service and student advising, research and publications, models of evaluation, and a guide to successful faculty evaluation.

Seldin, P. (1991). *The teaching portfolio*. Boston: Anker Publishing.

This paperback introduces teaching portfolios, describes how to prepare them, lists 30 items that might be included, and discusses how they might be used. One of the book's strengths is eight sample teaching portfolios. A possible limitation is the focus on portfolios as samples of teachers' best work--analogous to artists' portfolios. This approach emphasizes using portfolios more for development than for evaluation, although the latter also receives some attention.

Newsletter

Assessment Update: Progress, Trends, and Practices in Higher Education. San Francisco: Jossey-Bass.

This quarterly newsletter's articles address such topics as linking assessment and teaching, assessment activities at large research universities, politics and assessment in the University, faculty response to assessment at a major research University, and how assessment results are being used. Regular columns present profiles of assessment activities on specific campuses, the status of state-mandated assessment initiatives, and available assessment measures. The annual cost is \$60.00.

Periodical Literature

Among the numerous articles published about faculty evaluation, here are some worth considering.

Braskamp, L. A., Fowler, D. L., & Ory, J. C. (1984). Faculty development and achievement: A faculty's view. *Review of Higher Education*, 7, 205-222.

Braskamp, L. A., Muffo, J. A., & Langston, I. W. (1978). Determining salary equity: Policies, procedures, and problems. *Journal of Higher Education*, 49, 231-246.

Craven, E. C. (1980). *Academic program evaluation: New directions for institutional research*, No.27. San Francisco: Jossey-Bass.

Feldman, K. A. (1987). Research productivity and scholarly accomplishment of college teachers are related to their instructional effectiveness: A review and exploration. *Research in Higher Education*, 26, 227-298.

Holley, W. H. & Field, H. S. (1977). The law and performance evaluation in education: A review of court cases and implications for use. *Journal of Law and Education*, 6, 427-448.

Root, L. S. (1987). Faculty evaluation: Reliability of peer assessments of research, teaching, and service. *Research in Higher Education*, 26, 71-84.

Smith, K. L. & Clark, R. W. (1987). Executive Development Center: Taking the guesswork out of staff selection. *Journal of Extension*, XXV (Winter), 19-22.

Smock, H. R. (1982). Planning for an evaluation network and institutionalization. In R. Wilson (Ed.), *Designing academic program reviews: New directions for higher education*, No. 37 (pp. 67-73). San Francisco: Jossey-Bass.

Verna, O. S. & Bhaskar, V. V. (1988). Scientists' productivity: Miscellaneous activities are integral part of assessment system. *Journal of Extension Systems*, 4(1), 29-36.

Workshops and Seminars

Some organizations provide workshops and seminars on faculty evaluation and faculty development. The names and addresses of some of these organizations are given below.

On-Campus Resources

Center of Faculty Evaluation and Development
Kansas State University
College Court Building, 1615 Anderson Avenue, Manhattan, KS

Planning and Evaluation Services
Kansas State University, Fairchild Hall

Off-Campus Resources

Comprehensive Data Evaluation Services, Inc.
6730 North Camino Padre Isidoro, Tucson, AZ 85718

Division of Continuing Education
Ohio University, Athens, OH 45701

Consultants

Consultants willing to work with universities and colleges on faculty evaluation include the following:

On-Campus Resources

William Cashin
Center for Faculty Evaluation and Development
Kansas State University
College Court Building 1615 Anderson Avenue

Victoria Clegg
Educational Improvement, Fairchild Hall, Kansas State University

Off-Campus Resources

Lawrence A. Braskamp
University of Illinois -- Chicago Circle, Chicago, IL

Peter Cohen
Medical College of Georgia, Augusta, GA

John Hammonds
University of Arkansas, Fayetteville, AR

Peter Seldin
Pace University, Pleasantville, NY

Personnel Development
Ohio Cooperative Extension Service, The Ohio State University, Columbus, OH

Literature on Specific Domains of Faculty Responsibility

Teaching

Bennett, J. B., & Chater, S. S. (1984). Evaluating the performance of tenured faculty members. *Educational Record*, 65, 38-41.

Braskamp, L. A., Brandenburg, D. C., & Ory, J. C. (1984). *Evaluating teaching effectiveness: A practical guide*. Beverly Hills, CA: Sage.

Cashin, W. E. (September, 1988). *Student ratings of teaching: A summary of the research*. IDEA Paper No. 20. Manhattan, KS: Center for Faculty Evaluation and Development.

Cashin, W. E. (September, 1989). *Defining and evaluating college teaching*. IDEA Paper No. 21. Manhattan, KS: Center for Faculty Evaluation and Development.

Cashin, W. E. (January, 1990). *Student ratings of teaching: Recommendations for use*. IDEA Paper No. 22. Manhattan, KS: Center for Faculty Evaluation and Development.

Cashin W. E. (1990). Assessing teaching effectiveness. In P. Seldin (Ed.). *How administrators can improve teaching*. San Francisco: Jossey-Bass.

Cashin, W. E. (1990). Students do rate different academic fields differently. In M. Theall, & J. Franklin (Ed.). *Student ratings of instruction: Issues for improving practice: New Directions for Teaching and Learning*, No. 43. San Francisco: Jossey-Bass.

Centra, J. A. (1979). *Determining faculty effectiveness: Assessing teaching, research, and service for personnel decisions and improvement*. San Francisco: Jossey-Bass. See Chapters 2-5, pages 17-118.

Dunkin, M. J. (1986). Research on teaching in higher education. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd Ed.) (pp. 754-777). New York: Macmillan.

French-Lazovitz, G. (Ed.). (1982). *Practices that improve teaching evaluation*. San Francisco: Jossey-Bass.

McLaughlin, G. W., Montgomery, J. R., Gravelly, A. R., & Mahon, B. T. (1981). Factors in teaching assignments: Measuring workload by effort. *Research in Higher Education*, 14, 3-17.

Miller, R. I. (1987). *Evaluating faculty for promotion and tenure*. San Francisco: Jossey-Bass. See Chapter 3, pages 31-55 and Appendices A and B.

Seldin, P. (1980). *Successful faculty evaluation programs: A practical guide to improve faculty performance and promotion/tenure decisions*. Crugers, NY: Coventry Press. See Chapters 3-6, pages 36-107.

Seldin, P. (1991). *The teaching portfolio*. Boston: Anker Publishing.

Research

Blackburn, R. T., Behmyer, C. E., & Hall, D. E. (1978). Research note: Correlates of faculty publications. *Sociology of Education*, 51, 132-141.

Centra, J. A. (1979). *Determining faculty effectiveness: Assessing teaching, research, and service for personnel decisions and improvement*. San Francisco: Jossey-Bass. See Chapter 6.

Creswell, J. W. (1985). Faculty research performance: Lessons from the sciences and social sciences. *ASHE-ERIC Higher Education Report No. 4*. Washington, DC: Association for the Study of Higher Education.

Jones, L. V. (1980). The assessment of scholarship. In E. H. Loveland (Ed.), *Measuring the hard-to-measure: New directions for program evaluation*, No. 6 (pp. 1-20). San Francisco: Jossey-Bass.

Miller, R. I. (1987). *Evaluating faculty for promotion and tenure*. San Francisco: Jossey-Bass. See Chapter 4, pages 56-70 and Appendices A and B.

Mortimer, K. P., Bagshaw, M., & Masland, A. T. (1985). Flexibility in academic staffing: Effective policies and practices. *ASHE-ERIC Higher Education Research Reports*, No. 5. Washington, DC: Association for the Study of Higher Education.

Root, L. S. (1987). Faculty evaluation: Reliability of peer assessments of research, teaching, and service. *Research in Higher Education*, 26, 71-84.

Seldin, P. (1980). *Successful faculty evaluation programs: A practical guide to improve Faculty performance and promotion/tenure decisions*. Crugers, NY: Coventry Press. See Chapter 8, pages 127-141.

Directed and Non-Directed Service

Boyer, C. M., & Lewis, D. R. (1985). And on the seventh day: Faculty consulting and supplemental income. *ASHE-ERIC Higher Education Research Reports*, No. 3. Washington, DC: Association for the Study of Higher Education.

Braskamp, L. A., Muffo, J. A., & Langston, I. W. (1978). Determining salary equity: Policies, procedures, and problems. *Journal of Higher Education*, 49, 231-246.

Centra, J. A. (1979). *Determining faculty effectiveness: Assessing teaching, research, and service for personnel decisions and improvement*. San Francisco: Jossey-Bass. See Chapter 6, pages 119-135.

Elman, S. E., & Smock, S. M. (1985). *Professional service and faculty rewards: Toward an integrated structure*. Washington, DC: National Association of State Universities and Land-Grant Colleges.

Miller, R. I. (1987). *Evaluating faculty for promotion and tenure*. San Francisco: Jossey-Bass. See Chapter 4, pages 56-70 and Appendix A.

Root, L. S. (1987). Faculty evaluation: Reliability of peer assessments of research, teaching, and service. *Research in Higher Education*, 26, 71-84.

Seldin, P. (1980). *Successful faculty evaluation programs: A practical guide to improve Faculty performance and promotion/tenure decisions*. Crugers, NY: Coventry Press. See Chapter 7, pages 108-126.

Extension

Gerber, J. H. (1985). Extension specialists: A self-analysis. *Journal of Extension*, XXIII (Winter), 8-11

Ladewig, H., & Shiao, K. S. (1983). *Development of personnel appraisal procedures to measure the job performance of County Extension agents: Summary of research*. No. 30. Columbus, OH: Ohio State University Department of Agricultural Education.

Patterson, T. F. (1987). Refining performance appraisal: A reliable and useful method. *Journal of Extension*, XXV (Winter), 16-18.

Thompson, O. E., & Gwynn, D. (1989). Improving Extension: Views from agricultural deans. *Journal of Extension*, XXVII (Spring), 7-10.