The purpose of the conference, "Wagging the Dog, Carting the Horse: Testing vs. Improving California Schools," was to discuss alternative perspectives on testing and evaluation in education and their role in improving teaching and learning. Four papers were presented: (1) "Using Educational Evaluation for the Improvement of California Schools," by Elliot Eisner; (2) "Evaluating Educational Quality: A Rational Design," by Eva L. Baker; (3) "The Influence of Testing on Teaching and Learning," by Norman Frederiksen; and (4) "Beyond Outcome Measures: An Agenda for School Improvement," by John Goodlad. These papers, and the small group discussions at the conference, are summarized in this paper. The conference participants are listed, and the conference program is appended. (BW)
DELIVERABLE - NOVEMBER 1984
RESEARCH INTO PRACTICE PROJECT

Summary of Conference Proceedings
"Wagging the Dog, Carting the Horse: Testing and Improving Schools"

Joan Herman
Project Director

Grant Number
NIE-G-84-0112, P4

Center for the Study of Evaluation
UCLA Graduate School of Education
The project presented or reported herein was supported in part pursuant to a grant from the National Institute of Education, Department of Education. However, the opinions expressed herein do not necessarily reflect the position or policy of the National Institute of Education and no official endorsement by the National Institute of Education should be inferred.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>2</td>
</tr>
<tr>
<td><strong>Summary of Presentations</strong></td>
<td></td>
</tr>
<tr>
<td>Using Educational Evaluation for the Improvement</td>
<td>4</td>
</tr>
<tr>
<td>by Elliot Eisner</td>
<td></td>
</tr>
<tr>
<td>Evaluating Educational Quality: A Rational Design</td>
<td>6</td>
</tr>
<tr>
<td>by Eva L. Baker</td>
<td></td>
</tr>
<tr>
<td>The Influence of Testing on Teaching and Learning</td>
<td>8</td>
</tr>
<tr>
<td>by Norman Frederiksen</td>
<td></td>
</tr>
<tr>
<td>Beyond Outcome Measures: An Agenda for School Improvement</td>
<td>10</td>
</tr>
<tr>
<td>by John Goodlad</td>
<td></td>
</tr>
<tr>
<td><strong>Summary of Small Group Discussions</strong></td>
<td>12</td>
</tr>
<tr>
<td><strong>Conference Participants</strong></td>
<td>14</td>
</tr>
<tr>
<td><strong>Appendix A: Conference Program</strong></td>
<td>18</td>
</tr>
</tbody>
</table>
Introduction

This document summarizes proceedings of "Wagging the Dog, Carting the Horse: Testing vs. Improving California Schools," a conference sponsored jointly by the UCLA Center for the Study of Evaluation and the UCLA Laboratory in School and Community Education. The one day conference was held on June 7, 1984 at the Sheraton Miramar Hotel in Santa Monica, California.

The purpose of the conference was to discuss alternative perspectives on testing and evaluation in education and their role in improving teaching and learning. The conference considered whether, in the current rush to issue "report cards" on our schools, there may be danger that we are losing sight of what is meant by a "quality education." More particularly, speakers posed -- and provided alternative answers to -- the questions of what should be assessed and how evaluation can best contribute to our understanding of schools and to their improvement.

The conference attracted a diverse audience of professional educators, school board members, educational researchers and policy-makers. It featured presentations by Professor Eva Baker, Directors of the UCLA Center for the Study of Evaluation; Professor Elliot Eisner, Stanford University; Dr. Norman Frederickson, Educational Testing Service; and Professor John I. Goodlad, former Dean of the UCLA Graduate School of Education. Presentations were followed by questions and answers and small group discussions. (A copy of the program is provided in the appendix.) A summary of the presentations and small group discussions is provided in the following pages, followed by a listing of conference participants. The
full text from the presentations can be found in "1984 Policy Studies: The use of testing and evaluation for assessing educational quality and improving school practice."
Measurement, evaluation, and testing should be viewed as three independent processes. Measurement is an arbitrary means of quantifying and describing something without making value judgments about its quality. Evaluation involves making value judgments about something on the basis of some relevant criteria. Testing is a way of getting information by eliciting a response to something.

However, while we can measure without evaluating, can evaluate without measuring and testing, and can test without measuring, the three are often confounded. For example, when newspaper headlines signal decline in test scores, these interpreted test scores, though they provide no information about school context and practice, have an effect on the setting of educational priorities and the educational climate of schools. One has only to spend time in a classroom to see the effect of tests and test scores on teachers' practice.

What we need to do, instead of over-reliance on test scores, is to conduct evaluations that examine classrooms in the context of schooling, evaluations which would thus have the potential to improve the quality of schooling. What we need to do is to subject educational planning, curriculum development, and instructional content and strategies to evaluation, and in order to make effective uses of evaluation, we need to focus on the classroom unit.

Unfortunately, we have created a situation that makes it difficult for evaluators to spend time in classrooms and equally difficult for teachers...
to get feedback on how well they are running their classes. We have
created a situation in which it is sometimes impossible for teachers to
discover ways to become better at their jobs. Over-reliance on tests as
"evaluative" information has contributed to this problem.

To help overcome this situation, we need to begin stressing
evaluation's educational role, to use evaluation to inform teachers about
significant educational practice. To be able to do this, we must provide
teachers with access to each other and to establish a climate of trust so
that teachers will be willing to accept the observations of their
colleagues. Such observation should be designed to describe the subtle but
significant events that take place in a school and to provide feedback to
teachers that they can use to modify their classroom practice. Of
necessity, this description must move beyond the purely quantitative
information provided in test scores.

A prime ingredient in the process outlined above is to reconceive our
notions of inservice. We need to view inservice as a means of stressing
professionalism, as a vehicle which offers teachers the critical support of
their colleagues, as a way of stimulating teachers to become connoisseurs
of educational practice. Such a climate will do much more to improve
education than current attempts to humiliate teachers into excellence by
publishing their students' test scores.

Instead of trying to bully schools into quality education, we need to
give teachers a stake in what they teach, we need to have diverse programs
which use multiple criteria, and we need to create a climate which fosters
and encourages teachers' professional growth.
Evaluating Educational Quality: A Rational Design

Eva L. Baker

Evaluation reflects the viewpoint that we can influence the course of educational events by planning, implementing, and assessing. However, evaluation does not work this way very often.

Though many models of evaluation have been proposed -- criterion-referenced, norm-referenced, goal-oriented, responsive, and so forth -- the one that we need -- effective -- has eluded us for a variety of reasons.

For evaluation to be used it must be usable, meaning that it should reach people who can act on it, it should reach them in a timely manner, and it should be valid and credible. Above all, to be used for school improvement, evaluation must be aimed at the principal unit of change -- the school.

Most evaluations, unfortunately, are driven by a different reality. They are mandated from above, usually to meet the legitimate questions of school boards and government agencies about the effectiveness of education. These questions deal with educational processes such as quality of services as well as questions about what and how well students learn.

But while evaluation needs to generate information that will contribute to responsible oversight of the educational system, it needs also to provide information useful at the point of change, the local school. And therein lies the dilemma: the mismatch between top-down, externally mandated evaluation requirement and bottom-up, locally responsive efforts.

A system for accountability and oversight is driven from the top down and demands comparability of assessment in areas requiring policy
decisions. Point of change evaluation is driven from the bottom up and emphasizes the uniqueness of each school and its staff, setting, students, and social context.

Top-down evaluations usually rely upon commercial achievement tests to generate the information they need for comparison purposes. Bottom-up evaluations require more finely-grained information about student performance. While the two systems make overlapping demands, they also differ tremendously. Those within the schools often find little use in the information provided solely for top-down, policy needs. It is possible, however, to reconcile and merge the two viewpoints efficiently so that policy needs are met while maintaining the personality, integrity, and idiosyncracy of individual schools.

Such a system, embracing both top-down and bottom-up needs, would allow for cross-student and school comparison. However, it would also allow for local option, quick turn-around outcomes measured across time, with possible multiple data sources. It would also address quality of school life, quality of effort, instructional resources, and include measures of process/outcome, affect, and overall context. It would provide a comprehensive evaluation system that could help direct a school improvement agenda.
The Influence of Testing on Teaching and Learning

Norman Frederiksens

There is little question that tests influence what is taught and what is learned. Students, for example, adopt different study methods for different test formats. If a multiple choice test is expected, they will try to learn factual material. If an essay test is expected, they will be more inclined to look for broader concepts and their relationships.

This kind of test influence would not be bothersome if students were exposed to a variety of test formats. But it seems that the numbers of multiple choice tests given to students each year has grown enormously. For example, because it is easier to write multiple choice items that measure factual knowledge, item writers tend not to write items measuring skills in analysis, problem solving, and application. Further, due to increased pressure to teach minimum competency skills, there is less effort to teach important skills that are difficult to measure with multiple choice tests.

Certain trends seem to be emerging from these practices. There is research showing that while performance on test items measuring the basic skills has not declined, performance on items tapping the more complex cognitive skills has. It seems clear that we need tests, then, which measure not only the basic skills but also the ability to process information rapidly and accurately, to apply principles in new situations, and to solve problems not previously encountered.

There are various alternatives to multiple tests. The essay test is one such possibility. And although essay tests are sometimes criticized for their scoring time and low reliability, a variety of procedures exist
both for decreasing the time required for scoring and for increasing rater reliability.

Other testing alternatives, many of which are quite different from conventional tests, have grown out of theories of cognition. One such idea is concerned with measuring speed in performing cognitive tasks. Further, it is possible to combine both speed and power (the more conventional approach) in a test.

Similarly, it is possible to devise tests which tap both short- and long-term memory, and there are various approaches to assessing the processes a student brings to bear in representing a given problem. Different scoring procedures exist for each of these alternatives.

An important feature of the alternatives outlined above is that they represent tasks as well as constituting tests. Greater consideration needs to be given to task assignments such as writing papers, solving assignments, and taking tests. If we begin to view tests as tasks, we will be helping students to acquire not only the knowledge base but also the information-processing skills that are necessary to developing high levels of proficiency in thinking.
Beyond Outcome Measures:  
An Agenda for School Improvement  
John Goodlad

The current furor over school reform needs to be placed in perspective. The decline in competence in schooling and the increasing disaffection in schooling that occurred in the 1970s is closely linked to declining faith in our institutions in general and to economic downturns in the same period.

Publication of *A Nation at Risk* led to a galvanic connection of achievement test scores with school health. That is just as mediocrity in the schools was seen as reflected in declining achievement scores, so was improved school health to be seen in increased achievement scores. But if the schools are in the poor condition that many suggest, it is going to take a long period of care to bring them back to a condition of health. Further, achievement test scores will continue to be a poor indicator for judging that health.

How do we arrive at a more accurate diagnosis of the health of schooling, one that we can then use to prescribe a remedy? First, we need to view the healthy school as one which assures comprehensive democratic access to the domains of knowledge that constitute a good general education. Second, we need an overall evaluative system commensurate with these expectations. That system, essentially, will place much greater emphasis on the context of schooling, on the conditions of schooling which promote or impede healthy growth. We need to examine, for example, satisfaction, school climate, classroom climate, principal-teacher relationships, school-community climate, and the like. Features such as these are not tapped by achievement tests.
Having assessed critical features of the schools environment, we can then bring to bear the value system of the professionals in the school to set an agenda for improvement. And instructional assessment plays an important role in monitoring progress in the agenda for improvement.

However, such assessment will bear little resemblance to the typical achievement test. Rather, it will be concerned with the provision and assessment of an array of learning experiences commensurate with our expectations of a healthy school. Such an assessment system would engage children in solving real problems. It might have children working on problems for which there is no reward. It would engage in modes of inquiry commensurate with what we think real learning is. It would provide them with good guidance and feedback.

This kind of evaluation system will allow us to take the longitudinal view required to gradually bring our schools back to health.
Summary of Small Group Discussions

The conference provided participants with the opportunity to come together in small groups to discuss the implications of the conference for testing policies and school renewal. Small groups were constituted to enable interaction between researchers, practitioners and policy makers.

The conclusions of the small group discussions were as follows:

Group 1:

1. There is a fundamental conflict between the ideal and the real. We know that standardized test scores are not adequate or valid indicators of schools, but given the political realities, we must pursue them.

2. We need to build better coalitions to influence the political process. We need to organize researchers and practitioners so that they can influence program mandates and help enact programs which actually facilitate real change and renewal.

3. We need to find better ways to communicate our progress to our school constituencies. We need to educate them about the limits of standardized tests and share with them a broader and more comprehensive picture of school progress.

Group 2:

1. We need to assume the challenge of reeducating our staffs, community, and districts to issues of renewal. We need to provide time for dialogue between and among units to set goals and plan solution strategies.

2. We need to concentrate on developing professional leadership that can refocus and broaden educational reform. We need to build trust, support risk-taking and experimentation, provide incentives and rewards, and harness the time and resources needed for renewal.

3. We need to develop alternative assessment measures to help us analyze and improve our progress. We need measures that are sensitive to what teachers are trying to accomplish, that can serve as formative checkpoints, and that tap higher level critical thinking and problem-solving skills.
Group 3:

1. We need a strong, confident offensive for change. Rather than being defensive, we need to be proactive in setting the agenda for schools.

2. We need to provide quality time to promote dialogue among teachers, principals, districts and their communities to assess their needs and goals and to plan for improvement. We particularly need to bring teachers back into the dialogue.

3. We need a strong, professional teaching staff in order to promote renewal, we need to find the time, resources and incentives to "reprofessionalize" them and to facilitate their continued growth and satisfaction.

Group 4:

1. We need to face the realities of current testing practices. We must do well to satisfy our public.

2. We need to develop a variety of ways to assess school programs. We need measures of higher level thinking skills, attitudes, and other indicators of school climate and process.

3. We need to reeducate the public, and the media in particular as to what are accomplishing.
## Conference Participants

**WAGGING THE DOG, CARTING THE HORSE:**
**TESTING VS. IMPROVING CALIFORNIA SCHOOLS**

**June 7, 1984**

### NAME

- Delores Allen
- Christine Amato
- Eleanor Anderson
- Carol Anglan
- John Avila
- Jim Bailey
- Eva Baker
- Pam Bailis
- Adrianne Bank
- Carolyn Banks
- Amy Baumann
- John Bay
- Marge Beckman
- Iris Berke
- Roque Berlanga
- Judy Beuerman
- Tom Bishop
- James Black
- Ruth Bloom
- Lane Bove
- Delbert Bredy
- David Brown
- Douglas Brown
- Warren Bryld
- James Burry
- Leigh Burstein
- Margaret Butt
- Maggie Carrillo
- James Catterall
- Norman Chaffin
- Jean Cohen
- Dean Conklin
- Margaret Connel
- Betty Coogan
- Jim Cox
- Linda Daniels
- William Davis
- Greg Demuth
- Jan Dennis-Rounds
- Bill DeSimone
- Susan Dollar
- Martha Dong
- Elliot Eisner
- Donald Empey
- Nancy Enell
- Sydney Farivar

### DISTRICT

- Moorpark USD
- Vallecito School District
- Orange County
- Santa Clara USD
- Ceres USD
- L.A. County
- UCLA
- UCLA
- UCLA
- Fillmore USD
- Hacienda-La Puente USD
- Ventura USD
- Norwalk-La Mirada USD
- Santa Clara County
- ABC USD
- Merced City Schools
- L.A. County
- Burlingame School District
- L.A. County
- Loyola Marymount
- Rialto USD
- Walnut Valley USD
- L.A. County
- Redlands USD
- UCLA
- UCLA
- UCLA
- Lamont
- Anaheim Union H.S. District
- UCLA
- Elsinore Union H.S. District
- League of Women's Voters
- Charter Oak USD
- Santa Barbara City School Dist.
- Norwalk-LaMirada USD
- Anaheim Union H.S. Dist.
- Del Mar Union
- Grossmont Union H.S. Dist.
- Tulare USD
- Cerritos College
- Merced City Schools
- Fillmore USD
- Bassett USD
- Stanford University
- Glendale USD
- Sacramento City Schools
- UCLA
Attendees - 6/7/84 Conference

Jane Fauero
Gerald Figgins
Tom Fitzgerald
Norman Frederiksen
Helen Fried
Pat Ganther
Jane Gawronske
Blossom Getty
Larry Gilham
Bruce Givner
Claude Goldenberg
Stephen Goldstone
John Goodlad
Marsha Graham
Stu Greenfield
Julie Hadden
Michael Halloran
Lee Hancock
Keith Hartwig
Bruce Hauger
Paul Heckman
Claire Hobbs
Patricia Horkan
Ginger House
Leonard Hughart
Kathleen Huburd
Emma Hulett
Stephen Isaac
Tom Jacobson
Vera Jashni
Nancy Jenkins
Bob Jones
Michael Karpman
Robert King
Gerry Klor
Charles Kondrit
Al Koshiyama
Jack Kriege
Carol Kurtz
David Kuzmich
Alex Law
Deolores Lena
Kent Lewis
Tod Likins
Gary Luke
Peggy Lyons
Rudy Macioge
Dorothy Maloney
Tom Martin
John Martofse
Leo Martucci
Jack McLaughlin

Glendora USD
Redlands USD
Charter Oak USD
Ed. Testing Service
ABC USD
UCLA
Walnut Valley USD
Novato USD
Exeter Public Schools
Irvine USD
UCLA
Albany USD
UCLA
ABC USD
State Dept. of Ed.
L.A. City Schools
San Ramon Valley USD
L.A. City College
Private Consultant
Anaheim Union H.S. Dist.
UCLA
Freemont Union H.S. Dist.
Del Mar Union
Norwalk-LaMirada USD
Tulare City School Dist.
Freemont Union H.S. Dist
Norwalk-LaMirada USD
San Diego City Schools
Grossmont Union H.S. Dist.
Culver City USD
Norwalk-LaMirada USD
Norwalk-LaMirada USD
Redlands USE
Grossmont Union H.S. Dist.
Petaluma School Dist.
Rialto USD
State Dept. of Education
Grossmont Union H.S. Dist.
Tom Hayden's Office
Anaheim Union H.S. Dist.
State Dept. of Education
ABC USD
UCLA
Los Gatos Joint Union H.S. Dist.
Napa Valley USD
Santa Monica/Malibu USD
Fontana Unified
Conejo Valley USD
ABC USD
L.A. County
Santa Monica/Malibu USD
Sunnyvale
Attendees - 6/7/84 Conference

Judith Miller
Shirley Mills
Thelma Moore
T. C. Moshier
Bobbi Mulholland
Millie Murray
Gerald Nagata
Dan Nasman
Bruce Newlin
Warren Newman
Frank Nicassio
Wilhelmine Nielsen
Jeannie Oakes
Phillip Oakes
James Oberg
Kenneth Ogden
Jane L'Loughlin
Gary Olson
Jack Owens
Joni Padgett
Robert Parker
Jackson Partin
Barbara Patterson
Bert Pearlman
Terry Pearson
John Perry
Carole Pierce
Marie Plakos
Richard Quaglino
Tom Rabone
Mary Ann Rapuano
Richard Renheim
Jerry Richardson
Patrick Rooney
Thomas Ross
Julia Rothrock
Ray Rudkin
Vicie Rush
June Sale
Glenna Scheer
Neil Schmidt
Jodi Servatius
Norman Seward
Don Shalvey
Keith Shattuck
Frank Shultz
Ken Sirotkik
Rodney Skager
James Smith
Darryl Stucker
Lloyd Swanson
Bonnie Thiele

Private Consultant
San Gabriel School Dist.
Anaheim Union H.S. Dist.
Kings River USD
Irvine USD
UCLA
Kings River USD
San Diego County
Norwalk-LaMirada USD
So. Pasadena USD
Fillmore USD
Escondido Union School Dist.
UCLA
San Juan USD
Alum Road School Dist.(?)
San Marino USD
Culver City USD
Vista USD
Milpitas USD
Fillmore
Bear Valley USD
Ventura USD
Goleta Union
Santa Barbara County
Santa Monica/Malibu USD
Laguna Salada Union
Charter Oak USD
Norwalk-LaMirada USD
ABC USD
Fontana Unified
Manhattan Beach USD
Tulare USD
Fontana Unified
Ventura USD
West Covina USD
Glenn County
Sacramento City Schools
Tulare City School Dist.
UCLA
L.A. Community College Dist.
Fillmore USD
Santa Clara County
Laguna Salada Union
Merced City Schools
Norwalk-LaMirada USD
Goleta Union
UCLA
UCLA
San Ramon Valley USD
Anaheim Union H.S. Dist.
Monterey Peninsula USD
Tulare City School Dist.
Attendees - 6/7/84 Conference

Carol Thomas
Lories Tolbert
Rudie Tretten
Gene Tucker
Helen Turner
Marsha Viger
Dean Waldfoel
Irwin Wapner
J.C. Warneke
Steven Waterman
Chuck Weis
Debra Weiss
Jeff Wells
Robert Wenkert
Gary Wexler
Gail Wickstrom
Dick Williams
Marlus Williams
Robert Williams
Jim Wilson
Tom Wilson
John Wise
Pat Wolfe
Mrs. Freddie Wynberg
Paul Zimmerman
Ellen Zimmet

UCLA
Orange County
Laguna Salada Union
ABC USD
UCLA
Vista USD
Irvine USD
Lompoc USD
Bakersfield City School Dist.
Berkeley USD
Fillmore
UCLA
Orange County Dept. of Education
UCLA
Wm. S. Hart Union H.S. Dist.
Torrance USD
UCLA
Vista USD
Rialto USD
Tulare City School Dist.
UC Irvine
Elsinore Union H.S. Dist.
Napa County
Vallejo City USD
UCLA
L.A. City Schools
Appendix A

9:00 - 9:30  REGISTRATION/COFFEE

9:30 - 9:45  INTRODUCTION

Dr. Paul E. Heckman, Assistant Director, Laboratory in School and Community Education, University of California, Los Angeles

9:45 - 10:30  Dr. John I. Goodlad, Professor and Co-Director of the Laboratory in School and Community Education, University of California, Los Angeles

BEYOND OUTCOME MEASURES:
AN AGENDA FOR SCHOOL IMPROVEMENT

10:30 - 11:15  SMALL GROUP DISCUSSIONS

11:15 - 12:00  Dr. Eva Baker, Professor and Director of the Center for the Study of Evaluation, University of California, Los Angeles

ASSESSING LOCAL EDUCATIONAL QUALITY:
A COMPREHENSIVE SYSTEM

12:00 - 1:00  LUNCH
1:00 - 1:45  Dr. Norman Frederiksen, Educational Testing Service, Princeton, New Jersey
INFLUENCES ON TESTING AND LEARNING

1:45 - 2:30  SMALL GROUP DISCUSSIONS

2:30 - 2:45  BREAK

2:45 - 3:30  Dr. Elliot Eisner, Professor, Stanford University
USING EDUCATIONAL EVALUATION FOR THE IMPROVEMENT OF CALIFORNIA'S SCHOOLS

3:30 - 4:15  PANEL DISCUSSION
Dr. Eva Baker
Dr. Elliot Eisner
Dr. Norman Frederiksen
Dr. John I. Goodlad

4:15 - 5:00  RECEPTION