

1990

294

263

213

NATIONAL RESEARCH COUNCIL

NATIONAL RESEARCH COUNCIL

REPORTING
**DISTRICT-LEVEL
NAEP DATA**

SUMMARY OF A WORKSHOP

REPORTING DISTRICT-LEVEL NAEP DATA

NATIONAL ACADEMY PRESS

REPORTING
DISTRICT-LEVEL
NAEP DATA

S U M M A R Y O F A W O R K S H O P

Committee on NAEP Reporting Practices:
Investigating District-Level and Market-Basket Reporting

Pasquale J. DeVito and Judith A. Koenig, editors

Board on Testing and Assessment

Commission on Behavioral and Social Sciences and Education

National Research Council

NATIONAL ACADEMY PRESS
Washington, D.C.

NATIONAL ACADEMY PRESS 2101 Constitution Avenue, N.W. Washington, D.C. 20418

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competences and with regard for appropriate balance.

The study was supported by the U.S. Department of Education under contract number E95083001. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the organizations or agencies that provided support for this project.

International Standard Book Number 0-309-06893-2

Additional copies of this report are available from National Academy Press, 2101 Constitution Avenue, N.W., Washington, D.C. 20418

Call (800) 624-6242 or (202) 334-3313 (in the Washington metropolitan area)

This report is also available online at <http://www.nap.edu>

Printed in the United States of America

Copyright 2000 by the National Academy of Sciences. All rights reserved.

Suggested citation: National Research Council (2000) *Reporting District-Level NAEP Data: Summary of a Workshop*. Committee on NAEP Reporting Practices: Investigating District-Level and Market-Basket Reporting. Pasquale J. DeVito and Judith A. Koenig, editors. Board on Testing and Assessment, Commission on Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

THE NATIONAL ACADEMIES

National Academy of Sciences
National Academy of Engineering
Institute of Medicine
National Research Council

The **National Academy of Sciences** is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters. Dr. Bruce M. Alberts is president of the National Academy of Sciences.

The **National Academy of Engineering** was established in 1964, under the charter of the National Academy of Sciences, as a parallel organization of outstanding engineers. It is autonomous in its administration and in the selection of its members, sharing with the National Academy of Sciences the responsibility for advising the federal government. The National Academy of Engineering also sponsors engineering programs aimed at meeting national needs, encourages education and research, and recognizes the superior achievements of engineers. Dr. William A. Wulf is president of the National Academy of Engineering.

The **Institute of Medicine** was established in 1970 by the National Academy of Sciences to secure the services of eminent members of appropriate professions in the examination of policy matters pertaining to the health of the public. The Institute acts under the responsibility given to the National Academy of Sciences by its congressional charter to be an adviser to the federal government and, upon its own initiative, to identify issues of medical care, research, and education. Dr. Kenneth I. Shine is president of the Institute of Medicine.

The **National Research Council** was organized by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and advising the federal government. Functioning in accordance with general policies determined by the Academy, the Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in providing services to the government, the public, and the scientific and engineering communities. The Council is administered jointly by both Academies and the Institute of Medicine. Dr. Bruce M. Alberts and Dr. William A. Wulf are chairman and vice chairman, respectively, of the National Research Council.

**COMMITTEE ON NAEP REPORTING PRACTICES:
INVESTIGATING DISTRICT-LEVEL AND
MARKET-BASKET REPORTING**

PASQUALE DeVITO (*Chair*), Office of Assessment, Rhode Island
Department of Education

LINDA BRYANT, Westwood Elementary School, Pittsburgh

C. MELODY CARSWELL, Department of Psychology, University of
Kentucky

MARYELLEN DONAHUE, Planning, Research & Development, and
District Test Coordination, Boston Public Schools

LOU FABRIZIO, Division of Accountability Services, North Carolina
Department of Public Instruction

LEANN GAMACHE, Assessment and Evaluation, Education Services
Center, Littleton Public Schools, Littleton, Colorado

DOUGLAS HERRMANN, Department of Psychology, Indiana State
University

AUDREY QUALLS, Iowa Testing Program, Iowa City, Iowa

MARK RECKASE, Department of Educational Psychology and Special
Education, Michigan State University

DUANE STEFFEY, Department of Mathematical and Computer
Sciences, San Diego State University

JUDITH KOENIG, *Study Director*

KAREN MITCHELL, *Senior Program Officer*

KAELI KNOWLES, *Program Officer*

DOROTHY MAJEWSKI, *Senior Project Assistant*

BOARD ON TESTING AND ASSESSMENT

- ROBERT L. LINN (*Chair*), School of Education, University of Colorado, Boulder
- CARL F. KAESTLE (*Vice Chair*), Department of Education, Brown University
- RICHARD C. ATKINSON, President, University of California
- CHRISTOPHER F. EDLEY, JR., Harvard Law School
- RONALD FERGUSON, John F. Kennedy School of Public Policy, Harvard University
- MILTON D. HAKEL, Department of Psychology, Bowling Green State University
- ROBERT M. HAUSER, Institute for Research on Poverty, Center for Demography, University of Wisconsin, Madison
- PAUL W. HOLLAND, Graduate School of Education, University of California, Berkeley
- RICHARD M. JAEGER, School of Education, University of North Carolina, Greensboro
- DANIEL M. KORETZ, Center for the Study of Testing, Evaluation, and Education Policy, Boston College
- RICHARD J. LIGHT, Graduate School of Education and John F. Kennedy School of Government, Harvard University
- LORRAINE McDONNELL, Departments of Political Science and Education, University of California, Santa Barbara
- BARBARA MEANS, SRI International, Menlo Park, California
- ANDREW C. PORTER, Wisconsin Center for Education Research, University of Wisconsin, Madison
- LORETTA A. SHEPARD, School of Education, University of Colorado, Boulder
- CATHERINE E. SNOW, Graduate School of Education, Harvard University
- WILLIAM L. TAYLOR, Attorney at Law, Washington, D.C.
- WILLIAM T. TRENT, Associate Chancellor, University of Illinois, Champaign
- GUADALUPE M. VALDES, School of Education, Stanford University
- VICKI VANDAVEER, The Vandaveer Group, Inc., Houston, Texas

LAURESS L. WISE, Human Resources Research Organization,
Alexandria, Virginia

KENNETH I. WOLPIN, Department of Economics, University of
Pennsylvania, Philadelphia

MICHAEL J. FEUER, *Director*

VIOLA C. HOREK, *Administrative Associate*

LISA D. ALSTON, *Administrative Assistant*

Acknowledgments

At the request of the U.S. Department of Education, the National Research Council (NRC) established the Committee on NAEP Reporting Practices to examine the feasibility and potential impact of district-level and market-basket reporting practices. As part of its charge, the committee sponsored a workshop in September 1999 to gather information on issues related to district-level reporting for the National Assessment of Educational Progress (NAEP). A great many people contributed to the success of this workshop, which brought together representatives from state and local assessment offices, experts in educational measurement, and others familiar with the issues related to district-level reporting for NAEP. The committee would like to thank the panelists and discussants—many of whom traveled to the workshop during a hurricane—for their contributions to a lively and productive workshop. The full participant list appears in Appendix A.

Staff from the National Assessment Governing Board (NAGB), under the leadership of Roy Truby, executive director, and staff from the National Center for Education Statistics (NCES), under the direction of Gary Phillips, acting commissioner, were valuable sources of information. Sharif Shakrani, Mary Lyn Bourque, and Raymond Fields of NAGB and Arnold Goldstein of NCES provided the committee with important background information on numerous occasions. We also thank Keith Rust of Westat for his valuable information on NAEP sampling methodology.

Special thanks are due to a number of individuals at the National Research Council who provided guidance and assistance at many stages during the organization of the workshop and the preparation of this report. We thank Michael Feuer, director of the Board on Testing and Assessment (BOTA), for his expert guidance and leadership of the project. We are indebted to BOTA staff officer, Karen Mitchell, for her assistance in planning the workshop and writing the report; she was a principal source of expertise in both the substance and process for the workshop. We also wish to thank BOTA staff members Patricia Morison, Alix Beatty, Meryl Bertenthal, Naomi Chudowsky, Viola Horek, and John Shephard for their assistance with this work. Special thanks go to Dorothy Majewski, who capably managed the operational aspects of the workshop and the production of this report. We also thank Christine McShane for her advice on structuring the content of the report, her expert editing of the manuscript, and her deft guidance of the report through the publication process.

The committee is particularly grateful to NRC project staff, Judith Koenig, study director, and Kaeli Knowles, program officer, for their efforts in putting together the workshop and preparing the manuscript. Judith acted as the coordinator of the activities while Kaeli played a major role by contacting workshop participants and discussants, soliciting their involvement, guiding them in developing their presentations, and assisting with the writing.

This report has been reviewed in draft form by individuals chosen for their diverse perspectives and technical expertise, in accordance with procedures approved by the National Research Council's Report Review Committee. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making the published report as sound as possible and to ensure that the report meets institutional standards for objectivity, evidence, and responsiveness to the study charge. The review comments and draft manuscript remain confidential to protect the integrity of the deliberative process.

We wish to thank the following individuals for their participation in the review of this report: Susan A. Agruso, Office of Assessment, South Carolina Department of Education; Jonathan Dings, Director of Assessment, Boulder Valley Public Schools, Colorado; Richard Jaeger, School of Education, University of North Carolina, Greensboro (emeritus); Roderick J.A. Little, Department of Biostatistics, School of Public Health, Univer-

sity of Michigan; William D. Schafer, College of Education, University of Maryland; and Roger Trent, Director of Assessment, Ohio Department of Education. Although the individuals listed above have provided constructive comments and suggestions, it must be emphasized that responsibility for the final content of this report rests entirely with the authoring committee and the institution.

Pasquale J. DeVito, *Chair*
Committee on NAEP Reporting Practices:
Investigating District-Level and
Market-Basket Reporting

Contents

1	Introduction	1
2	Background	6
3	NAEP's Influence on State Instructional and Assessment Programs	14
4	Comparisons with National Benchmarks: Pros and Cons	19
5	Factors That Influence Interest in District-Level NAEP	24
6	Summing Up: Issues to Consider and Resolve	38
	References	42
	Appendix A: Workshop on District-Level Reporting for NAEP Agenda and Participants	45
	Appendix B: Background Information on NAEP	51

Introduction

The National Assessment of Education Progress (NAEP) has earned a reputation as one of the nation's best measures of student achievement in key subject areas. Since its inception in 1969, NAEP has summarized academic performance for the nation as a whole and, beginning in 1990, for the individual states. Increasingly, NAEP results get the attention of the press, the public, and policy makers. With this increasing prominence have come calls for reporting NAEP results below the national and state levels. Some education leaders argue that NAEP can provide important and useful information to local educators and policy makers. They want NAEP to serve as a district-level indicator of educational progress and call for NAEP results to be summarized at the school district level.

At the same time, others have called for simpler, more intuitive and meaningful reporting of NAEP results. Advisers to the National Assessment Governing Board (NAGB) and the National Center for Educational Statistics (NCES) have proposed the use of market-basket reporting methods as one means to accomplish this. Market-basket reporting would allow results to be reported as percentages of items correct on sets of representative items. As part of their evaluation of NAEP, the National Research Council's Committee on the Evaluation of National and State Assessments of Educational Progress stressed the need for clear and comprehensible reporting metrics that would simplify the interpretation of results and endorsed the concept of market-basket reporting for NAEP (National Re-

search Council, 1999a). Market-basket reporting would provide an easier-to-understand picture of students' academic accomplishments.

In pursuit of improved reporting and use of test results, NAEP's stewards are exploring the feasibility and potential impact of district-level and market-basket reporting practices. Accordingly, at the request of the U.S. Department of Education, the National Research Council established the Committee on NAEP Reporting Practices to examine the feasibility and potential impact of district-level and market-basket reporting practices. Because these two topics are intertwined, the committee is examining them in tandem, focusing first on district-level reporting.

During the course of the study, the committee is seeking to answer the following questions regarding district-level reporting for NAEP:

- (1) What are the characteristics of district-level NAEP?
- (2) If implemented, what information needs might it serve?
- (3) What is the degree of interest in participating in district-level NAEP? What factors would influence interest?
- (4) Would district-level NAEP pose any threats to the validity of inferences from national and state NAEP?
- (5) What are the implications of district-level reporting for other state and local assessment programs?

To begin to address these questions, the committee convened the Workshop on District-Level Reporting for NAEP on September 16 and 17, 1999. Although this workshop relates to one of the committee's charges, it was not intended to bring closure on issues related to district-level NAEP. The committee's work will continue with a workshop on market-basket reporting in February 2000, and joint consideration of the two issues will be taken up in the final report.

WORKSHOP ON DISTRICT-LEVEL REPORTING

The purpose of the National Research Council's Workshop on District-Level Reporting for NAEP was to explore with various stakeholders their interest in and perceptions regarding the likely impacts of district-level reporting. NCES has, to date, had two experiences with district-level reporting. In 1996, NCES contacted several of the larger school districts in the country to gauge their interest in receiving district-level results. The

data collected by these districts could potentially meet the requirements for district-level reporting through augmentation of the state NAEP samples, although there would be a fee associated with the augmentation procedures. In 1998, NCES identified several districts that met the sample size requirements “naturally” as a result of the state NAEP sampling procedures; these districts are referred to as the “naturally occurring districts.” Additional details on these experiences with district-level reporting appear in Chapter 2.

The workshop consisted of four panels, each with a specific goal. The opening panel was designed to provide broad context for the two-days of workshop discussions. This panel explored the purposes that district-level reporting might serve, discussed who might use the results and how they might be used, and highlighted the key issues that should be considered. Panelists included two individuals who had earlier authored papers discussing the advantages and disadvantages of “below-state” reporting for NAEP as well as representatives from the Council of Chief State School Officers and the Council of Great City Schools. A representative from a district that would qualify for receiving NAEP results served as the discussant.

The second panel provided an opportunity for the committee and workshop participants to hear several state assessment directors discuss the impact that state NAEP has had on their state and local education policy, instruction, and assessment. The concerns that have been expressed regarding district-level reporting at previous committee meetings and elsewhere parallel those considered when state NAEP was implemented in 1990. Thus, the committee thought it would be useful to reflect on the lessons learned from state NAEP as they consider the likely impact of district-level reporting. The presenters focused on the ways in which state NAEP has affected their assessment and instructional programs, the types of comparisons made between NAEP results and state and local assessment results, and what happens when results from various assessments portray differing pictures of achievement. The states represented on this panel were Colorado, Connecticut, Nevada, and Washington. Steve Dunbar, a member of the National Research Council’s earlier Committee on the Evaluation of National and State Assessments of Educational Progress, served as the discussant.

The third panel brought together district and state assessment directors to discuss their interests in district-level results, the types of information district-level NAEP would provide, the ways in which district-level results might be used, factors that would bear on their decisions to partici-

pate, and issues regarding who should make participation and score release decisions. This panel had four subpanels: the first three subpanels included representatives from naturally occurring districts; each subpanel paired a district representative with a representative from the respective state assessment office, and each speaker addressed the issues in turn. The final subpanel consisted of representatives from districts that had expressed interest in receiving district-level NAEP data in 1996.

Representatives from NAGB, NCES, and the contracting organizations that work on NAEP (the Educational Testing Service and Westat) sat on the final panel. This panel highlighted the technical issues related to sampling and scoring methodologies for district-level reporting and the policy issues related to participation and reporting decisions. Laress Wise, also a member of the earlier Committee on the Evaluation of National and State Assessments of Educational Progress, served as the discussant.

The workshop was structured so as to permit considerable discussion by presenters as well as participants, much of which is woven into this summary. Time was allotted for each speaker, and following each presentation, substantial time was devoted to open discussion. In preparation for the workshop, speakers were given sets of questions to address during their presentations and asked to supply written responses in advance.¹ Questions posed to the various panelists are included on the agenda, which appears in Appendix A.

ORGANIZATION OF THIS REPORT

The purpose of this summary is to capture the discussions and major points made during the workshop in order to assist NAEP's stewards in their decision making about implementing below-state reporting and to provide information for those who would make decisions about whether or not to participate. The summary is organized as follows. Chapter 2 provides background information on NCES's past experiences with reporting district-level results, along with a discussion of the benefits associated with and the concerns expressed about the implementation of state NAEP, since

¹Due to inclement weather (Hurricane Floyd), a number of participants were unable to attend the meeting, although some did participate via speaker phone. Their written responses were used in this summary.

it is expected that these parallel issues related to district-level NAEP. Chapter 3 summarizes the information presented by Panel 2 speakers regarding the impact that state NAEP has had on state and local instruction and assessment programs.

The next two chapters reflect the common themes that emerged from discussions at the workshop. There was considerable overlap in the nature of the comments made across the four workshop panels. Thus, instead of summarizing each panel's discussions separately, we have organized these two chapters around the common issues raised during the workshop. Much of the discussion focused on issues related to comparing results from different districts. Chapter 4 is therefore devoted to the subject of interdistrict comparisons. Chapter 5 highlights participants' comments regarding factors that bear on their interest in district-level data. Issues to consider and resolve are summarized in the Chapter 6. Appendix A contains the workshop agenda and list of participants; Appendix B contains general background information on NAEP.

Background

This chapter provides background information on experiences with state NAEP and the reporting of district-level NAEP results. The first section describes some of the concerns expressed during the early implementation stages of state NAEP, discusses findings from initial evaluations of the program, and highlights their relationship to district-level reporting. The second section describes prior experiences NCES has had with reporting district-level results through the Trial District Assessment in 1996 and the reporting of results for naturally occurring districts in 1998.

THE STATE NAEP EXPERIENCE

The Trial State Assessment (TSA) was designed with several purposes in mind: (1) to provide states with information about their students' achievement and (2) to allow states to compare their students' performance with that of other students in the states (National Academy of Education, 1993). Implementation was on a trial basis to allow for congressionally mandated evaluations of the program's feasibility and utility before committing resources to an ongoing state-by-state assessment. Prior to its implementation, a number of concerns were expressed about its possible impact. The text below describes some of these concerns, cites some of the benefits reported in reviews of the TSA, and notes how these concerns relate to district-level NAEP.

Early Concerns About Implementation of State NAEP

Concerns about state NAEP centered around the anticipated uses of state-level data and the consequent effects on test preparatory behaviors. Reporting of national-level results had been regarded as having low stakes, since decisions at the state, district, school, or classroom level could not be based on NAEP reports. National-level data were not being used for accountability purposes, and participants were relatively unaffected by the results. But the provision of state-level data prompted concerns about the effects of increasing the stakes associated with NAEP.

As enumerated by Stancavage et al. (1992:261) in discussing the TSA in mathematics, NAEP's stakeholders asked:

- (1) Would the reporting of the NAEP TSA cause local districts and states to change the curriculum or instruction that is provided to students?
- (2) Would local or state testing programs change to accommodate NAEP-tested skills, would they remain as they are, or would they simply be pushed aside?
- (3) Would any such changes in curriculum or assessment, should they occur, be judged as positive by mathematics educators, and others, or would the changes be viewed as regressive and counter-productive?
- (4) Finally, would it be found that the entire NAEP TSA effort had no impact at all and was, therefore, a wasteful expenditure of time and money?

These questions stemmed from concerns about the emphases attached to and the inferences drawn from NAEP results. Increasing the stakes associated with NAEP was seen as a move toward using NAEP results for accountability purposes. It was feared that such uses would degrade the value of the assessment. Koretz (1991:21) warned that higher stakes would bring inappropriate teaching to the test and inflated test scores, adding that NAEP results, so far, had been free from "this form of corruption." While this is an important concern, it should also be noted that when state standards mirror the NAEP frameworks, having schools teach the content and skills assessed by NAEP is a desirable result.

Beaton (1992:14) used the term "boosterism" to describe the activities that might be used to motivate students to do their best for the "state's