

# **Evaluating Teaching Effectiveness**

**A Practical Guide**

**Larry A. Braskamp  
Dale C. Brandenburg  
John C. Ory**

# ***Evaluating Teaching Effectiveness***

-----



# **Evaluating Teaching Effectiveness**

## **A Practical Guide**

**Larry A. Braskamp  
Dale C. Brandenburg  
John C. Ory**

With the Assistance of

**Eileen Kohen  
Paul W. Mayberry**



SAGE PUBLICATIONS, INC.  
Beverly Hills / London / New Delhi

Copyright © 1984 by Sage Publications, Inc.

All rights reserved. No part of this book may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

*For information address:*

SAGE Publications, Inc.  
275 South Beverly Drive  
Beverly Hills, California 90212

SAGE Publications India Pvt. Ltd.  
C-236 Defence Colony  
New Delhi 110 024, India



SAGE Publications Ltd  
28 Banner Street  
London EC1Y 8QE, England

Printed in the United States of America

### **Library of Congress Cataloging in Publication Data**

Braskamp, Larry A.

Evaluating teaching effectiveness.

1. College teachers—Rating of—United States.
  2. College teaching—United States—Evaluation.
- I. Brandenburg, Dale C. II. Ory, John C. III. Title.  
LB2333.B7 1984 378'.125 84-9942  
ISBN 0-8039-2341-4  
ISBN 0-8039-2342-2 (pbk.)

**FIRST PRINTING**

## **CONTENTS**

<b>Preface and Acknowledgments</b>	<b>7</b>
<b>How to Use this Guidebook</b>	<b>11</b>
<b>1. Defining Good Teaching</b>	<b>15</b>
<b>2. Evaluating Teaching—Some Considerations</b>	<b>19</b>
<b>3. An Approach for Evaluating Teaching</b>	<b>29</b>
<b>4. Collecting Evaluative Information about Teaching</b>	<b>35</b>
Section I: Students as Sources	36
A. Rating Scales	38
B. Written Appraisals	54
C. Interviews	57
D. Tests of Achievement	60
Section II: Colleagues as Sources	63
A. Classroom Observations	64
B. Review of Classroom Materials	68
C. Review of Instructional Development and Advising	69
Section III: Self as a Source	70
Section IV: Alumni as Sources	73
Section V: Records as Sources	75
<b>5. Using Evaluative Information</b>	<b>77</b>
<b>Appendixes</b>	<b>88</b>
<b>References</b>	<b>127</b>
<b>About the Authors</b>	<b>135</b>



## PREFACE AND ACKNOWLEDGMENTS

This book is first and foremost a practical guide intended to assist faculty and administrators critique, design, and implement evaluation of teaching on their campuses.

Although this guidebook is practical in its intent and contents, we have written it from the point of view that the evaluation of teaching should be assessed from a variety of perspectives; that is, no single piece of evidence (e.g., ratings) collected from one source (e.g., students) is sufficient to judge the competence of a teacher. When put into practice, this principle becomes a multiple purpose, criteria, source method approach in this guidebook. A second major principle in evaluating teaching effectiveness is that the purpose of the evaluation, such as personnel decision and improvement, needs to be taken into account when evaluating. Purpose is related to use in our thinking; that is, the use to be made of an evaluation needs to be determined before an evaluation is undertaken.

We consider our approach to evaluation to be within the current mainstream of thinking of faculty evaluation, at least the thinking of our colleagues who write and conduct research in this area. The current consensus is that evaluation is a complex, dynamic undertaking and that sole reliance on student ratings, the most common strategy in evaluation to date, is not sufficient. We also regard our approach to be consistent with the position on evaluation of teaching taken by the American Association of University Professors as noted in the report of Committee C on College and University Teaching, Research, and Publications (1975).



The numerous origins of our thinking about evaluation as presented in this book are based both on our experiences as evaluators in our office (Measurement and Research Division, Office of Instructional Resources at the University of Illinois, Urbana—Champaign) and the writings, ideas, conversations, and debates with many colleagues over the past decade or more. Although we can't mention everyone, we want to acknowledge some of our colleagues and briefly state how they have influenced our thinking. Robert Brown (University of Nebraska—Lincoln) helped form our view of evaluation as being fundamentally a human enterprise in which communication between the parties is a key part of the evaluation process; Barbara Gross Davis (University of California—Berkeley), has demonstrated a thoroughness and informal style of working with her clients worth copying; Ernest House (University of Illinois) helped persuade us that since evaluation is not infallible, we should regard it as a form of argument and treat it as a guide to action. Donald Hoyt (Kansas State University) has greatly helped promote the important role of credibility in evaluation by viewing it as a political and perceptual issue apart from the technical problems of validity and reliability; Martin Maehr (University of Illinois) has helped us see the need to place evaluation within the context of a person's motivational pattern; and Wilbert McKeachie (University of Michigan) has presented three conditions that are necessary for faculty to change their teaching behaviors based on feedback about their teaching, one being that alternative ways of behaving must be presented before a person can change. This idea has influenced us on the role of feedback in facilitating improvement. Barak Rosenshine (University of Illinois) assisted in our conceptualization of a hierarchy of student rating item types, especially the differentiation of high and low inference items. Michael Scriven (University of Western Australia) has distinguished assessing the worth versus the merit of a faculty member, and this distinction is embedded in our thinking on

the uses of information. Richard Smock's (University of Illinois) emphasis on the symbiotic relationship between evaluation and development has made us more aware that these two activities are not totally independent of each other. Robert Stake (University of Illinois), who has tirelessly argued an evaluation should incorporate different value-perspectives, has not only influenced our thinking but many in the field of evaluation.

Many of our colleagues who hold similar positions at other institutions have also played an important part in our work as evaluators. Ken Doyle and John Centra also provided important leadership in the American Educational Research Association Special Interest Group, Instructional Evaluation, a group that has conducted the lion's share of the empirical research summarized in this guidebook.

In addition, a few of our colleagues played a specific part in developing this book. Many of the ideas have been field tested in workshops at the University of Illinois and elsewhere, including one at SUNY at Potsdam under the leadership of C. R. McKinstry and two at the University of Virginia under the leadership of Sam Kellams. Sam's critique of an earlier draft helped us in our ideas on defining good teaching as a prerequisite to the evaluation of teaching. We also received invaluable feedback on earlier versions of this book from our former staff colleague, David Frisbie (University of Iowa) and from Barbara Gross Davis (University of California—Berkeley).

Materials from many others have been included in this book. The following have kindly given us permission to reproduce their materials: Educational Testing Service; Center for Faculty Evaluation and Development, Kansas State University; University of Virginia; University of Southern California; Augustana College, Rock Island, Illinois; and the departments of business administration and horticulture, University of Illinois, Urbana—Champaign.

Finally, we are grateful to Charles McIntyre, Director of the Office of Instructional Resources at the University of Illinois for allowing us the time to write a guide for the University of Illinois, Urbana—Champaign campus use and for encouraging us to write this guidebook for use on other campuses. We also express our appreciation to Janet Osterbur and Debra Drake, who willingly typed yet another revision.

## HOW TO USE THIS GUIDEBOOK

In using this guidebook you should keep the following in mind:

- (1) We have broadly defined teaching, and it thus encompasses components such as classroom activities, organizing a course, developing a curriculum, and advising students.
- (2) We have emphasized the distinction between two major purposes of evaluating faculty—personnel decision and teaching improvement.<sup>1</sup> These purposes are to be viewed as complementary. Conflicts that emerge from evaluating faculty simultaneously for both purposes need to be recognized and dealt with, but if an evaluation is properly designed and implemented, both purposes can be served with a minimal amount of conflict and with increased efficiency and effectiveness.
- (3) “What is the use?” is a fundamental question to ask in any evaluation program of faculty competence. In our view of evaluation, the uses to be made of an evaluation is one of our two overriding principles of faculty evaluation.
- (4) The second key principle in our view of evaluation is multiple perspectives. In this guidebook we have labeled it a “multiple purpose, criteria, source method approach.” The net result of this view is a very comprehensive approach to evaluation, which, however, can seldom be fully implemented on any one campus. Thus we recommend that you adopt this approach as a conceptual framework (i.e., a way of thinking about evaluation) to help you organize your ideas and plans. Then, as you begin to implement this multiple-perspective approach, your “realism” should emerge and be an important factor in deciding what you can do given your local conditions.
- (5) The human side of evaluation is crucial. Evaluation of persons is a deeply personal and sensitive undertaking. We have yet to work with someone who has not been anxious, interested, or

concerned about an assessment of his or her work. But giving advice and suggestions about this side of evaluation is difficult, and thus our concern about the human element in evaluation may not come across as strongly in this guidebook as we wish it to be.

This guidebook is organized into five chapters, with evaluation for personnel decisions and improvement highlighted in Chapters 2 through 5. Chapter 4, the chapter on ways of collecting information about teaching, is organized around the five common sources of information—students, colleagues, self, alumni, and records. For each source we have included a discussion of the technical quality of the evaluative information that can be collected from each source, examples of techniques and instruments, and a list of suggestions for using information from these sources for both personnel decision making and improvement. For a quick overview of the ways that various components of instruction can be evaluated, please read Table 4.1. Table 4.2 lists where in the book the various ways are discussed.

This guidebook was written for three major audiences: (1) departmental and college administrators who have the responsibility of evaluating faculty for annual salary increases and for promotion and tenure, (2) departmental advisory and executive committees, and (3) faculty who desire to collect more and better information about their competence both for personnel decisions and for improving their own teaching.

If you are a college or departmental administrator or member of a committee with the responsibility for evaluating teaching for salary adjustments and/or promotion and tenure, Chapters 1, 3, and 5 and the sections in Chapters 2 and 4 that are headed “Suggestions for Personnel Decision” are the most relevant. If you wish to learn of possible ways to evaluate teaching for improvement, sections in chapters 1 through 4 headed “Suggestions for Improvement” are the most relevant.

If your institution (e.g., system, campus, college, department) has specific policies and practices already in force, you will need to take them into account in using this guidebook. From our experience the best plans are those that are developed locally.

## Note

1. Some have equated evaluation for personnel decisions to summative evaluation and evaluation for improvement to formative evaluation.



## CHAPTER 1

### DEFINING GOOD TEACHING

What is good teaching? How can we define meritorious teaching? Can we measure excellence in teaching? Questions of this type have been asked for centuries and they now embody a key issue in the minds of our politicians and civic leaders as the debate on merit pay for teachers escalates.

Unfortunately, there is no set of easy answers to these questions. Research on teaching effectiveness is voluminous and approached from a number of theoretical perspectives. (See Peterson & Walberg, 1980, for a summary of the trends in this research.) The research on effective teaching at the collegiate level has primarily been in two areas. One focus has been on discovering what teacher characteristics are associated with good teaching. The net result of this line of inquiry has been a list of teacher characteristics (closely aligned with attributes, traits, and personality factors) that are used to define the ideal, model, best effective teacher. Although it is impossible to capture the findings in a phrase, the one that comes as close as any in our opinion is the phrase, "hardness of the head and softness of the heart," which Goldsmid, Gruber, and Wilson (1977) used to summarize how students and faculty colleagues define excellence in teaching. (For more information on this topic, see reviews of research by Aubrecht, 1979, 1981; Costin, Gre-



nough, & Menges, 1971; Dowell & Neal, 1982; Feldman, 1976a,b, 1977, 1978, 1979, 1983; Kulik & McKeachie, 1975; Levinson-Rose & Menges, 1981; and McKeachie, 1979; books by Centra, 1980; Doyle, 1983; and Seldin, 1980; and the chapter by Scriven in the *Handbook of Teacher Evaluation* edited by Millman, 1980.)

The second line of inquiry has focused on the relative effectiveness of the lecture method and alternative modes of instruction. Kulik and Kulik (1980) in their review of the research in this area concluded that teaching by the lecture and discussion method are equally effective if the criterion is learning of factual information. However, teaching by discussion is more effective than lecturing if the criteria are problem-solving abilities, interesting subject matter, attitudes, and curiosity. In their review of research on the individualized approaches to teaching such as the Personalized System of Instruction (PSI), they concluded that the personalized instruction modes generally resulted in higher end-of-course achievement, better long-range retention, but not longer student time spent on courses when compared to the traditional lecture method.

Based on our review of the research on teaching at the collegiate level, we think that the research reflects a diversity of conclusions as much as a consensus. Thus no one definition of excellence in teaching is advocated as the standard against which all teaching is to be compared.<sup>1</sup> As a practical guide to evaluating teaching, we think that a good strategy for defining excellence in teaching initially is to consider three major areas that can be emphasized in defining teaching. They are *input*, *process*, and *product*. Figure 1.1 displays these three areas with some prominent factors in each area. In general, the evaluation of instruction can be divided by its emphasis on input (What do students and teachers bring to the classroom?), process (What do teachers and students do in a course?), or product (What do students learn or accomplish in the course?). A closer look at each emphasis