TEACHER EVALUATION: Guide to Effective Practice

Evaluation in Education and Human Services

Editors:

George F. Madaus, Boston College,
Chestnut Hill, Massachusetts, U.S.A.
Daniel L. Stufflebeam, Western Michigan
University, Kalamazoo, Michigan, U.S.A.

Other books in the series:

Madaus, G. and Stufflebeam, D.:
Education Evaluation: Classic Works of Ralph W. Tyler
Gifford, B.:
Test Policy and Test Performance
Osterlind, S.:
Constructing Test Items
Smith, M.:
Evaluability Assessment
Ayers, J. and Berney, M.:
A Practical Guide to Teacher Education Evaluation
Hambleton, R. and Zaal, J.:
Advances in Educational and Psychological Testing
Gifford, B. and O'Connor, M.:
Changing Assessments
Gifford, B.:
Policy Perspectives on Educational Testing
Basarab, D. and Root, D.:
The Training Evaluation Process
Haney, W.M., Madaus, G.F. and Lyons, R.:
The Fractured Marketplace for Standardized Testing
Wing, L.C. and Gifford, B.:
Policy Issues in Employment Testing
Gable, R.E.:
Instrument Development in the Affective Domain (2nd Edition)
Kremer-Hayon, L.:
Teacher Self-Evaluation
Payne, David A.:
Designing Educational Project and Program Evaluations
Oakland T. and Hambleton, R.:
International Perspectives on Academic Assessment
Nettles, M.T. and Nettles, A.L.:
Equity and Excellence in Educational Testing and Assessment

TEACHER EVALUATION: Guide to Effective Practice

by
Anthony J. Shinkfield
and
Daniel L. Stufflebeam

with contributions by:

Carol Anne Dwyer
Sandra Horn
Madeline Hunter
Edward Iwanicki
Richard Manatt
Thomas McGreal
Robert Mendro
William Sanders
William Webster
Graeme Withers

KLUWER ACADEMIC PUBLISHERS
Boston/Dordrecht/London

Distributors for North America:

Kluwer Academic Publishers
101 Philip Drive
Assinippi Park
Norwell, Massachusetts 02061 USA

Distributors for all other countries:

Kluwer Academic Publishers Group
Distribution Centre
Post Office Box 322
3300 AH Dordrecht, THE NETHERLANDS

Library of Congress Cataloging-in-Publication Data

Shinkfield, Anthony J.

Teacher evaluation : guide to effective practice / by Anthony J.
Shinkfield and Daniel L. Stufflebeam : with contributions by Carol
Anne Dwyer ... [et al.].

p. cm. -- (Evaluation in education and human services)

Includes bibliographical references and indexes.

ISBN-13: 978-0-7923-9674-1

e-ISBN-13: 978-94-009-1796-5

DOI: 10.1007/978-94-009-1796-5

1. Teachers--Rating of. I. Stufflebeam, Daniel L. II. Title.

III. Series.

LB2838.S53 1995

371.1'44--dc20

95-14669

CIP

Copyright © 1995 by Kluwer Academic Publishers

Softcover reprint of the hardcover 1st edition 1995

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, mechanical, photo-copying, recording, or otherwise, without the prior written permission of the publisher, Kluwer Academic Publishers, 101 Philip Drive, Assinippi Park, Norwell, Massachusetts 02061.

Printed on acid-free paper.

Contents

ACKNOWLEDGMENTS	ix
ABOUT THIS BOOK	1
Foreword	1
An Overview of Contents	2
Audiences	6
Contributors	8
 1	
HISTORICAL PERSPECTIVES OF TEACHER EVALUATION	9
Pre-World War II	11
Post-World War II Until The Mid-1970s	13
The Late 1970s To The Present	23
Future Challenges In Evaluation	30
A National Center for Research on Teacher Evaluation and Dissemination of Outcomes	34
The Praxis Series: Professional Assessments for Beginning Teachers	35
The National Board for Professional Teaching Standards	36
Conclusion	37
References	38
 2	
STANDARDS AND CRITERIA FOR TEACHER EVALUATION	43
Preamble: The Place and Importance of Standards	43
References	47

Professional Standards for Assessing and Improving Teacher Evaluation Systems	47
Criteria for Performance-Based Teacher Assessments: Validity, Standards, and Issues, <i>by Carol Anne Dwyer</i>	62

3
SCHOOL PROFESSIONALS' GUIDE TO IMPROVING TEACHER EVALUATION SYSTEMS

Preamble: A GUIDE to Improving Teacher Evaluation Systems	81
Introduction	82
Teacher Evaluation: Its Purpose, Meaning, and Improvement	85
Organizing a Participatory Project to Improve Teacher Evaluation	102
Profiling the Current Teacher Evaluation System	106
Determining which Personnel Evaluation Standards Are Met by the Present Evaluation System	111
Deciding and Planning How to Improve the Teacher Evaluation System	117
Improving the Current Teacher Evaluation System	121
Planning and Implementing the Evaluation Improvement Project	127
References	131
Form for Documenting a Teacher Evaluation System	133
Questions to Be Answered in Addressing the Personnel Evaluation Standards	153

4
MODELS FOR TEACHER EVALUATION

Preamble: Overview of Alternative Models	173
Madeline Hunter: Instructional Effectiveness Through Clinical Supervision	187
Thomas McGreal: Characteristics of Successful Teacher Evaluation	208
Appendix: An Example of an Evaluation System That Reflects the Commonalities of Successful Systems	230
Edward Iwanicki: Contract Plans—A Professional Growth-Oriented Approach to Evaluating Teacher Performance	245
Getting Value from Teacher Self-Evaluation, <i>by Graeme Withers</i>	261
Richard Manatt: Teacher Performance Evaluation	271
Toledo School District: Intern and Intervention Programs	289
Principal and Peer Evaluation of Teachers for Professional Development, <i>by Anthony Shinkfield</i>	302
The National Board for Professional Teaching Standards: Assessing Accomplished Teaching	319

CONTENTS

vii

The Tennessee Value-Added Assessment System (TVAAS): Mixed Model Methodology in Educational Assessment, by <i>William L. Sanders and Sandra P. Horn</i>	337
An Accountability System Featuring Both "Value-Added" and Product Measures of Schooling, by <i>William Webster and Robert Mendro</i>	350
5	
AN ANALYSIS OF ALTERNATE MODELS	377
Introduction	377
A Summary of the Purposes of the Ten Models Selected in Chapter 4	378
An Examination of the Models Against the Joint Committee's <i>Standards</i>	385
The Worth of the Presented Models for Decision Making	392
References	394

Acknowledgments

This work was supported by the U.S. Department of Education, Office of Educational Research and Improvement (OERI) through grant No. R117Q00047 to Western Michigan University for the work of the National Center entitled Center for Educational Accountability and Teacher Evaluation (CREATE). The opinions expressed are those of the authors, and do not necessarily reflect those of the U.S. Department of Education.

ABOUT THIS BOOK

Foreword

Formal teacher evaluation has long been considered important by the public. However, until recently most schools evaluated teachers in only the most cursory manner, e.g., by the principal's annual brief observation of the teacher's classroom performance. The last 15 years has seen a dramatic development of the technology of teacher evaluation. Also, there has been much more attention in the schools to making teacher evaluation extensive, systematic, and valid. Much of this growth arose from state mandates and from educational institutions whose leaders seek improved means of evaluating teachers. Closely connected to these worthwhile enterprises has been the often expressed need to elevate the respect accorded to teacher evaluation by using professional standards. This book is an attempt to present and examine important developments in teacher evaluation, show their interrelationships, and offer practical guidelines for using teacher evaluation models.

We organized the book around 4 dominant, interrelated core issues: professional standards, a **GUIDE** for applying the Joint Committee's *Standards* (which are featured in Chapter 3), 10 alternative models for the evaluation of teacher performance, and an analysis of these selected models. The book draws heavily upon the research and development conducted by the federally funded national Center for Research on Educational Accountability and Teacher Evaluation (CREATE).

We hope that the reader will grasp the essence of the experience of sound teacher evaluation and apply its principles, facts, ideas, processes, and procedures. To this end, we offer information that is as up-to-date, and practically useful, as possible. For example, Chapter 3 presents a thoroughgoing, logical sequence of steps to guide readers in examining teacher evaluation systems for adherence to definitive pro-

fessional standards. Moreover, the book invites and assists school professionals and other readers to examine the latest developments in teacher evaluation.

Evaluation, both formal and informal, is inextricably interwoven with the entire process of education. Since the quality of learning depends largely on the quality of teaching, teacher evaluation clearly is essential in effective schools. Administrators, other leaders in education, and school communities have long realized the often disturbing truth of this situation. Some schools and districts have failed to plan and carry out an appropriate teacher evaluation scheme because they could not locate a suitable model. Others failed because they lacked the knowledge or will to make a sound model work in the face of difficulties. By addressing these kinds of issues, this book should guide readers toward attaining important goals in teacher evaluation. These include improving the performance of teachers, students, and the organization as a whole.

An Overview of Contents

The book has five chapters. Together they explain and amplify the four main cores already listed.

Chapter 1 offers a historical perspective of teacher evaluation. Our literature search brought us to the realization that extensive material on this topic does not exist (and certainly not in any consolidated form). Formal teacher evaluation is largely a recent phenomenon. We hope, therefore, that this section is a useful, interesting contribution to the emerging field of assessment of teacher performance.

Chapter 1 first examines pre-World War II ideas of teacher evaluation. The common view was that students were responsible for their learning. Accordingly, schools attributed learning deficits to the student rather than the teacher.

However, during the Victorian era, England initiated the first coordinated, nationwide program to assess teachers and reward them accordingly. This was called payment by results. This led to the English inspectorial system that has persisted in that country, and others. Many annual reports and other sources reveal that, as one would expect, principals were conducting informal assessments of teachers in the U.S. during the 1900s.

From the conclusion of World War II until the mid-1970s, the literature and research reports reveal a growing consensus in the major purposes of teacher evaluation. The predominant factor, it seems, was the growing belief that the entire educational system must gain from improved teacher performance arising from widely acceptable evaluation processes. During this time states and school districts made tentative beginnings in areas like systematic accountability of teachers and appraising teachers based on the learning of their students.

The final section of Chapter 1 treats the late 1970s to the present. It looks into the emerging ideas of evaluation and accountability, teacher certification, and the legal and political aspects of evaluation.

Chapter 1 concludes with a view of future challenges in evaluation. Formal teacher evaluation may be in its infancy, but the growth rate has been most marked, particularly during the past decade. Recent historical perspectives necessarily reference contrasting models for evaluating teachers; the importance of using professional standards to assess and improve teacher evaluation systems; CRE-ATE's extensive, ongoing, nationwide studies of teacher evaluation; and the emergence of an annual national teacher evaluation institute. Thus, in the 1980s and 1990s there has been a considerable advance in the theory and practice of this important aspect of education.

Chapter 2 centers on standards and criteria for teacher evaluation. If the evaluation field is to achieve its potential contribution in any area, it must yield dependable assessments of all aspects of a discipline or system. It follows that evaluators must develop and apply professional standards to help ensure that all aspects of evaluation attain the highest levels of fairness and quality.

Chapter 2 also discusses **Professional Standards for Assessing and Improving Teacher Evaluation Systems**. It outlines the major undertaking by a widely-representative national Joint Committee during the 1980s culminating in the publication of *The Personnel Evaluation Standards* in 1988. These *Standards* provided education with a widely endorsed set of guiding principles. The rigorous application of the *Standards* strengthens and adds credibility to systems and practices of personnel evaluation and protects teachers and others from corrupt evaluation practices. They should also help to mitigate evaluation-related conflicts among different interest groups. The chapter discusses development of the Joint Committee's *Standards*. It delineates areas where they are applied. Then it provides a brief introductory statement about each of the four basic principles of sound evaluation—propriety, utility, feasibility, and accuracy.

Professional standards are the foundation for this book. They play a remarkably pervasive role, directly and indirectly in every chapter. We believe that school professionals take a serious view of teacher evaluation only if they are knowledgeable about the Joint Committee's *Standards*. Chapter 3, which is a user's guide to assessing and improving teacher evaluation systems, and the Preamble to Chapter 4, an overview of alternative models, exemplify this point most markedly.

In **Chapter 2** Carol Anne Dwyer writes from her recent experiences in leading the development of Praxis (the Educational Testing Service successor to NTE). Her concern is the historic lack of standards for assessing teacher competence. Thus, her article's emphasis on criteria for teacher evaluation relates strongly to the emphasis placed on professional standards for evaluations in Chapter 2. She observes that PRAXIS, inter alia, keys to 3 different aspects of the pedagogy:

content-specific pedagogical knowledge, knowledge of general principles of teaching and learning that transcends different disciplines, and the application of this knowledge and skills to actual classrooms. These require different assessment methods. Dwyer also summarizes the recent progress in improving performance assessment. She documents progress in closing the gap between standards for sound performance assessment and actual practice.

This article is on the leading edge in both defining and addressing issues in the validation of teacher assessments. It also emphasizes the complexity and considerable expense involved in defining and applying assessment standards.

Chapter 3 is the School Professionals' GUIDE to Improving Teacher Performance Evaluation Systems. In many ways, this chapter is a companion document to Chapter 2, particularly in respect to the Joint Committee's *Standards*. This part shows how the *Standards* can be effectively and systematically used to examine extant or contemplated teacher evaluation systems.

Chapter 3 begins with a discussion of which particular standards are most important in assessing the adequacy of a teacher evaluation system. It next presents a conceptual framework delineating factors that define and influence such systems. Advice is offered on how to use this framework to examine systems. The chapter concludes with recommendations on how to organize a participatory project involving all stakeholder groups.

These activities in the GUIDE lead to formulating a new system for teacher evaluation or improving an existing one. The GUIDE refers to the *Standards* in discussing these processes. The GUIDE concludes with 3 appendices. These offer highly relevant exercises plus a table of contents for a manual on teacher evaluation.

We hope that the GUIDE will prove very beneficial to school districts. Its authors designed it to be easy to understand, easy to use, and to apply to evaluation needs in schools. It recommends and provides direction for involving all stakeholders in improving a district's teacher evaluation system.

Chapter 4, Models for Teacher Evaluation, has two main objectives. The first is to give a general overview of teacher evaluation models. The second is to present 10 models that are widely used or especially interesting.

In the preamble to Chapter 4, we explain our intended meaning of the term model. In effect, we redirect emphasis from various ideas of teacher evaluation that portray the actual process of evaluation to those that prescribe a preferred, ordered set of steps for conducting teacher evaluations. The presented models vary from being highly directive to less directive. However, they address a common purpose. It is to evaluate teachers so well that there are clear and practical benefits for schools, school districts, and students.

We based our overview of teacher evaluation models on a listing provided by Michael Scriven in his CREATE publication called TEMP Memo 2 (September 1991). We added self-evaluation to the list and briefly outlined each model. The 15

models contained in the overview are Traditional Impressionistic, Clinical Supervision, Research-Based Checklist, High Inference Judgments, Interviewing, Paper and Pencil Tests, Management by Objectives, Job Analysis, Duties-Based Approach, Theory-Based Approach, Student (Learning) Improvement Outcomes, Consumer Ratings, Peer Ratings, Self-Evaluation, and Metaevaluation of Existing Models.

The 10 selected models that comprise the remainder of Chapter 4 cover most of the listed TEMP Memo approaches. We believe that they show the range of models available for school professionals to consider.

We define *formative* evaluation as systematically assessing the merit and worth of some enterprise to guide its continuing revision during the process. We define *summative* evaluation as a comprehensive assessment of the merit and worth of the enterprise, including especially its outcomes at the end of the process. The remainder of Chapter 4 depicts 10 evaluation models divided into (1) formative, (2) formative and summative, and (3) summative approaches to teacher evaluation.

The first four models, Chapters 4.1–4.4, are essentially formative: Madeline Hunter's *Instructional Effectiveness Through Clinical Supervision*, Thomas McGreal's *Characteristics of Successful Teacher Evaluation*, Edward Iwanicki's *Contract Plans - A Professional Growth-Oriented Approach to Evaluating Teacher Performance*, and Graeme Withers' *Getting Value from Teacher Self-Evaluation*.

The next three models, Chapters 4.5–4.7, are both formative and summative. They include Richard Manatt's *Teacher Performance Evaluation*, Toledo School District's *Intern and Intervention Programs*, and Anthony Shinkfield's *Principal and Peer Evaluation of Teachers for Professional Development*.

The final three models, Chapters 4.8–4.10, are summative. They include the National Board for Professional Teaching Standards' work on *Assessing Accomplished Teaching*, William Sanders and Sandra Horn's *The Tennessee Value-Added Assessment System (TVAAS)—Mixed Model Methodology in Educational Assessment*, and William Webster and Robert Mendro's *An Accountability System for School Improvement*.

Chapter 5 analyzes the 10 models presented in Chapter 4. It tries to achieve the following:

- A summary display of the purposes of each model, which is placed under one of three headings: formative, formative and summative, and summative.
- An examination of the models against the Joint Committee's *Standards* to find their main strengths and weaknesses
- Our discretionary value judgment about the worth of these models considered against the main uses of teacher evaluation models for decision making.

Taken together, these elements comprise a useful consumer's guide. It stems logically from the development of the first three of this book's four main cores: standards for teacher evaluation, the GUIDE to improving teacher evaluation

systems by applying the Joint Committee's *Standards*, and the presentation of alternative models for teacher evaluation.

Audiences

In the improvement of teacher evaluation, it is critical to organize collaborative work by those involved in the evaluation process and those influenced by its outcomes. Almost without exception the approaches to teacher evaluation given in this book bear this out. Most of the approaches either state or imply the need for collaboration in their models and guidelines for their use.

Thus, we wrote the book purposely to provide useful guidance to professional personnel connected with schools—particularly teachers, principals and other school administrators, and superintendents and board members. We recommend that officials in state education departments and professional education organizations use this book to develop better state and district systems of teacher evaluation. Parents, and the public generally, are also stakeholders in the products of the educational process. Many of them should find the book useful for assessing their school districts' teacher evaluation practices. Concerted actions by all the stakeholders to improve teacher evaluation should lead to improved public credibility of the schools and the education establishment.

Teacher evaluation is growing in importance both quickly and pervasively. We therefore recommend the book to those whose working lives could see them involved with teaching, or evaluation, or both. These include education professors and college and university students at both undergraduate and graduate levels.

There is a strong reason for research to continue in teacher evaluation because it is an emerging area of national interest and need. Any group of professionals using the GUIDE presented in Chapter 3, where we assess models in the light of accepted standards, will discover the imperfections of any model. The outcomes of sound, focused, research studies should help improve many aspects of teacher evaluation practice. These include validity, potential utility, feasibility, consonance with human rights legislation, and the effective and efficient application of validated models.

Many people, therefore, have a legitimate interest in teacher evaluation. These include:

- Those who are immediately involved in teacher evaluation—to lists already given, students are added because the basic purpose of any program within education is to improve student learning.
- Those who provide leadership for improving teacher evaluation practices at the state and national levels

- Parents and the wider public whose educational accountability requirements should be met
- Those persons responsible for teacher training and student teachers themselves
- Evaluation researchers

Considered as a whole, this is a widely disparate group, often holding different views about education and its many complex functions. Therefore, we wrote the book straightforwardly in the hope that it will be readable, interesting, and practicable.

Responsibilities and Reassurances

A discussion of teacher evaluation often leads to the question of what other persons should be evaluated. That is, who else should communities hold accountable for producing quality outcomes from educational processes? There is a growing recognition that the acceptance of a teacher evaluation system is strengthened if all administrators in the district (superintendent, principals, vice-principals, and perhaps school board members could be added) are also evaluated, and particularly so if common principles and procedures are used. Someone must evaluate the evaluators. Otherwise, the public and other stakeholders can have no assurance that the evaluators are employing sufficient accountability measures to safeguard student interests and motivate educational improvements. There is also growing realization that school districts must evaluate support and special staff and that the process should not differ significantly from that for teachers and administrators. This country is a considerable distance from this happening universally, and even further from it occurring in highly credible and acceptable ways.

We did not address these important adjunct areas for personnel evaluation in this book. However, CREATE is developing methodology designed to give very considerable direction to evaluation of administrators, support personnel, and board members.

The major stakeholders should share responsibility for developing teacher evaluation in a school or district. The major actors include the teachers, administrators, board members, and teachers' organization. However, responsibility for carrying out the process satisfactorily rests with different groups, or individuals, according to the decision-making involved. For instance, administrators will clearly strengthen a district and its schools by doing a thoroughly professional job of assessing and selecting newcomers. Also, principals are largely responsible for using best possible formative evaluation for staff professional development. Unfortunately, evaluations associated with tenure, promotion, and reassignment are often haphazard and often lack the credibility needed to guide and defend personnel

decisions. The board must enact policies to assure that the district professionally evaluates teachers and other district personnel to benefit the district and its students.

As professional people, teachers themselves must engage in evaluation for both professional development and accountability. Regular external evaluation acts as a supplement to self-evaluation, an ongoing process essential for any professional. Teacher involvement in developing evaluation models is important. The school teacher, school administrator, and school board member can assure themselves and their clients that accountability is being practiced properly only when they conduct, report, and act on the results of systematic, valid evaluation.

Contributors

We owe a great debt to those who have made this book possible. We refer to those who have allowed us to present a version of their model for teacher evaluation: the late Madeline Hunter, Richard Manatt, Thomas McGreal, and Edward Iwanicki. We hope that we have done justice to the Toledo School District and the National Board for Professional Teaching Standards in our discussion of their models. Thanks to Carol Anne Dwyer, Graeme Withers, William Sanders and Sandra Horn, and William Webster and Robert Mendro for allowing us to reproduce their articles; to Thomas McGreal for permission to print in full the appendix to his book, *Successful Teacher Evaluation*; and to Daniel Stufflebeam's coauthors David Nevo, Bernard McKenna, and Rebecca Thomas for agreeing to our inclusion of a version of the *School Professionals' GUIDE*.

We especially appreciate the support given to us by the U.S. Department of Education, Office of Educational Research and Improvement (OERI) through a grant to Western Michigan University for the work of CREATE. The opinions expressed in this book do not necessarily reflect those of the U.S. Department of Education. This book helps to fulfill one of CREATE's chief missions, which is to extend knowledge of the place and importance of teacher evaluation and to offer guidance about the use of models. Finally, we very much appreciate the ready and most able assistance given us by the book's reviewers. CREATE secures external evaluation of all CREATE products by appropriately qualified people. Graeme Withers, Senior Research Officer with the Australian Council for Educational Research, and Bernard McKenna, retired official of the National Education Association, fulfilled this task in a thoroughly professional manner. We appreciate their significant contributions.

— Anthony J. Shinkfield

— Daniel L. Stufflebeam