

October 1994

Educational policy analysis archives

Arizona State University

University of South Florida

Follow this and additional works at: https://digitalcommons.usf.edu/usf_EPAA

Recommended Citation

Arizona State University and University of South Florida, "Educational policy analysis archives" (1994).
Education Policy Analysis Archives (EPAA). 277.
https://digitalcommons.usf.edu/usf_EPAA/277

This Book is brought to you for free and open access by the USF Faculty Collections at Digital Commons @ University of South Florida. It has been accepted for inclusion in Education Policy Analysis Archives (EPAA) by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact scholarcommons@usf.edu.

Education Policy Analysis Archives

Volume 2 Number 13

October 7, 1994

ISSN 1068-2341

A peer-reviewed scholarly electronic journal.

Editor: Gene V Glass, Glass@ASU.EDU. College of Education,
Arizona State University, Tempe AZ 85287-2411

Copyright 1994, the EDUCATION POLICY ANALYSIS
ARCHIVES. Permission is hereby granted to copy any article
provided that EDUCATION POLICY ANALYSIS ARCHIVES is
credited and copies are not sold.

Carrot or Stick? How Do School Performance Reports Work?

Mark E. Fetler
California Community Colleges Chancellor's Office

Abstract: State and federal government espouse school performance reports as a way to promote education reform. Some practicing educators question whether performance reports are effective. While the question of effectiveness deserves study, it accepts the espoused purposes of performance reports at face value, and fails to address the more basic, tacit political and symbolic roles of performance reports. Theories of organization, modern government, and regulation provide a context that helps to clarify these political and symbolic roles. Several performance report and assessment programs in California provide illustrations.

Introduction

"The buck stops here." --Harry S. Truman

"State and federal governments view performance reports as instruments of policy to help promote education reforms. "Decisions about desired outcomes and conditions will determine the nature of any indicator system ... these decisions will be political." (Oakes, 1986, p. 24) Accountability systems are "very powerful policy tools." (OERI, 1988, p. 31) "An apparent strategy imbedded in most state indicator systems is that they will be used to guide future policy." (Brown, 1990, p. 5) An education indicator information system "can be used by policymakers responsible for defining the nation's education agenda to monitor the education outcomes they consider most significant." (NCES, 1991) The function of an accountability mechanism "is to oversee (monitor and evaluate) the performance of the education system and to propose needed changes to policy makers." (Wohlstetter, 1991, p. 31) "It is hardly a novelty for testing and assessment to figure prominently in policymakers' efforts to reform education." (Linn, 1993, p. 1)

The above statements presume a relationship between accountability and politics. The authors are clear about the meaning of "accountability," "indicator system," and "assessment." Although the authors may be equally clear about "politics," "policy," and "power," these terms are not well defined. Traditionally, educators define accountability as a system with goals (educational reform), inputs (indicators), processes (reporting, incentives), and results (school change). How does "politics" enter the picture? How effective are accountability systems?

One barrier to answering these questions is that the traditional view of accountability tends to jumble political matters (decisions about the use of limited resources) with structural issues, (roles and responsibilities of management and staff), with human resource issues, (authoritarian versus need oriented management styles), or culture, (the symbols, rituals, myths, and theater that cloak public schools). For example, Mitchell and Encarnation (1984) consider together such diverse "policy mechanisms" as structural organization, revenue generation, resource allocation, program definition, personnel training, assessment, and curriculum--an approach that clusters diverse aspects of organizations into the single category of "policy." "Politics" is a word that through frequent repetition in many contexts appears to have lost any precise meaning. A second barrier to answering the questions is the common view of a monolithic government, which fails to discriminate the internal from the external orientations of government, and the forces that mediate between these orientations. These different aspects of government often have distinct perspectives on policy.

An alternative to the traditional approach, suggested by Bolman and Deal's (1991) multifaceted analysis of organizations, examines accountability not only from a structural viewpoint, but also from political, cultural, and human resource perspectives. Galbraith's discussion of modern government (1983) and Thurow's (1981) analysis of regulation deepen the understanding of the relationship between accountability, policy, and power.

Of course, there are other non-standard approaches to accountability, assessment, and evaluation. For example, Sirotnik and Oakes (1990) observe that school culture (roles, norms, expectations, assumptions, beliefs) influences school conduct. This cultural view presumes a sense of wholeness of schools and consideration of sources of resistance to change. Moreover, the evolution of program evaluation bears witness to the creative use of various metaphors -- law, art, history, hermeneutics, etc. -- to develop contrasting methods of evaluation.

Conventional Wisdom

"Taking a car's measure ... The testers note whether the gauges and displays are easy to read, whether the controls are logically placed, and whether drivers of various sizes can easily work the controls and see out." (Consumer Reports, p. 220)

Educators often view accountability almost mechanically as a steering system that monitors progress toward goals. The system sounds an alarm when schools veer off track, do not function efficiently, or do not meet goals. Accountability indicators signal an opportunity to make adjustments. Key accountability system design issues are who has oversight, the conditions of education to be monitored, the measures or indicators, performance standards, who is responsible, and what the consequences are for meeting or falling short of the standards. Resolving these issues defines the type of accountability system. While there are many different kinds of accountability systems -- program review, compliance monitoring, fiscal audits, etc. -- this paper focuses on the recent development of performance reports.

The current vogue for school performance reporting traces its roots back to the early 1980s when various national reports portrayed public schools as at risk and on the brink. The US. Department of Education published an "Education Wallchart" or "State Reportcard" that described the condition of state education systems using test results, dropout rates, funding, and

staffing. Many states followed suit with similar school performance reports. (Council of Chief State School Officers, 1987) States hoped to use performance reports for a number of purposes. Educators could use the information in order to promote school improvement. The reports would motivate parents to lobby school boards, PTA groups, and other organizations to make desired changes. Newspaper comparisons would spur schools to compete for favorable standings. Performance incentives based either on money, recognition, or embarrassment would help to motivate school improvement.

While state approaches vary, there is consensus on the elements of a standard performance report system. (Oakes, 1986; OERI, 1988; Kagan and Coley, 1989; Blank, 1993) Very briefly, a performance report system begins with a systemic model based on research of how schools and the education system function. The model specifies major inputs (fiscal, staffing, and students), processes (curriculum, instruction, and services) and desired results (student achievement, skills, attitudes, college attendance, or employment). To be useful, the indicators, defined as "measures of the condition of education," must meet certain criteria, e.g., measure the central features of schooling, measure what is actually taught, provide policy relevant information, focus on the school site, allow for fair comparisons, and maximize usefulness and minimize burden. Some authorities emphasize outcome indicators (Murnane, 1987), others argue for contextual information (Oakes, 1989), and yet others recommend measures of process. (Porter, 1991)

Dissemination strategies encompass the potential audiences for the report and include quality standards that rely on comparisons. Schools can compare their own present to past performance by tracking an indicator over a period of years. Norms can help to judge school performance in comparison to an overall population, or to a subgroup of schools that are socially or demographically similar. (Fetler, 1991) School performance can be predicted by statistical regression using relevant background measures that are not readily controlled by schools, e.g., parent education or economic status. (Salganik, 1994) Comparisons can also be made after standardizing all school scores to a common (state or national) demographic mixture. (Wainer, 1994) Content standards specify the ability of some percentage of students to complete a given task or demonstrate a skill. Ernest Boyer describes the current interest with performance standards: "I think we've gone about as far as we can go in the current reform movement dealing with procedural issues." By establishing national academic standards and exams, schools "would be held accountable for outcomes rather than the current situation of heavy state regulation that nibbles them to death over procedures." (Center for Research on Evaluation, Standards, and Student Testing, 1989)

In summary, the conventional approach to performance reporting tends to adopt a rational perspective. The focus is on structural characteristics (budgets, staffing, processes) and the outcomes that this structure is designed to produce. The rational approach pays less attention to the ways that people interact to do work, cultural values and norms, and the political process of dividing up resources. However, understanding how performance reports affect organizations requires consideration not only of structure, but also of other organizational dimensions.

Organization

The American Heritage Dictionary defines "organization" as something comprising elements with varied functions that contribute to the whole and to collective functions. Different theories of organization--structural, human resource, cultural, and political--focus on different elements and functions. To examine the relationship between politics and accountability, it helps to characterize these theories and suggest how they apply to schools. Consideration of these different organizational perspectives also helps to clarify the idea of "politics" in the context of accountability.

Structure

Structurally, schools are relatively closed systems with explicit inputs, processes, and outcomes. Two central issues are how to divide and coordinate the work. For example, faculty are classified by grade level or subject matter. Schools are organized by grade-span, and sometimes by curriculum or type of student (academic, regular, continuation, or special needs). Different types of facilities best meet the specific needs of the elementary, middle, or high schools. Vertical lines of control run from the board to the superintendent, administrative staff, principals, to faculty and staff. Committees, or matrix management schemes address lateral coordination. There are explicit rules and procedures for providing instruction, delivering services, and administering the system. Plans, budgets, and accounting systems help to monitor school operations.

A number of factors influence structure. Bigger schools tend to require more coordinating mechanisms and clearer lines of authority. The "core technology,"--teaching and learning-- and beliefs about cause and effect relations are important factors. For example, vocational and college preparatory programs often involve different curricula, equipment needs, instructional techniques, kinds of students, and student outcomes. The social and demographic environment is an influence. Districts in high growth versus older established neighborhoods face different challenges and must organize differently. Different school goals will produce different structures, e.g., academic excellence, citizenship, character development, custody and control, efficiency, or equity. An emphasis on equity might result in highly diverse classrooms, collaborative teaching styles, and a wide array of services for disadvantaged students. An efficiency goal to standardize instruction might encourage tracking and sorting of students.

The traditional view of educational accountability has a strong structural bias. Categorical program regulation and accreditation review, for example, tend to focus on structural features, such as staffing, organizational charts, published rules and procedures, budgets, expenditure reports, plans, and record keeping. Despite a focus on student outcomes and results, the current trend in school performance reports is no less structural in its emphasis.(Oakes, 1986; Shavelson, 1991; Shavelson, McDonnell, and Oakes, 1989; OERI, 1988) That is, performance report indicators should reflect a rational, systemic model of the educational process. Reports of student achievement are helpful if they causally refer to structural factors that can be changed, e.g., staffing, budgeting, planning, facilities, curriculum, instructional techniques, technology, etc.

Human Resources

The human resource frame deems the central task of managers to be harmonizing the needs of schools with the needs of people who work in them. One view assumes that managers direct and control the work of subordinates, who prefer to be led and who resist change. A different, needs based, view is that managers must arrange conditions so that the employee's self-interest coincides with the organization's interests. In practice, managers may find it difficult to attain harmony. As people mature and develop they become more independent, attain a broader perspective and range of skills, and develop a longer time perspective. This developmental process appears to hold not only for faculty, but also for other professionals who provide education services. However, organizations often treat people like children by requiring higher level managers to direct and control subordinates. Managerial domination can result in psychological failure, passivity, and dependence.

The conventional approach to performance reporting often presupposes an inadequate and demeaning approach to human resource management. Performance reports typically focus on the institution as a whole, and not on relationships and the process of implementing change. The institutional focus, and lack of concern with individuals, would appear to be consistent with a

top-down, controlling management style. Performance reports aim to produce change by aligning a school's need for recognition, incentive funding, or avoidance of bad publicity with the government's desire to meet goals for graduation, college attendance, or employment. Schools decide for themselves how to meet goals and manage change. An example is Arizona's Student Assessment Program (ASAP). (Noble and Smith, 1994) ASAP's goals are instructional improvement and accountability. Arizona designed the assessment to be consistent with current cognitive theory. To score well schools must adopt new curricular and instructional practices. Arizona widely publicized the results, including school comparisons. However, ASAP does not specify how schools are going to change from their traditional models of teaching to the desired models. Districts vary in their capacity to accomplish the outcomes. Some lack the money to train faculty, and some are loyal to their traditional methods.

Politics

The political frame views schools as arenas where coalitions compete for limited or scarce resources, e.g., money, staff, facilities. Conflict is inevitable and decisions are the product of bargaining and negotiation. Coalitions arise when people have common enduring interests, values, or beliefs. Hierarchy, budgeting, and diversity can stimulate coalition formation. Collective bargaining highlights the faculty union and management groups. A school's budget may profile faculty, instructional support groups, student services groups, athletics organizations, or extra-curricular groups. Diversity can lead to advocacy groups for underprivileged, special education, or gifted students. Similar coalitions arise as the result of state and federal budget processes. Coalitions at either the local, state, or federal level seek to influence decisions made by groups at other levels.

The distribution and use of power to distribute resources is central to politics. The authority that inheres in an organizational position is one kind of power. Power also flows from individual expertise or control of information, the control of rewards, the coercive ability to punish, alliances, control of agendas, the ability to persuade with symbols and myths, and personal charisma. Given the multiple sources of power, school goals rarely come those who are nominally in charge. Goals are more frequently a product of negotiation among those people who have power. Conflict among people with power is not primarily a matter of selfishness or incompetence, but rather results from enduring differences between groups, limited resources, and the distribution of power.

Murnane (1987) suggests that the political basis for accountability programs derives from the hierarchy of federal, state, and local agencies. The federal government delegated major responsibility to states that in turn delegated to local school districts. Schools draw their funds from all three levels and in principle must respond to their requests for accountability. As public policy issues become more important (e.g., illiteracy, equity, or diversity), government seeks to measure and report on relevant conditions (e.g., student achievement of underrepresented groups). In practice, state and federal data collection must accommodate local officials who can refuse to participate or who can delay for reasons of cost or for fear of inappropriate or unfair comparisons. State and federal government cannot unilaterally require data collection without endangering the quality and timeliness of reports. Most performance report systems result from negotiation with local educators. The conflict that arises during those negotiations is a natural and inevitable result of the political process.

Culture

"Evaluation is a ritual whose function is to calm the anxieties of the citizenry and to perpetuate an image of government rationality, efficiency, and accountability.

The very act of requiring and commissioning evaluations may create the impression that government is committed to the pursuit of publicly espoused goals, such as increasing student achievement" (Floden and Weiner, 1978, cited in Bolman and Deal, p. 284)

How people assign meaning and interpret their experience in organizations depends on culture. The hard-headed realist may dismiss cultural symbols, rituals, or myths. However, the symbolic point of view values "meaning" more highly than "reality." The more ambiguous and uncertain a situation, the less easy it is to rationally analyze, and the more likely that people will create symbols that support faith in order and predictability. Diplomas, textbooks, tests, grades, report cards, chalkboards, etc., all have one kind of meaning for a school planner. They are also deeply embedded symbols that define what many people expect from a "school."

Rituals help the members of an organization to define what it means to be a member, what to believe, how to behave, and what is important. Rituals embody the accreted culture of a school and help to instruct new members in that culture. Schools open in September, hold classes for two semesters with set holidays, and let out for summer vacation. These and other patterns of behavior and events define educational roles responsibilities, and processes. In another sense they are rites of matriculation, passage, and commencement that help people make sense of the educational process.

A common way of discounting an explanation is to label it as a "myth" or "story." While myths and stories are not empirically testable, they present clear and vivid messages that establish meaning, solidarity, certainty, and legitimacy. People judge the legitimacy and worth of a school by the correspondence between prevailing myths and actual structural characteristics. Public esteem depends on properly maintained facilities, professional behavior of faculty, availability of instructional materials, and appropriate procedures for teaching, testing, and grading. Ideally, an appropriate structure, combined with the right technology and processes will result in effective teaching and learning. However, to many the appearance of a school is more important than effectiveness.

The symbolism of a school performance report relies on the connotations of accountability for responsibility, integrity, and trust. People entrust their children directly to schools. Indirectly, through government taxation and education funding, they entrust their money to schools. A report on the measurable benefits of education for children can have several meanings. The report can affirm school integrity by documenting the consistency of the school's mission, e.g., teaching and learning, with the goals of the educational program and student outcomes. The act of reporting can also affirm the school's willingness to take responsibility for carrying out its mission. The connotations of integrity and responsibility are stronger if people perceive schools as accepting accountability. This might be termed the "Truman effect," from the hand lettered sign on the President's desk, "The buck stops here." By contrast, if the perception is of government coercion, the connotation is negative.

Government

"Few words are used so frequently with so little seeming need to reflect on their meaning as power, and so it has been for all the ages of man." - John Kenneth Galbraith

Many perceive government as a monolithic agent wielding power in order to implement its goals. Many perceive that government imposes accountability on schools in order to promote reform policies. The fallacy in this perception is that modern governments do not typically behave as large disciplined units. Galbraith (1983) distinguishes three aspects of government; an

inner orientation, an exterior orientation, and a force that mediates between the two orientations. These orientations respond to different constituencies and in practice have different viewpoints on accountability. The distinctions made here between the types of government will be useful in the case studies of California's performance reporting programs, below.

The exterior orientation comprises the legislature, voters, and the many organizations that seek to influence both the legislators and voters. Organized groups, e.g., faculty unions, administrator associations, political action committees, book, or test publishers, may seek to sway legislators and voters, either by lobbying or by public information campaigns. For example, a union might oppose student assessment in the context of performance review, but might support assessment as a justification for increased funding. A test publisher might oppose an assessment program that relies on state developed tests, but support one that relies on commercially available instruments.

The inner orientation refers loosely to the bureaucracy and the many organizations that administer the tasks of government. Continuity and relative autonomy characterize the inner orientation. The power of the bureaucracy is in preparing budgets, overseeing programs, and developing regulations. The inner orientation also promotes its goals to the public by providing information -- speeches, memoranda, advisories, press conferences, etc.

The chief executive, her staff, cabinet, and appointees embody the force that stands between and mediates the external and internal orientations. In California, for example, the Governor controls the budgets of government agencies, can blue-pencil budget line items, can veto or sign legislation, and persuades by means of press conferences, speeches, and news releases.

Galbraith considers society to be in an equilibrium between those who use power and those who oppose it -- the result of a "dialectic of power." For instance, when the bureaucracy opposes an employee union, the union can turn to the legislature or the executive. A governor can also weaken a bureaucracy that is too effective in enforcing claims on the state budget.

Regulation

Regulation is one way that government can exert power and enforce accountability. Historically, government regulations have accompanied new funding for categorical programs, e.g., Title I, vocational education, bilingual education, special education, etc. Perhaps in recognition of local political pressure on budgets, state and federal agencies have invoked regulation to encourage schools to spend funds for their earmarked purposes. Whatever the original intent, the public has come to see regulation as burdensome and as an ineffective way of attaining program goals. Performance reports are sometimes proposed as an alternative to regulation that be a more effective tool for meeting program goals. However, Thurow's (1981) analysis of regulation suggests that performance reports actually are a kind of regulation. Further, an attempt to substitute performance reports for existing regulation is likely to encounter resistance.

Thurow observes that government justifies regulation as a way to accomplish a worthwhile social goal. Implicitly, however, regulations always alter the distribution of income, and this is the real reason for their existence. Judgments about the goodness of regulations presume a vision of what distribution of income should exist, and how to create this distribution. Two related issues are how regulations come into existence, and the failure of deregulation. In the US regulations are created in response to real problems. Markets are not performing some task and/or public tolerance for failure has weakened.

It is difficult to eliminate established regulations. Regulations persist because they create vested interests. People -- faculty, students, counselors -- enter schools expecting certain kinds of employment and services, and meeting this expectation depends on regulations. There may be

more people who pay for these regulations (taxpayers) than there are students and teachers. The per capita loss to students and teachers is much larger than the per capita gain to taxpayers. "Intensity overwhelms numbers, and the resistance to deregulation may be much stronger than the pressures for it." Thurow's view is that regulation is inevitable. If a majority wishes to provide services to disadvantaged students or establish equity in educational outcomes, government will adopt regulations. Regulations may be frustrated or they may be ineffective. The result is not a return to the "status quo ante," but the adoption of new and more stringent regulations.

There are two basic sets of regulatory instruments. One set influences the production of goods or services. Thurow refers to these as q-regulations since they affect quantities of things. Another set of regulations attempts to levy taxes or subsidies to encourage or discourage the production of goods and services. Thurow refers to these as p-regulations since they affect prices. With p-regulations the government agency takes advantage of market incentives. Q-regulations attempt to fight market incentives.

Thurow's analysis suggests that performance reports are regulatory. In education q-regulations prescribe how schools spend money, e.g., to hire staff, purchase equipment, or serve certain groups of students. These are the conditions for accepting categorical program funds. P-regulations involve establishing incentives for desired student outcomes or results, and sanctions for undesirable outcomes. The incentives can be intangible, as in the publicity surrounding a newspaper report or a recognition program, or they could be fiscal, relating to competition for enrollment or performance based funding. To the extent that performance reports are effective they establish incentives and are a type of p- regulation.

A strategy of justifying accountability systems in exchange for deregulation may be difficult to implement. Thurow's analysis suggests that deregulation of categorical programs would threaten the incomes of faculty and service providers. An effective accountability program will change school budgets. There will be pressure not to deregulate. Consistent with this view, Lorraine McDonnell (Rothman, 1993) observes that Americans support procedural equity, or a process to ensure that everyone has access to valued goods. Substantive equity, or equal results, does not enjoy public support, in part because it demands redistribution of resources.

California

California joined many other states in using accountability programs to promote educational reform goals. Three of these programs stand out for their visibility, ties to education funding, and amount of effort. They include a school performance report program, a school report card mandate, and the California Learning Assessment System. Different aspects of government took the lead in designing these programs, with varying consequences for schools and for the programs themselves.

Performance Reports

The California Department of Education (CDE) started a school performance report program in 1983. (Fetler, 1986) Research staff designed and implemented the program at the request of CDE's elected superintendent. Although the program trailed major school reform legislation, it had no legal mandate. The staff had primary responsibility for design, production, and dissemination of the report, as well as its use in school recognition programs. While CDE controlled the program and made the final decisions, it also consulted local school administrators, researchers, and various education organizations.

CDE claimed that the selection of indicators and ways of displaying them reflected the state's judgment of what is most important for schools. Larger enrollments in academic courses

and higher achievement test scores indicate school quality. CDE considered these indicators to be consistent with the reforms enacted in 1983 -- a strengthened curriculum, higher grading standards, and more rigorous graduation standards. Special interest group lobbying added to the performance reports. Initially, in 1984, the California report limited itself to school average test scores. Next years' report disaggregated results by ethnicity and English language fluency. Initially, the report included enrollments in specific academic courses. This expanded to include enrollment in courses that met the University of California entrance requirements. For the first few years the reported included SAT scores. ACT test results were soon added. Other groups managed to stay out of the report. For instance, research staff attempted unsuccessfully for several years to define and collect indicators for vocational programs.

CDE annually produced a quality indicator report for the state, each district, and school. The high school indicators reflected academic course enrollments, attendance, dropouts, state achievement test results, SAT, and ACT scores. School and district reports showed progress over time, its statewide rank, and its standing compared to other demographically similar schools. CDE set absolute improvement goals for each indicator. Schools could meet normative goals by performing in the highest 25 percent of a demographically similar group. Goal attainment was the primary criterion for selection into state and federal school recognition programs. School district and county offices received the reports about two weeks before release to the press.

A 1984 survey of board members, superintendents and administrators foreshadowed the question of performance reporting effectiveness. CDE conducted workshops to explain the program to local school officials. Four questions were asked in order to evaluate the workshops. How important are the proposed quality indicators? How realistic are the state goals? How comprehensive are the school profiles? How fair are the analysis and approach? Possible responses to these questions were "very," "some," and "not at all." Analysis of 674 questionnaires from 21 sites statewide found that 43 percent responded that accountability was "very important," 21 percent responded that the goals were "very realistic," 15 percent said the profiles were "very comprehensive," and 10 percent felt that the approach was "very fair." There was more support for the idea of accountability than for the state's proposed performance report program. This uneven local support may have weakened program effectiveness. For example, one large Southern California school district apparently confiscated the school reports in the central mailroom. Anecdotally, some superintendents used the reports for conferences with principals, but others criticized the reports to local newspapers.

Oakes (OERI, 1988) surveyed 350 educators in California and several other states to examine how accountability systems influence schools and classrooms. The survey found that performance accountability systems caused schools in California, Florida, and Georgia to change the way they plan and teach. Apparently, schools did not direct these changes at improved effectiveness. Accountability systems caused administrators and faculty to change because they want students to perform well on State tests.

Presumably, performance reports stimulate public interest, raise the stakes for schools, and cause school improvement. However, the effectiveness of California's performance reports has never been carefully evaluated. For example, do course enrollment indicators really result in greater student participation in rigorous academic courses? Do they result in greater participation in less stringent courses with academic titles? Do achievement indicators elicit a higher priority for academic instruction and curriculum, or do they encourage narrow teaching to the test and irregular practices in test administration and scoring? The studies reported by OERI (1988) and Fetler (1986) only assessed perceptions, not the attainment of program goals.

While the effectiveness of performance reports for school improvement remains in question, the existence of the program per se, and the information provided by the indicators became a staple weapon in the annual battle over K-12 education's portion of the state budget, in campaigns for education bonds, and in the passage of legislation favoring schools. In particular,

the campaign for Proposition 98, which has formed the basis of education finance in California since 1988, depended in part on performance report data.

Proposition 98 and School Report Cards

California voters passed a constitutional ballot initiative in 1988, Proposition 98, which addressed school funding and accountability. (Fetler, 1990) Many considered the school report card requirement of Proposition 98 as significant in swaying public support for the measure, which succeeded by a very thin margin. A number of politically significant events preceded Proposition 98. During the 1980s California's growing and increasingly diverse population pressured the public education system, while strained state finances made school funding more difficult. A 1987 study by Policy Analysis for California Education (PACE) described California's school finance in the early 1980s as "unstable and uncertain." Tax reform measures in the 1970s had cut local property taxes in half, limited the ability of governments to expend revenues, and shifted the balance of funding away from local to the state level. Funding in constant dollars declined and student enrollments increased. A growing population of disadvantaged and underrepresented students brought increased linguistic diversity, poverty, and threw up new barriers to teaching and learning. Rising public expectations helped pass major reforms in 1983, which encouraged higher pupil achievement, standards for faculty, increased instructional time, tougher graduation requirements, and incentives for student, faculty and school performance.

Before 1988, the California Constitution capped the amount of taxes that government, including school districts, could appropriate. Any excess was returned to taxpayers. Proposition 98 altered the cap by specifying a minimum funding level for schools, which was set either at the percentage allocated in 1986-87, or at the same amount received in the prior year, whichever was larger. In addition, schools would receive any revenues which exceeded the cap. These excess revenues permanently increased the minimum school funding levels and must be used for "instructional improvement and accountability." The initiative also required the Department of Education to develop a Model School Accountability Report Card, which contained information on thirteen school conditions, including:

- Student achievement in and progress toward meeting reading, writing, arithmetic, and other academic goals.
- Progress towards reducing dropout rates
- Estimated expenditures per student and types of services funded
- Progress toward reducing class sizes and teaching loads
- Any assignment of teachers outside their subject areas of competence
- Quality and currency of textbooks and other instructional materials
- The availability of qualified personnel to provide counseling and other student support services
- Availability of qualified substitute teachers
- Safety, cleanliness, and adequacy of school facilities
- Adequacy of teacher evaluations and opportunities for professional improvement
- Classroom discipline and climate for learning
- Teacher and staff training and curriculum improvement programs
- Quality of school instruction and leadership

Each district board must issue an annual report card for each school that addresses the thirteen conditions. Local boards can develop their own report card, but must compare their own document with the state model every third year.

Support for Proposition 98 coalesced from several directions. The California Teachers Association spent over seven million dollars on the ballot campaign. Other major fund raisers were California's Superintendent of Public Instruction, the state PTA, and the Association of California School Administrators. The supporting coalitions naturally were involved in designing the school report cards. The Department of Education convened a task force, composed of teachers, administrators, parents, school board members, classified employees and researchers. Representatives and friends of the California Teachers Association were well represented. The consultation resulted in a 27-page document, calling for new data collection, and reporting in more areas than the 13 required by Proposition 98.

Even though administrators had a voice on the original task force, the Association of California School Administrators developed its own model "in order to assist the State Board in its deliberations, and to insure that the specifications and reporting requirements of the state's model were not burdensome on school site administrators." (ACSA, 1989; Stephenson, 1989) ACSA lobbied the State Board of Education not to accept the original model and assembled its own group to develop an alternative. ACSA's guiding principles were to keep the report card simple, to focus on the assessment areas required by law, and to use data sources that are readily available to all schools. The primary audience for the report card should be the parents within a school's attendance boundaries. Ultimately, The State Board of Education accepted ACSA's model with some minor changes. (California Department of Education, 1989)

ACSA's model differed from the state's in using a simpler display of expenditures, requiring less detailed reporting of class sizes and teaching loads, allowing optional use of comparative data on student achievement, and permitting a simpler description of student support services. Both models cite the main objective as informing the local school community about conditions and progress being made at schools. ACSA goes on to state that "comparisons with other school sites or to statewide standards is permissible, but is not required." The relationship between the report cards, school accreditation reports, and program quality reviews is another area of difference. ACSA notes that the report cards are not intended to replace or duplicate accreditation reports or program reviews, and that data collection should rely on existing sources where possible. The Department of Education stresses that the report cards should be viewed as complementing, not duplicating these other assessments, that is, "review teams should use report cards as a source of valuable information and self-assessment, and report cards should draw important information about school conditions from the various reviews and reports."

The PTA, in a news release issued after the State Board adopted its model (California Congress of Parents, Teachers, and Students, Inc., 1989), stressed its view that the School Accountability Report Card should help parents to exercise their right "to be informed about their schools' conditions and progress." There should be a role for PTA units in collecting information for the Report Cards, in serving on relevant committees, and in sharing of the Report Card information through the sponsorship of public meetings. Although the California PTA did not issue its own model, it did propose that parents ask certain questions when reviewing the school accountability report card.

- If class size and teaching load are being reduced, how is this being done?
- If teachers are being assigned outside their subject areas, what are the reasons for their assignments?
- If the drop-out rate is being reduced, which methods for reduction are proving effective?
- If a classroom discipline policy is in effect at a school, what is the policy and how is it being enforced?
- If the school facilities have been determined to be safe and adequate, what were the criteria for that determination?
- As academic goals are being addressed, how is student achievement being measured in

reading, writing, arithmetic, etc.?

- As textbooks and other instructional materials support the school's instructional program, how current are the materials and are they in adequate supply?

Individuals voiced a number of issues about use of the school report cards. (Kossen,1989) A member of the State Board of Education asked whether schools might not use the report cards as a tool to tell the public what it wants to hear. An official with the California School Boards Association suggested that the report cards are missing the point. Parents really want to know how their individual children are performing, and the report cards will not provide this information. Also, the press is likely to look only at controversial areas and take information out of context. An ACSA spokesman noted that schools may be unfairly judged without considering socioeconomic and other background information. The PTA also expressed concerns that marketing campaigns designed to attract students could be confusing to parents.

The campaign for Proposition 98 and the development of the school report cards illustrates how different political coalitions--bureaucrats, administrators, teachers, and parents--can cooperate to obtain funding, but can also differ on the closely related issue of accountability. The model school report card was the result of numerous political compromises on a variety of both major and minor issues.

Performance Assessment

"Assessment, at all levels, is seen as the key strategy in bringing about significant educational reform. It gives educators more tools to evaluate the quality of learning-- and then make necessary adjustments. Moving beyond the standardized, multiple choice test as the primary accountability tool, the new approach to assessment helps educators measure what matters--including a student's ability to analyze, organize, interpret, explain, synthesize, evaluate, and communicate important experiences." (California Department of Education, 1992, p. 32)

The idea of using assessment to implement state education reform is relatively new. Cronbach (1984) omits education reform as a use of tests. In the mid 1980s Mitchell and Encarnation (1984) summarized the policy goals of testing and assessment to include student placement, program evaluation, and certification of competence. Policy makers had not yet made a strong connection between assessment and reforms in curriculum, instructional methods, or staff development. By the mid 1990s the terms of the debate had changed. Prominent educators (Tucker, M., Sizer, T., Resnick, L., and Anrig, G., 1992) viewed performance assessment, in harmony with curriculum and staff development, as a way to implement educational reform. The evolution of state assessment in California reflects this change in thinking.

The California Department of Education administered a program to assess student achievement for many years. During the 1970s CDE required schools to select a commercially available standardized achievement test, administer it, score the results, and submit a record of the distribution of the test scores to CDE. Unfortunately, CDE could not obtain a clear picture of student achievement statewide from the variety of tests in use. The legislature enacted the California Assessment Program (CAP) in order to provide a statewide picture. CAP relied on pools of 400 - 600 multiple choice test items that were distributed in a matrix fashion into 30 or 40 roughly similar short forms of the test. The test forms were spiraled into packets to insure that each form was taken by about the same of students at every school. Schools administered all three CAP reading, mathematics, and written assessment tests during one regular class period. The assessment produced accurate and reliable statewide estimates of achievement both in

general subjects and at a more detailed skill level. The accuracy and reliability of school results varied, depending on the size of the school, and the participation of students in the test administration. Because each student responded to only a few questions, there were no meaningful scores for individuals.

Initially, CAP produced an annual statewide report for the State Board of Education. Individual school reports were discretely shared with local superintendents. During the 1980s CAP assumed a new high-stakes role in school reform. CDE began by disseminating individual school results to most newspapers in California. In 1984 CDE made the CAP scores a centerpiece of the school performance reports. In 1989 the CAP results became a part of the Proposition 98 School Accountability Model Report Card. A widely publicized annual school recognition program identified schools that met test score goals. CAP's resultant high visibility may have contributed to its demise. California's Governor, just before retiring from public service, during a highly publicized feud with the State Superintendent, eliminated CAP funding in 1990 from the CDE budget.

Despite strained state finances, in 1991 California's Legislature, the new Governor, and CDE created the California Learning Assessment System (CLAS). (Intersegmental Coordinating Council, 1993). CLAS' designers wanted a multipurpose test that would "first and foremost" assess pupil progress, but would also measure program effectiveness, school and district effectiveness, and the condition of education statewide. Unlike CAP, CLAS' designers expected assessment to be an instrument for school reform, and wove the test into existing school programs and activities. The new statewide assessment program is "performance based," requiring students to write, speak, do research, work cooperatively, solve problems, create, and experiment. Teachers assist in the test construction, test scoring, and receive individual student scores. CDE made teacher designed and field tested tasks available for classroom instruction across the state as a "curriculum embedded assessment." CDE aligned CLAS is with statewide curriculum frameworks. The assessment system is consistent with strategic plans published in the CDE sponsored "grade span initiative" documents.

CLAS requires more time for administration than its forerunner. (CDE, 1994) The 1993 English Language Arts tests required one class period for each of three sections. Students responded independently to a reading passage on the first day. On the second day they worked in small collaborative groups to prepare for writing. On the final day, students independently wrote down their responses. The mathematics assessment required an additional class period during which students answered two open-ended problems and several enhanced multiple choice questions. The test development teams produced scoring guides that teachers used to evaluate students' work. Roughly two thousand teachers participated in the scoring as a staff development exercise at 34 regional sites. CDE views the training of teachers for scoring as a way to shape effective teaching practices.

CLAS produces a distribution of student scores, compared to CAP's subject area and skill scores. Teachers use performance standards with six levels to rate student responses. The performance standards and levels refer to CDE's curriculum frameworks. For example, in Reading a level six response "demonstrates insight," is "confident and willing to take risks," is "open to considering and developing new ideas," and "explores complexities in depth." A level one response "demonstrates understanding of only an individual word, phrase, or title," and does not "demonstrate understanding of the ideas or experiences offered or developed." The school report includes the percentage of students scoring at each level, as well as an average of the results obtained by 100 schools that are most similar socially and demographically.

Controversy over reports on school performance led the Superintendent for Public Instruction to appoint a committee to evaluate CLAS. The committee's report (Cronbach, Bradburn, and Horvitz; 1994) addressed operational problems relating to the management of the assessment, measurement difficulties, and touched upon political issues. Many of the problems

arose from the newness of CLAS, and were not anticipated. Considered as a field trial, CLAS succeeded because it uncovered these problems and is overcoming them.

The performance assessment aspects of CLAS posed operational requirements that were much greater and more complex than those associated with traditional multiple choice testing. Test development, administration, scoring, and reporting all involve more material, are more time consuming, and require more careful coordination. The limited budget, considering the magnitude of the task, and the short time frame contributed to the operational difficulty.

Some of the measurement problems related to the requirement that CLAS report on school as well as student performance. A design to assess schools includes many different test forms and numerous assessment tasks. The time available for testing individual students is limited, and the luck of the draw determines whether a student receives an easy test form, or a hard one. Fairness in student assessment requires allowing for unequal opportunity to learn. Despite a clearly stated preference in the law for individual student reports, the limited funding available permitted scoring only a fraction of the students' tests. The relatively small number of performance tasks on each test form resulted in unreliable student scores. Deficiencies in the sampling and scoring also resulted in unacceptably imprecise school results for a number of schools. Apparently the scoring rules for the mathematics test were much more rigorous than the other tests. The low mathematics results for most schools may have contributed to the misperception of a crisis in California's mathematics instruction.

The committee's report noted that a number of assessment tasks became controversial when they became public. These tasks required students to express personal opinions about their families. Some of the written prompts referred to sensitive racial stereotypes. CDE's reluctance to public disclosure of the test in the face of the controversy, appeared to cause more dissent and resistance. Cronbach et al. (1994, p. 30) noted that the "complaints have reminded Californians of the ill-defined border between political sensitivity and censorship. The dilemma is to be resolved by political mechanisms, not by measurement professionals." The dissent relates to curriculum goals, not to the fidelity of the test to California's adopted curriculum frameworks. While this dissent is not primarily an operational or measurement problem, some parents and school boards resisted the requirement to test students. The resistance created non-response that may have contributed to measurement and operational difficulties. More significantly, the dissent can eroded political support for CLAS.

CLAS has provoked more lively controversy than California's earlier state assessment programs ever did. Some of this controversy relates to technical measurement aspects (Linn, 1993) but other aspects are cultural and may be more difficult to address. Mitchell and Encarnation (1984) describe the development and selection of curriculum materials as a "legal and moral issue," not just a matter of professional judgment. Parents will naturally have interest in CLAS as an assessment that is designed "first and foremost" to assess individual pupil progress. CLAS' other goal, to measure the effectiveness of statewide reforms, will then focus the attention of those parents on the purposes and values that underlie the reforms. Perhaps unavoidably, CLAS is caught in the same legal and moral controversies that crop up around curriculum and staff development. If this analysis is accurate, the context and the process for designing CLAS will sharpen the conflicts. First, CLAS is intentionally highly visible as the centerpiece of California's accountability programs. Moreover, those political and cultural groups that are interested in curriculum and in the treatment of their children, but with no role in test development, will be dissatisfied and will continue to lobby for their interests.

Conclusions

Practicing educators who work in schools, unions, or associations sometimes comment that performance reports do not work. (Fetler, 1990) These educators do not see improved

academic performance, the reports are costly and time consuming, and they cite unintended side effects, e.g., narrowed curriculum, poor teacher morale, irregular testing practices, and biased data gathering. While these perceptions are worthy of study, they are not an evaluation of performance report effectiveness. Such an evaluation should include not only the espoused goals of performance reports (outcomes and restructuring), but also the tacit political and cultural goals.

During the 1980s and 90s the California Department of Education intensely pursued its own bureaucratic goals of a more academic curriculum, more rigorous instruction, and higher student achievement. CDE's strategy was to align its assessment and performance report programs with its goals. Schools and professional organizations were in a sense subject to these programs and had little to say about their operation. At the same time, the assessment and performance report programs were a part of CDE's advocacy campaign for K- 12 with the legislature and in the development of the state's budget. Although many districts disagreed with CDE's policy of school accountability, there was even more intense conflict over the budget with non-education interests, e.g., the justice system, social services, and health services. When the Governor abolished the assessment program, CDE's ability to negotiate grew weaker.

The California Teachers Association and the Association of California School Administrators are two organizations that have influence on legislation and during elections. CTA and ACSA worked to pass Proposition 98 as a guarantee of school funding, and used school accountability report cards to persuade the public. CTA and ACSA negotiated with the legislature on the process for drawing up the report cards and arguably had dominant roles in their design. The State Board rejected the CDE model in favor of ACSA's proposal. CDE has little to say in the administration of the report cards. Local schools have primary authority over production and dissemination.

Some educators consider performance reports to be an alternative to regulation. For example, in exchange for relief from prescriptive regulations, schools agree to disclose how well students perform and to meet specified standards. While this exchange is difficult to transact with established regulations, it may work in agreements for new funding, e.g., the funding guarantees in Proposition 98 and the school accountability report cards. New funding obtained through Proposition 98 is to be used for school improvement, including salaries. Without the report cards, it is possible that government would more narrowly prescribe the use of new funds.

The effectiveness of performance reports for school restructuring and improving student outcomes is poorly understood. Performance reports work politically in the sense that state and federal government continue to devote time and resources to their design and implementation. The old symbols of accountability and cudgels in budget battles--program regulation and compliance reviews--have become somewhat tarnished over time. New symbols of accountability-- performance reports, standards, and incentives--are gaining in favor.

References

Association of California School Administrators. (1989). Model accountability report card. Sacramento: Author.

Brown, P. (1990). Accountability in public education. Far West Laboratory Policy Brief Number Fourteen. San Francisco: Far West Laboratory.

Bolman, L. and Deal T. (1991). *Reframing Organizations*. San Francisco: Jossey-Bass.

California Department of Education. (1992). Second to None. A Vision of the New California High School. Sacramento: Author, pp. 35-35.

- California Department of Education. (1994). Statewide student assessment results released. (Press release). Sacramento: author.
- California State Department of Education. (1989). Technical assistance manual for the California model school accountability report card. Sacramento: author.
- Consumer Reports. (1994). How Consumer Reports tests. Taking a car's measure on and off the track. *Consumers Union*. 59 (4), pp. 220-222.
- Center for Research on Evaluation, Standards, and Student Testing. (1989). Accountability in the Post-Charlottesville Era. Los Angeles: University of California Center for the Study of Evaluation.
- Council of Chief State School Officers. (1987). Survey of the 50 State Performance Accountability Systems. Washington DC.: Author.
- Cronbach, L. (1984) *Essentials of Psychological Testing*. Fourth Edition. New York: Harper & Row.
- Cronbach, L., Bradburn, N., and Horvitz, D. (1994). Sampling and Statistical Procedures Used in the California Learning Assessment System. Sacramento: California Department of Education.
- Fetler, M. (1986). Accountability in California public schools. *Educational Evaluation and Policy Analysis*. 8 (1), pp. 31-44.
- Fetler, M. (1990). California's mandate for school accountability. Boston: American Educational Research Association Annual Meeting.
- Fetler, M. (1991). A method for the construction of differentiated school norms. *Applied Measurement in Education*, 4 (1), pp. 53-66.
- Galbraith, J. (1983). *The Anatomy of Power*. Chapter XV. Organization and the State. Boston: Houghton Mifflin Company, pp. 144 - 159.
- Intersegmental Coordinating Council. (1993). K Through 12 School Reform. Implications and Responsibilities for Higher Education. Sacramento: author. pp. 38-40.
- Kirst, M. (1990). Education: The solution to California's problems. *California Journal*. January, 1990. 49-51.
- Kossen, B. (1989). Grading the schools. *Golden State Report*. August, 1989, 25-28.
- Linn, R. (1993). Educational assessment: Expanded expectations and challenges. *Educational Evaluation and Policy Analysis*. 15 (1). pp. 1-16.
- Mitchell, D. and Encarnation D. (1984). Alternative state policy mechanisms for influencing school performance. *Educational Researcher* 13 (5), pp. 4-11.
- Murnane, R. (1987). Improving education indicators and economic indicators: The same problems? *Educational Evaluation and Policy Analysis*. 9 (2), pp. 101-116.
- National Center for Education Statistics. (1991). Education Counts. An Indicator System to Monitor the Nation's Educational Health. Washington, DC: Author.

- Noble, A. and Smith, Mary Lee. (1994). Old and new beliefs about measurement-driven reform: "Build it and they will come." *Educational Policy*, Vol. 8, No. 2, pp. 111-136.
- Oakes, J. (1989). What educational indicators? The case for assessing the school context. *Educational Evaluation and Policy Analysis*, 11, pp. 181-199.
- Office of Educational Research and Improvement. (1988). Creating Responsible and Responsive Accountability Systems: Report of the OERI State Accountability Study Group. Washington, DC.: U. S. Department of Education.
- Policy Analysis for California Education. (1987). Conditions of Education in California. Berkeley, California: University of California School of Education. 45-70.
- Porter, A. (1991). Creating a system of school process indicators. *Educational Evaluation and Policy Analysis*. Spring, 1991, 13 (1), pp. 13-29.
- PTA California Congress of Parents, Teachers, and Students, Inc. (1989). School Accountability Report Card: PTA urges parents to become involved. Los Angeles: Author.
- Rothman, R. (1993). Assessment Questions: Equity Answers. Proceedings of the 1993 CRESST Conference.
- Rudner, L. and Boston, C. (1994) Performance assessment. *The Eric Review*. 3 (1) 1994, 2-12.
- Salganik, L. (1994). Apples and apples: Comparing performance indicators for places with similar demographic characteristics. *Educational Evaluation and Policy Analysis*. 16 (2) pp. 125-141.
- Sirotnik, K. and Oakes, J. (1990). Evaluation as critical inquiry: School improvement as a case in point. *New Directions for Program Evaluation*, No. 45, Spring. pp. 37-59.
- Stephenson, A. (1989). School accountability report cards: The principal's role. *Thrust*. September. 27-31
- Thurow, L. (1981). *The Zero-Sum Society*. Distribution and the Possibilities for Economic Change. Spreading Rules and Regulations. New York: Penguin Books. pp. 122-154.
- Tucker, M., Sizer, T., Resnick, L., and Anrig, G. (1992). By all measures: The debate over standards and assessments. *Education Week*, June 17, 1992
- Wainer, H. (1994). On the academic performance of New Jersey's public school children: Fourth and eighth grade mathematics in 1992. *Education Policy Analysis Archives*. Available URL:<http://olam.ed.asu.edu/epaa/>.
- Wohlstetter, P. (1991). Accountability Mechanisms for State Education Reform: Some Organizational Alternatives. *Educational Evaluation and Policy Analysis*. Vol. 13, (1), pp. 31-48

Copyright 1994 by the *Education Policy Analysis Archives*

EPAA can be accessed either by visiting one of its several archived forms or by subscribing to the

LISTSERV known as EPAA at LISTSERV@asu.edu. (To subscribe, send an email letter to LISTSERV@asu.edu whose sole contents are SUB EPAA your-name.) As articles are published by the *Archives*, they are sent immediately to the EPAA subscribers and simultaneously archived in three forms. Articles are archived on *EPAA* as individual files under the name of the author and the Volume and article number. For example, the article by Stephen Kemmis in Volume 1, Number 1 of the *Archives* can be retrieved by sending an e-mail letter to LISTSERV@asu.edu and making the single line in the letter read GET KEMMIS V1N1 F=MAIL. For a table of contents of the entire ARCHIVES, send the following e-mail message to LISTSERV@asu.edu: INDEX EPAA F=MAIL, that is, send an e-mail letter and make its single line read INDEX EPAA F=MAIL.

The World Wide Web address for the *Education Policy Analysis Archives* is <http://olam.ed.asu.edu/epaa>

To receive a publication guide for submitting articles, see the *EPAA* World Wide Web site or send an e-mail letter to LISTSERV@asu.edu and include the single line GET EPAA PUBGUIDE F=MAIL. It will be sent to you by return e-mail. General questions about appropriateness of topics or particular articles may be addressed to the Editor, Gene V Glass, Glass@asu.edu or reach him at College of Education, Arizona State University, Tempe, AZ 85287-2411. (602-965-2692)

Editorial Board

John Covalesskie Syracuse University	Andrew Coulson
Alan Davis University of Colorado--Denver	Mark E. Fetler mfetler@ctc.ca.gov
Thomas F. Green Syracuse University	Alison I. Griffith agriffith@edu.yorku.ca
Arlen Gullickson gullickson@gw.wmich.edu	Ernest R. House ernie.house@colorado.edu
Aimee Howley ess016@marshall.wvnet.edu	Craig B. Howley u56e3@wvnm.bitnet
William Hunter hunter@acs.ucalgary.ca	Richard M. Jaeger rmjaeger@iris.uncg.edu
Benjamin Levin levin@ccu.umanitoba.ca	Thomas Mauhs-Pugh thomas.mauhs-pugh@dartmouth.edu
Dewayne Matthews dm@wiche.edu	Mary P. McKeown iadmpp@asvm.inre.asu.edu
Les McLean lmclean@oise.on.ca	Susan Bobbitt Nolen sunolen@u.washington.edu
Anne L. Pemberton apembert@pen.k12.va.us	Hugh G. Petrie prohugh@ubvms.cc.buffalo.edu
Richard C. Richardson richard.richardson@asu.edu	Anthony G. Rud Jr. rud@purdue.edu
Dennis Sayers dmsayers@ucdavis.edu	Jay Scribner jayscrib@tenet.edu
Robert Stonehill rstonehi@inet.ed.gov	Robert T. Stout aorxs@asvm.inre.asu.edu