The Best Path to Systemic Educational Policy: Standard/Centralized or Differentiated/Decentralized? 

William H. Clune
University of Wisconsin–Madison

Systemic educational policy recommends creating more ambitious instruction in schools through a centralized strategy of mandatory curriculum frameworks, high-stakes student assessments, and coordinated teacher training. This article suggests serious problems with such a strategy and recommends a decentralized alternative involving local choice of curricula, technical assistance, and professional development. Arguments for, problems with, and alternatives to the centralized strategy are examined in eight dimensions of educational policy: curriculum development, curriculum stratification, educational indicators, high-stakes examinations for students, school improvement, educational finance, educational governance, and teacher training. The article concludes with recommendations for policy and research.

Systemic educational policy (Smith & O'Day, 1990) rests on five basic insights: (a) Curriculum in the United States is poor by almost any standard (Baker, 1993); (b) a more ambitious curriculum would have a strong, positive impact on the achievement of most students; (c) some combination of coherent signals and capacity-building from outside schools is needed to encourage and sustain positive curricula within schools; (d) existing policies push the curriculum toward mediocrity or in fragmented, contradictory directions; and (e) well-documented and officially accepted high minimum goals for student achievement are useful as a standard for establishing the resources needed in high poverty schools (Clune, in press; O'Day & Smith, 1993).

The policy implication usually drawn from these premises is that we should move toward something resembling the national curricula and educational ministries of countries that have high levels of student achievement. Besides a national (or perhaps state-by-state) curriculum, we would adopt high-stakes examinations (e.g., for entrance into higher levels of academic education) and teacher training organized around the national curriculum (Smith, O'Day, & Cohen, 1991). Problems with this vision of change have been raised from many directions: for example, on grounds of curriculum diversity, school capacity, and test validity (Darling-Hammond, 1991; Education Week, 1992; Jaeger, 1991; Koretz, Madaus, Haertel, & Beaton, 1992; Lieberman, 1991; Madaus, 1991; McLaughlin, 1991; Shepard, 1991; Stake, 1991). And proposals for change have been modified to reflect such criticisms—a “core curriculum” with local modifications, “delivery standards” to raise school capacity, and new kinds of high-stakes tests that supposedly would not have a distorting effect on teaching and learning (O'Day & Smith, 1993).

In this article, I am making a different point: The basic means-ends analysis of systemic educational policy may simply be mistaken in light of current social, political, and educational conditions in the United States. I agree with the five premises of systemic educational policy but think that a mandatory system of strong instructional guidance from the state has four basic problems:
1. **Standard curriculum, diverse needs.** State adoption of a standard curriculum is not required to make ambitious curricula widely available to schools and teachers. Acceptable models already exist, and more will regularly become available through multiple channels. Further, if the effort to standardize curricula succeeds in the face of strenuous resistance, it will inevitably block a vital process of matching curricula with local needs and diverse student populations.

2. **Absence of attention to delivery.** Whatever the curriculum, defining new goals for instruction is the easy part. The real challenges are at the delivery stage (extra resources used effectively in low-income schools, professional development, technical assistance, more cost-effective governance, and school improvement). Yet the official enactment of high standards may distract attention from the problems of delivery and be confounded with real help for poor schools and real gains in student achievement.

3. **Problems with high-stakes exams.** In the absence of a realistic delivery structure, state-centered instructional guidance will probably rely on high-stakes examinations of students to force change in schools and classrooms. But the motivational power of high stakes automatically creates a vast, probably oppressive regime of teaching to the test, requires a highly prescriptive curriculum, creates misleading grade inflation (Linn, 1993), is inconsistent with the new constructivist and contextualized goal of teaching for understanding (Cohen & Barnes, 1993b), and may well become transformed into an official system of student stratification (since the passing grade, or performance standard, must be set at realistic levels for the full-time effort of lower achieving students). Further, a high-stakes examination will almost certainly provoke litigation over opportunity to learn, thrusting the difficult delivery question into the awkward policy-making forum of the courts.

4. **Incompatibility with U.S. educational governance.** Adoption of a standard curriculum by the national government and 50 states is difficult enough (e.g., states tend to create multiple kinds of new standards and simply add them to the old ones), but the adoption problem is dwarfed by the implementation problem. Whatever is adopted at the top is filtered unpredictably through multiple intermediate layers of discretion shaped by preexisting cultures and agendas (Cohen & Spillane, 1992). The possibility of a prescriptive yet unpredictable system is difficult to imagine, but is the likely outcome of a standardized system. This governance problem is really the delivery problem in disguise, since the only way to escape the unmanageability of top-down regulation in a fragmented governance system is to create customized coherence at the level of local districts and schools.

Thus, the challenge is to design policies that combine the high standards of systemic policy with a broad diversity of curricular options and a powerful local delivery system. To frame the question in a specific way, one might imagine a choice confronting educational policymakers in California. To date, California has developed curriculum frameworks, relied on professional development and school improvement, and avoided high-stakes exams for students. The question considered here is, in which of two different directions should a state such as California proceed? The centralizing route would make it clear that the frameworks are mandatory (i.e., prohibit alternative approaches), incorporate the frameworks into a set of high-stakes student examinations, and centrally coordinate a statewide system of teacher training around the frameworks. The decentralizing option would move in the opposite direction: offering the state frameworks as one possibility but approving other options, avoiding high-stakes exams, and relying primarily on school improvement and professional development.

However, anyone who asserts that curriculum upgrading can be achieved in a differentiated, decentralized world must face the argument that differentiation is the problem rather than the solution, in effect, that coherent, centralized policy is necessary to overcome a highly differentiated social environment, especially on behalf of the disadvantaged. Again, the question is how to get coherence without centralization.

One problem with answering this question is how to analyze centralization, a complex...
and multifaceted issue (Clune & Witte, 1990). The case for systemic policy is also complex, identifying multiple problems, diagnosing causes, and suggesting remedies. In order to unpack this complexity, I break the discussion down into eight components of educational policy: curriculum development, curriculum stratification, educational indicators, high-stakes examinations for students, school improvement, educational finance, educational governance, and teacher training. Then, for each component, I consider the arguments for centralization, problems with centralization, and alternatives. From this discussion of the alternatives within each area, the outlines of a suggested alternative system emerge. To summarize and elaborate those suggestions, the eight sections are followed by ones on the policy and research agendas that emerge from a more decentralized vision and, finally, a conclusion reflecting on the evolution of systemic policy.

Curriculum Development

The case for centralization. What I mean by curriculum development is the availability of high quality, ambitious curricula on a widespread basis in schools. Those who advocate, standardization consider it necessary for curriculum development for three reasons: school level coherence, comparative advantage, and equity. School level coherence refers to the idea that only a regionwide effort can harmonize all the fragmented and conflicting influences on school curriculum that now emanate from multiple policy sources (e.g., standardized testing, Chapter 1 requirements, commercial textbooks, various kinds of teacher training). Comparative advantage is the idea that curriculum development is a job for experts and cannot be managed by ordinary teachers and administrators on a school-by-school basis (Clune, 1987). Equity is the idea that a high quality curriculum is a resource especially for low-income schools, or more generally that any system that leaves power in local schools will result in progress for high-capacity schools and stagnation for those with low resources and many problems (O'Day & Smith, 1993).

Problems with centralization. Whatever the merits of these arguments, a centralized curriculum poses an enormous problem in achieving social and political consensus. It is one thing to agree on “higher standards” and a “more ambitious curriculum,” but serious arguments exist about the specific meaning of these phrases. Old, unresolved conflicts between traditional and progressive visions of the curriculum and new conflicts around multiculturalism abound (Newmann, 1993). We have seen a long period of differentiation of the curriculum, first into academic, general, and vocational education, more recently with the addition of special education, gifted and talented, compensatory, and bilingual education (Oakes, Gamoran, & Page, 1992). Many of these curricula are embedded in separate institutional patterns—special classes and schools, specially trained teachers, distinctive academic careers, and special interest groups. At the same time that we hear calls for a common academic curriculum, an equally strong call can be heard for much sharper differentiation of the secondary curriculum into academic and vocational career tracks. The widespread movement toward restructured schools seems to intensify all the other forces leaning toward diversity, because restructuring schools tend to select and adapt existing curriculum models according to the perceived needs of the schools’ own staff and students (Fuhrman & Elmore, 1992; Levin, 1988).

Furthermore, it is not generally well understood that the new curriculum models look in very different directions and raise many unresolved conflicts (e.g., see the various approaches in New American Schools Development Corporation schools, 1992). Perhaps the greatest is a battle just surfacing between the modern versions of traditional and progressive education. For example, while some groups push for more meaningful applications in mathematics (National Council of Teachers of Mathematics, 1989), other researchers emphasize the importance of fluency in traditional calculations (Geary, Fan, & Bow-Thomas, 1992). Not everyone warms to the goal of “higher order thinking”; the phrase is a red flag to traditionalists. Exactly what is “higher” and how to reach it are legitimate matters of discussion, a discussion that is just beginning rather than approach-
ing consensus. Divisions and much uncertainty also exist about how much acceleration of learning is possible within any particular model for various groups of students—college-prep students, low-income students, uneducated, non-English speaking immigrants, and so forth. For example, research on the effects of curriculum apparently does not support transplanting fast-track academic instruction directly to low-income students; adaptations are required in pacing, attitude, and individual assistance (Oakes, Gamoran & Page, 1992). And, of course, many scholars and practitioners vehemently reject the academic model for these and other students, preferring applied education, which is more motivating and realistic and, thus, in the view of its supporters, better preparation for both higher education and the workplace (Shanker, 1990; Wehlage, Rutter, Smith, Lesko, & Fernandez, 1989).

Divisions also exist within the more applied school; there are the academically oriented applied visions like NCTM math and 2061 in science (NCTM, 1989; Rutherford & Ahlgren, 1990), and there are the vocationally oriented visions embracing a much wider range of skills and attributes, such as the SCANS report, which includes communication and leadership as well as generic skills like gathering and analyzing data (Secretary’s Commission on Achieving Necessary Skills, 1991). Proclamations about higher order thinking frequently embrace a wide spectrum of specific curricula (SCANS, 1991; University of Pittsburgh, 1992), but this commendable tolerance implies that the definition of higher order thinking is broader than any possible particular set of standardized frameworks. Perhaps the debate is so wide-ranging and multifaceted because of the uncertainty and disagreement about the purposes of schooling and how best to achieve them. Ideas about how to prepare for higher education or the workplace and conceptions of the skills relevant to each goal are both myriad and dynamic (Murnane & Levy, in press).

To be certain, this is not the same as saying that we are in a period of utter confusion and chaos. There is a rather exciting consensus about the need for and feasibility of a much more ambitious and interesting curriculum. Powerful pockets of consensus around improved curricula are emerging all over the place and are being implemented in many schools. But all this creativity also creates diversity. In fact we are just as likely to be witnessing another stage of profound curriculum differentiation as we are a trend toward consensus.

Alternatives to centralization. In light of all these problems with a common curriculum, the arguments in favor of centralization should be examined quite closely, and each seems to crumble on closer scrutiny and to open the door to consideration of a more decentralized system.

School level coherence involves two components: the removal of fragmenting forces (a kind of deregulation, if you will) and the delivery of a powerful, coherent curriculum to the school level. The evidence is quite clear that schools that have the will and capacity to implement a new curriculum are able to implement almost anything they choose even within the existing framework of unfavorable regulation (Levin, 1988; Sizer, 1989). Such discretion probably flows from the fact that the regulatory forces are relatively weak and conflicting, more of an annoyance and a transaction cost than an outright roadblock (Cohen & Spillane, 1992). On the other hand, the task of deregulation and change can proceed independently of curriculum-building at the school level.

Thus, the door seems to be open for the activity of the many networks of curriculum upgrading that are now flowering throughout the country (Lieberman & McLaughlin, 1992). Unlike a standardized curriculum, which is vague about the link between policy and practice, such networks focus directly on the process of instructional change in schools. Indeed, a strong standardized curriculum could be the one type of regulation that prevents creative change in high capacity systems without offering higher capacity for those unable to change (a different and disturbing perspective on the idea of regulation as a “boundary condition” for creative school restructuring) (Hannaway, 1993). For example, the clearest example of systemic policy in the United States, the New York State Re-
gents, developed two sets of exams, the regular Regents and the Regents competency exams (Tyree, 1993). This dual system is currently under criticism and study (New York State Curriculum and Assessment Council, 1992). On the other hand, even the regular Regents is rejected as too easy by high-status, Ivy-oriented, private prep schools (Arthur Powell, personal communication, Consortium for Policy Research in Education meeting held in Lumberville, New York, July 11–12, 1991), while several other districts in New York are seeking waivers from the Regents system to implement a different kind of math curriculum emphasizing applications.

Paradoxically, the argument for centralization from comparative advantage loses most of its force because of the success of the movement toward a more ambitious curriculum. There seems to be an ever-growing capacity in both state and nonstate enterprises to develop and deliver complete packages of curriculum and instructional change. This is true not simply for the networks, such as Slavin schools, Levin schools, OBE schools, and the College Board’s Equity 2000, but also for the curriculum control systems of many high-capacity school districts (Hanna-way, 1993). The growing capacity of the decentralized sector is in marked contrast to persistent and worsening problems in the state sector as seen, for example, in the lack of any demonstrated capacity in departments of education to operate as agents of instructional change because of recent drastic budget cuts. From the perspective of a theory of government, the time could hardly be worse for adoption of a standard curriculum. Social and scholarly consensus seems to be developing that effective central regulation of decentralized institutions is difficult at the same time that an explosion of capacity occurs outside the state. In light of the recent growth of nonstate capacity, I would question the relevance of earlier social science research on the difficulty of sustaining positive improvements at particular schools (Rivlin & Timpane, 1975). In a sense, this evidence showed only the obvious, the challenge of organizational change on a wide scale, and offers no support for the state as change agent. And the recent success of nonstate agents suggests that the optimum role for the state is probably something more like encouraging and sustaining the capacity of decentralized curriculum networks than assuming the task of curriculum change.

My sense is that the equity argument for a common curriculum also is falling apart. Low-income schools and students do not need a curriculum that is identical to a curriculum somewhere else in the state or country; they need a curriculum that is well adapted to produce dramatic gains in learning for their particular students (Darling-Hammond, 1992). Thus, the effort to forge a common curriculum for all students can only distract from the development of curricula well adapted for low-income schools, producing, for example, a kind of hybrid compromise of college prep and other goals. One of the problems with the New Jersey school finance case was setting spending in wealthy suburban districts as the baseline for effective programs in poor urban districts (Abbott v. Burke, 1990). A better approach would have been to establish the cost of specified instructional programs capable of reaching high minimum outcomes actually adopted by poorer districts (Clune, in press). Low-income schools desperately need serious attention devoted to problems of low capacity and inadequate resources, a type of response that may actually be deflected by the illusion that a common curriculum is a sufficient stimulus for change. In other words, a common curriculum seems to offer low-income schools little or nothing compared with a strategy (if we can imagine it for a moment) of subsidizing powerful curriculum development specifically intended for each school, backed up by adequate resources (implying a new conception of school finance, discussed below).

The problem of curriculum stratification

Argument for centralization. Few would disagree that the curriculum actually delivered to lower track students is in most cases boring, unambitious, and uninspiring. One logical response, would be to set a common standard. Thus, for systemic advocates, the differentiated curriculum is the problem, and the common curriculum is the solution. For
example, as a means of “improving the overall quality of schooling for all children.” Smith and O’Day recommend developing a common curriculum that then can be employed to expose “differences in appropriateness and quality of the curriculum and programs offered to different groups of students” (O’Day & Smith, 1993).

Problems with centralization. There are two kinds of curriculum differentiation with quite different implications for policy: the practice of giving a different curriculum to different groups of students and the practice of giving a clearly inferior curriculum to some groups of students (Oakes, Gamoran, & Page, 1992). While an inferior curriculum retards student learning (Bryk, Lee, & Smith, 1990), there appears to be no empirical basis (and certainly not a strong case) for believing that even within a single classroom, the most traditional form of curriculum differentiation (grouping within subjects) hurts student achievement (Slavin, 1987a, 1987b, 1990). The key question in the United States is whether to change both aspects (the common curriculum) or only the second (moving to more ambitious content where it is needed).

The problems with abolishing all stratification seem overwhelming. The range of student achievement in the United States with its incredibly diverse population is so vast both within and across schools that a common curriculum (in the sense of students’ reaching the same level of instruction and proficiency at the same age) seems patently impossible. It is also likely either to produce a huge mismatch of content and ability or to be a common curriculum in name only. This difficulty is sometimes masked by saying that all students should have the same content standards. We have already seen how agreement on content standards more specific than the general idea of higher order thinking is unlikely because of disagreement about the ends of the curriculum. But the real problem will come with the performance standards and pacing of instruction. Even within common content standards, we can count on the fact that some groups of students will reach more topics and higher levels of proficiency faster than others. Under these circumstances, the important point for disadvantaged students is to choose a goal that offers genuine access to society and to higher levels of instruction (e.g., getting to calculus a little later, but getting there).

Alternatives to centralization. The alternative to a common curriculum is achieving curriculum upgrading through the differentiated curriculum. Curriculum upgrading would be conceived as a series of projects each adapted for (and chosen by) a different sector (e.g., academic, vocational), group of schools or students (low-income, fast-track academic), or groups with different curriculum preferences (traditional, progressive, applied, multicultural).

This flexible approach to curriculum upgrading can be well illustrated with low-income students, who suffer the most under curriculum differentiation, not because of a different curriculum per se but because of an inferior one. In a paper on program equity in school finance (Clune, in press), I suggest that curriculum standards for such schools should be operationalized in terms of three standards:

1. Content. Because all children can learn all types of skills, the curriculum options for low-income students should be the same as all other students.
2. Performance. All students other than the 2% or so truly disabled should score at least at the proficient level on new tests including higher order thinking (and, of course, many students from all groups should score much higher—the minimum is not the maximum).
3. Certification. Every student should be qualified and certified for college entrance by, for example, having taken the required courses and being able to pass a college entrance exam (College Board, 1991).

While these are very ambitious standards for high-poverty schools, they are not the same standards that would be adopted by privileged prep schools or fast-track suburban schools, where standards might be that almost all students gain access to highly competitive colleges and universities. But the realistic specification of high minimum performance standards offers enormously more to low-income schools than the pretext of equality under common content standards. “Get-
ting serious” about performance standards permits local schools and communities to identify the instructional resources needed to reach such standards and thus to make a powerful claim for much higher levels of per pupil spending. On the other hand, failure to specify standards could easily lead to an officially approved system of performance stratification. For example, the path of least resistance on a high-stakes examination is to set the passing rate at a point producing a tolerable level of failures among the lowest achieving students.

The idea that a common curriculum for all is necessary to protect poor students is ultimately paternalistic and self-defeating. Given information, a choice, and adequate resources, what local community or school would intentionally choose an inferior curriculum for its students? On the other hand, the larger society has every incentive to withhold the resources necessary to produce high minimum levels of performance.

Student Assessments as Educational Indicators

Argument for centralization. Student assessments are needed as part of the system of educational indicators for planning and evaluation. The “condition of education” (U.S. Department of Education, 1992) is really a snapshot of students: how they enter the system, what they experience in the system, what they learn, and what happens when they leave. As for evaluation, great changes may occur in the rhetoric and process of reform at every level without any assurance of changes in important outcomes such as student achievement. Some kind of standardized assessment is extremely useful to compare the achievement of different groups of students in different geographical areas, schools, and conditions of education (Finn, 1991).

But systemic policy is not satisfied with centralized measurement of a core of student achievement in diverse curricula. Generic tests are said to be insensitive to most student achievement and drag the curriculum down to the lowest common denominator of basic skills (Smith & O’Day, 1990). Thus, the specific argument for alignment of assessments and curriculum is the need to assess whether the students have learned exactly what they have been taught. General assessments try to measure proficiency in skills that cut across curricula: reading comprehension, logical thinking, writing, general mathematical reasoning. But such tests completely ignore more specific content: trigonometry, history of the American Civil War, Spanish language, Hamlet. The only way to measure learning in such subjects is through an assessment tied to curriculum, whether this is the teacher-made classroom examination, an AP test, a university test for advanced placement, or a national or state exam based on a standard curriculum (e.g., the New York State Regents). The argument for a centralized, content-specific exam is the need for centralized measurements of this kind of specific subject matter. Put another way, we should standardize education in order to measure it. And it is true that only states with curriculum mandates can develop standardized measurements of how much their students know about specific subjects.

Problems with centralization. There is no insurmountable technical problem with preparing a common student examination reflecting a national or state curriculum. The existing NAEP could easily evolve into an exam for individual students (Smith, O’Day, & Cohen, 1991), and New York State Regents exams have been given for many decades. The problem is really with standardizing the curriculum itself, and with high stakes attached to such exams. In fact, the relative ease of developing such exams (as opposed to improving instruction) is one reason they are attractive to policymakers. Thus, the key question about student assessments is whether we can get the advantages of objective data on student learning of ambitious curricula without a single, centralized curriculum.

Alternatives to centralization—a split-level indicator system. We already have a system for measuring the general level of achievement among U.S. students, the National Assessment of Educational Progress. NAEP is being upgraded to include new content (Silver, Kenney, & Salmon-Cox, 1992) and is being administered to provide state-by-state data (Koretz, 1991). Various states also have student assessments. Thus, the unmet need is
for decentralized measurement of specific content.

In fact, at least two kinds of decentralized, standardized assessments are readily available. One is the standardized but optional curriculum such as the Advanced Placement (AP). Scores on AP subjects give colleges and universities a precise assessment of students' mastery of the AP curriculum. AP scores can also be used to measure learning in a population: for example, the number of students taking and passing AP exams. The College Board is now expanding this kind of system to a broader range of students with its Pacesetter exam, a system of standardized examinations developed by the same organization that creates the AP exams but aimed at a broader range of students (DeWitt, 1992; Pitsch, 1992). Another feasible system is to use samples of instructional products for assessment: student papers, experiments, oral presentations, and so forth. This kind of system is easiest to imagine as part of a certification process in which inspectors ascertain how faithfully a particular school has implemented a particular curriculum by looking at instructional materials, process, and products. I will sketch how such a system could work later in the article (see also the evaluation system recommended in NCTM evaluation standards 11–14; NCTM, 1989). Broad-scale inspection of this kind has the advantage of measuring process and outcome goals at the same time (Porter, 1991).

Of course, the more optional curricula that exist, the more difficult it will be to compare all students, regions, and schools with each other. But the population assessment (e.g., NAEP) gives us some information, and lack of comparability becomes less of a problem to the extent that we have confidence in the curriculum. A strong system of inspecting instructional practice in schools should be far more informative than scores on tests alone (which are influenced by a variety of factors other than instruction). Consider, for example, the confidence we place in simple grades at a strong academic institution—A's in science at MIT. Widespread adoption of voluntary high-stakes exams would also add to our confidence in improved instruction. For example, we do not really care if the student takes an AP exam or an International Baccalaureate. Some foreign countries that use a standard curriculum and high-stakes entrance examinations actually have a number of different curricula for different states, provinces, university entrance systems, and so forth (Koretz, Madaus, Haertel, & Beaton, 1992). In each case, we can tell if the student has reached a high level of achievement. At the same time, it would be helpful to have some method of comparing all students to check the effectiveness of the different curricular systems and make other judgments useful for educational evaluation and planning. Standardized entrance exams such as SAT, ACT, and GRE are useful.

All of this suggests the virtues and feasibility of a split-level system of indicators: a regional (national, state, etc.) low-stakes assessment, like the NAEP, coupled with various methods of measuring and communicating student mastery of ambitious curricula chosen at the local level. Later in this article, I describe a system of school inspection of instructional practice that would be strengthened and supplemented by voluntarily adopted high-stakes exams.

**High-Stakes Student Examinations**

**Argument for centralization.** It may not be immediately obvious, but high-stakes student examinations are a key component, perhaps the cornerstone, of the centralized version of systemic educational policy. Without some incentive for performance, the top-down qualities of systemic policy become an extreme liability; schools, teachers, and students might not care enough about performance on the exams to make difficult and often disagreeable changes in the content and pedagogy of courses. But with such powerful consequences as grade promotion, graduation, college entrance, and employment, the evidence is quite clear that teachers will teach to the test and students will be motivated (to the extent that they can be motivated by any eligibility incentive). Students in other countries with standard curricula and high-stakes exams work very hard, perhaps too hard, to make the grade (Nordquist, 1993). Thus high stakes for students are the link between policy and prac-
Decentralized Systemic Policy

The importance of high stakes to the systemic agenda has been matched by the intensity of criticism for this particular feature. No other component has occasioned such an outpouring of resistance (Education Week, 1992), and the reasons for the criticisms are directly related to the reasons for the policy. In one sense, the resistance is predictable because all other objections, such as to the common curriculum, can focus on the component that gives the policy teeth (in much the same way that objections to regulations often focus on regulatory enforcement). But the connections between power and problems are deeper than this. For critics, the power of high stakes to influence conduct from afar (i.e., from the centers of power) guarantees distortions of the educational process and fatally threatens the validity of the exam itself. Unless the exam can be made extremely broad and incorporate almost all objectives of the curriculum, critics believe that high stakes will narrow the curriculum through strategic teaching to the test and cause test scores to be inflated relative to the full range of what students should know (Koretz et al., 1992).

High-stakes exams may also foster and institutionalize stratification of the curriculum. The reason is the pressure to set the passing grade low enough to avoid the legal and political problems of excessive numbers of students, especially poor minority students, failing (Debra P. v. Turlington, 1985; Jaeger & Tittle, 1980). While it is true that the difficulty of the exam can be “ratcheted up” once practically everyone passes (supposedly the case with the competency exams of Connecticut, Florida, and Texas), high stakes lead to an emphasis on the lower end of achievement rather than the more “aspirational” aspects of curriculum development (Porter, in press).

Alternatives to centralization. In one sense, the resistance to high stakes involves an interesting puzzle. When not centralized, high stakes seem to cause little controversy. Grades given by classroom teachers and scores on the AP tests, for example, are not controversial. Grades given by classroom teachers and scores on the AP tests, for example, are not controversial. High stakes on a national exam cause fear and loathing; high stakes elsewhere are treated as routine. Why the difference?

The answer clearly suggests the superiority of a decentralized system—high stakes are acceptable when the exams match the curriculum chosen voluntarily by the school. One might be tempted to explain the difference in terms of cost: Sophisticated centralized exams clearly are very expensive to administer. But the classroom assessments of teachers are expensive in terms of the labor of both teachers and students; AP exams are expensive (Koretz et al., 1992). It seems obvious that the expense is accepted when the assessments match what the teachers want to do, when the assessments are in some sense the natural products of the instruction, and when the instruction itself is fully accepted at the school level.

An advocate of centralized systemic policy might respond by saying that this is not really an objection to the high stakes but to the underlying policy and that the system would begin to accept the policy as soon as it was enacted. But I think the problem goes much deeper and involves a kind of Catch-22. Narrow high-stakes exams cannot be objected to for coopting the entire curriculum, but for that reason they are technically invalid. Subgroups of students will spend widely different amounts of time on a narrow, mandatory curriculum. Such a curriculum will be the maximum for some and the minimum for others. As the assessments are made broader and broader to embrace the full range of curricular objectives, the suppression of local options becomes more and more complete until nothing is taught other than the centralized curriculum (except perhaps to fast-track students who can master the curriculum with time to spare and to lower track students who are indifferent to the stakes). Thus, not only do the high stakes give the policy teeth; they require that the policy be completely prescriptive of local curricular options (including even “digressions” by teachers who are intrigued by some aspect of the material not targeted for assessment, such as the moral dilemma of Hamlet). High stakes directly threaten the compromise suggested by some advocates that the national curriculum should contain only a core of instruction.

To some extent, this Catch-22 is a dilemma for any system of external accountability and
to the entire agenda of performance orientation in education. But the centralized exam presents the dilemma in the starkest possible form whereas decentralizing options can greatly soften it. When schools (or individual teachers) voluntarily choose a curriculum, they are more willing to see it as the major part of their instructional mission and to accept the external monitoring (Darling-Hammond, 1992). Consider the AP as an example. First of all, the AP is not required; for schools, teachers, and students it has a real attractiveness but is not mandatory. Beyond its voluntary character, stakes on the AP are subtle, more in the nature of incentives or bonuses than penalties. There is little downside risk—AP students get course credit for passing the AP with a score that does not qualify for advanced credit—combined with a nice up-side reward if they do qualify. This bonus character of the ambitious curriculum can be reinforced by giving rewards for other important outcomes, such as graduation and college placement, rewards that cannot compete with a powerful, high-stakes exam. And the voluntary character of the curriculum can be further enhanced by allowing a variety of instructional products to serve as the raw material for external assessments, for example, when a certification team reads samples of student papers but accepts any topic assigned by the particular classroom teacher (presumably the topics themselves could be evaluated for their conformity to the ambitious curricular goals adopted by the school).

School Improvement

Argument for centralization. Probably the most ingenious argument in favor of systemic policy is that it would encourage school improvement in the sense of the widespread adoption of more ambitious curriculum content. The first step of this argument is that change toward a more ambitious curriculum occurs all the time in isolated schools and classrooms, but that these “hothouse flowers” do not sustain themselves, and mediocre instruction eventually replaces them. The second step is that systemic policy would provide a powerful, coherent set of authoritative signals and resources that would nourish ambitious change efforts and sustain them over time (O’Day & Smith, 1993).

Problems with centralization. The most obvious problem with the argument is the lack of any obvious mechanism for linking coherent policy adopted at the top with widespread educational change. Scores of studies show little or no effect of policies at the ground level (McLaughlin, 1987). Analogies to other countries are of limited value because the correspondence of policy and practice is supported by centralized governance and may reflect underlying cultural homogeneity (Cohen & Spillane, 1992). The idea that educational change would come through massive teacher training is also highly optimistic, since systemic change would come through massive teacher training is also highly optimistic, since systemic policy includes this as a concept more than as a clearly developed program, and to totally transform the skills and attitudes of the entire teaching force would be a massive, expensive, and socially revolutionary task. High-stakes centralized exams are a candidate for the engine of change, but the difficulties with this approach have already been discussed.

The final problem is that the process of school improvement seems powerfully inconsistent with a standard curriculum. It may be true that school restructuring needs the “beef” of curriculum content to improve achievement (Newman & Clune, 1992), but it also seems that for school improvement to occur, curriculum goals that are distinctive to the context of the school must be developed (Darling-Hammond, 1992). More ambitious learning goals may be avoided in the process, but they also may become the focal point. The power of a coherent, standard curriculum is no answer to this objection; in fact, it may intensify the problem. A standard curriculum might dampen the process of change by eliminating locally developed projects without being able to inspire much enthusiasm in the schools for a common curriculum.

Alternative to centralization. The argument for school improvement needs to be understood on its head and made the focal point of a strategy oriented primarily to instructional change at the school level. School adoption has enormous advantages as the locus of curriculum reform because the school is the place with a definable student population and a distinctive staff capable of adopting, adapting, and implementing goals (Clune,
Decentralized Systemic Policy

1990). Of course, this still leaves the problem of encouraging and sustaining change. We should take very seriously the idea that schools will have a hard time going it alone. The unconsidered alternative is that schools could choose from a menu of ambitious curricula and accompanying networks of teacher training, and the state could adopt the role of facilitating such adoptions and delivering resources for change.

As far as I am aware, the United States does not have much experience with using sustained technical assistance and professional development as instruments of curriculum change. Previous research findings that innovations do not sustain themselves are therefore suspect. California has had, in essence, a system of professional development and school improvement for some time (Adams, 1992; Marsh & Odden, 1991). Chicago’s reform movement has come to rely heavily on technical assistance for locally tailored curriculum improvement. Change well beyond single school sites seems to have occurred in these places without a major investment of resources. Thus, we are really in no position to reject a partially successful capacity-building strategy (McDonnell & Elmore, 1987) in favor of a completely untried regulatory approach.

School Finance

Argument for centralization. The argument for a centralized system is nowhere stronger than in the area of school finance, because many schools lack fiscal and technical capacity to deliver an ambitious curriculum to their students, and only central governments can redistribute resources. Modern school finance litigation and policy have sharpened our understanding of the role for central government by focusing on both inputs and outcomes and by raising questions about the management system that is necessary to build a more effective bridge between the two (Clune, 1992). On the one hand, we see a stronger role for compensatory education; schools are poor to the extent that they lack local tax revenues but also to the extent that their students lack educational resources from outside the school. Conversely, there is growing acceptance for the principle that money is not enough. To improve outcomes, we must improve the delivery of educational services in at least two fundamental ways: (a) by delivering specialized resources from outside the school, especially information about effective programs and well-trained teachers, and (b) by better managing resources at the school level through a vigorous process of school improvement emphasizing higher expectations, new curriculum content, and new pedagogies (Clune, in press).

Problem with centralization. Within the modern conception of school finance, there is certainly a profound problem with centralization: how to compute the resources needed by diverse schools to deliver an ambitious curriculum (program equity or adequacy) and finding ways for policy to stimulate improved service delivery in decentralized locations. But at least as originally conceived, systemic policy ignored these issues. The main resources needed were assumed to be generic—improved curricular models and better trained teachers (Smith & O’Day, 1990). Two characteristics of systemic policy have contributed to this blind spot: focusing on curriculum in the abstract, rather than the full range of factors affecting teaching and learning, and focusing on policy at the top rather than educational change at the bottom.

The result is somewhat paradoxical in terms of our review of the arguments for centralization, because the area with the greatest need for a strong central role has received the least attention and analysis. This is the point that upsets advocates of educational change for poor children—the image of systemic policy as a model curriculum in the sky lacking realistic mechanisms for bringing that model down to earth for the students who would receive the most benefit. In this sense, systemic policy does have a problem of being overcentralized—being so acutely policy centered that the real problems of low-achieving students and decentralized change are never confronted.

Alternative to centralization. We are just beginning to look seriously at the linked problems of program equity and program management. Program equity in a fiscal sense requires calculating the level and kinds
of resources that must be delivered to the school; program management is ensuring that these resources are transformed into effective teaching and learning. The decentralized model of state-supported curriculum improvement networks does offer a link between policy and practice. In a first step, researchers must provide the state with realistic estimates of the full cost of ambitious curricula, such as extra staff and adequately trained teachers in districts with concentrations of poor children (Madden, Slavin, Karweit, Dolan, & Wasik, 1991). Preliminary research suggests that the cost of achieving high minimum outcomes in high poverty schools may be about $5,000 per pupil per year more than the national average for such schools (Clune, in press). To be effective in raising student achievement, such expenditures should be invested according to a long-range instructional plan developed jointly by schools and an external supervising agency. This instructional plan, technical assistance, and evaluation could all be provided by a curriculum network or evaluation team (Lieberman & McLaughlin, 1992; NCTM, 1989).

Educational Governance and Instructional Change

Argument for centralization. The need for a change in governance is central to the case for systemic policy because the existing system of school governance is seen as incapable of producing a major upgrading of instruction. The present system is characterized by multiple power centers issuing fragmented control over various components of the curriculum-guidance system (curricula sporadically issued at different levels of governance, standardized testing in basic skills from commercial test makers, generic textbooks from national publishers, splintered control of teacher education, and so forth) (Cohen, 1990a). The essence of systemic policy is to replace this fragmented nonsystem with a coherent system of curriculum controls at the level of national and state governments.

Problems with centralization. The problems with building effective centralized systems in the United States are nicely summarized in an article by David Cohen and James Spillane (1992). The basic argument against change is surprisingly similar to the argument for it; both rest on the same vision of the existing low level of instruction and fragmented educational policy in the United States.

These authors see educational governance in the United States as fragmented not simply in its own right but as a result of and in proportion to the level of policy making in general. In the United States, ambitious policies are implemented in the context of decentralized and pluralistic governance. The result is that each new, ambitious policy tends to create new administrative structures that cut across governing levels and add new layers of complexity. The governance systems in other countries that integrate control of all aspects of instruction in central agencies (educational ministries) are lacking in the United States. As a result, centralizing efforts under way in some states and districts may simply add to the complexity of the signals received by schools rather than succeed at producing a coherent system. The bottom line is that a coherent system of educational governance in the United States would require a revolution in governmental structure.

Cohen and Spillane (1992) also assert that the argument that coherent policy, even if we could enact it, could change instructional practice is ambiguous and weak. No other country has succeeded in changing educational governance first and then changing instructional practice; rather, systems of governance and practice have developed together and tend to reinforce each other. Presumably, this was also the case with the New York State Regents, which has become deeply institutionalized at all levels of the system (Tyniec, 1993). Cultures that emphasize high academic achievement and hard work by students also reinforce a standard system. Americans' dislike for the strict sorting of students at early ages favors fragmentation. Also, even assuming that policy could change practice, the magnitude of the desired change is so great that the cost of the hypothetical policies may be prohibitive. It would be necessary to invest enormous resources in research and development of new curricula and centralized exams. The cost of teacher training would be monumental since prac-
Decentralized Systemic Policy

Tically every teacher in the United States has been trained in a way inconsistent with the new curricula and would need to acquire completely new skills.

At its base, the debate over the potential role of centralized governance is a familiar one. Modern organizational theory is most illuminating in explaining the reasons why almost all change is incremental (March & Olsen, 1989). Systemic policy tries to cure the problem of incremental change at the bottom through systemic change at the top, but this merely relocates the problem without solving it. One way of characterizing the basic approach of this article is an incremental strategy for producing curriculum coherence at the bottom in a system of fragmented governance and culture.

Alternative to centralization. Despite the overwhelming difficulty of switching to a centralized system, the argument for greater coherence is so powerful that even advocates of a more decentralized system, such as myself, must explain how a decentralized system can provide coherence. Indeed, the argument here is that decentralized guidance is both more feasible and a more powerful tool for producing coherent, ambitious curricula at the school level.

Cohen and Spillane (1992) obliquely address the issue of incremental change by saying that a strategy of bypassing educational governance would be blocked by existing layers of governmental complexity. But in objecting to such a strategy, Cohen and Spillane seem to be thinking of nongovernmental organizations competing with government for top-down influence, thereby intensifying complexity and fragmentation. The bypass strategy recommended here is really quite different. Instead of defining a standard curriculum, governments would offer incentives for schools to join approved curriculum networks and would require data on the progress of curriculum improvement. This strategy takes advantage of our fragmented system by capitalizing on its one advantage: the ability of schools to ignore all the weak signals and create coherence at the school level. To take some current examples, the Levin (1988) and Slavin (Madden et al., 1991) networks are able to create powerful instructional coherence at the school level in spite of some obstacles from the policy environment. Removal of such obstacles would be an obvious part of a decentralized policy package. There is also good reason to think that decentralized curriculum development is the most effective system for making teachers both accountable for what occurs in their classrooms and enthusiastic about the process (Hannaway, 1993; Meier, 1992).

Teacher Training

The argument for centralization. The need for standardized teacher training is both a conclusion from and an argument for a common curriculum. If we have a common curriculum, the argument goes, teachers who come from a variety of places should receive common, rather than irrelevant or conflicting, training, but, more strongly, a diversity of training prevents the development of any kind of curriculum even in a single school. If we must standardize teacher training to produce school change, why not take the logical correlative step of formalizing the common curriculum?

Problems with centralization. Once again, whatever the merits of these arguments, the problems with implementing the vision are enormous. First, all the problems of obtaining consensus on the common curriculum apply to consensus over the content of teacher training, but, in addition, such consensus would need to be developed in a whole range of very diverse and autonomous institutions and activities: departments and schools of education in universities, national and state teacher examinations, the credentialing systems of every state, and sprawling, decentralized systems of in-service training. The diversity of teacher training is not just a problem for developing a common curriculum; it is a problem for implementing one.

As with curriculum development, there also seems to be some increasing diversification of teacher training. For example, growing support for alternatives to certification exist side by side with increased, and diverse, credentialing requirements (Darling-Hammond & Berry, 1988). A revolutionary change may occur wherein elementary teachers become subject matter specialists
rather than teachers of all subjects. Finally, David Cohen (1990b) reminds us that teacher training is a challenging problem relating to cultural change—that the most important part of teacher training may be the kind of teaching most teachers experienced as students before they ever had any training as teachers.

**Alternatives to centralization.** Such difficulties do not negate the need for a common curriculum, but they would make a more diverse, decentralized, and incremental strategy considerably more attractive. And, once again, on closer scrutiny, the arguments for standard training do not seem particularly powerful.

The need for changes in teacher training seems to break down into two parts: general competence and preparation in specific curricula. The main criticism of pre-service teacher training is not lack of preparation for a detailed curriculum but lack of preparation in the content of subjects, like mathematics, and pedagogies such as teaching for understanding and active learning (Cohen & Barnes, 1993a; Smith & O’Day, 1990). Reform of teacher training to meet this kind of need can proceed independently of the development of specific curricula as long as there is consensus about the broad type of knowledge and skill that is necessary. Indeed, broad teacher training has the same advantage in meeting a variety of teaching needs as the teaching of general occupational skills has for general employment (Darling-Hammond, 1992). Teacher training would be stronger, not weaker, for looking at a variety of curriculum visions and approaches. The development of national standards for advanced teacher certification, which is now in progress, is fully consistent with the model of improved general competency in subject matters and subject matter pedagogy. Teachers meeting such standards should be able to teach more effectively across a range of curricula.

On the other hand, training for specific curricula can and must be accomplished at the school level. Although radical changes will be required from most kinds of preexisting in-service, such changes become much more feasible when support comes both from the local school staff and externally. The optimum role for the state is to facilitate decentralization, for example, by redirecting financial support toward curriculum-specific training elected by the school.

**Policy Agenda**

The following comments should be understood as a discussion of guiding principles for policy, not the design of policy itself. Most of the hard thinking about detailed policy design (e.g., how to craft workable incentives) remains to be done.

Policy should be based on the belief that there is a need for curriculum change and that a powerful stimulus for change must come from outside the school. The trick as I see it is finding the right balance among competing forces. I submit that a state curriculum is ill-suited to achieve these objectives. A superior, more decentralized system would work roughly as follows:

1. Develop consensus about a range of desirable curricula. The policy goal of curriculum upgrading cannot be achieved through any curriculum that bears the label “higher order thinking” or “accelerated.” The institutions of educational policy and practice are far too sophisticated at the fine art of packaging and labeling for that. A decentralized change strategy requires some central guidance. Thus, central agencies (not necessarily governments) should continue to develop models for upgrading the curriculum. But the emphasis should change from developing the one, ideal curriculum to developing a selection, or menu, of curriculum models calculated to appeal to the broadest possible range of students and schools. For example, the menu could include both innovative and traditional mathematics adapted for low-income students.

2. Sponsor and facilitate a range of curriculum development efforts. Having models for curriculum development is not the same as having actual, working curricula including instructional materials. Thus, the state should encourage a multitude of curriculum development efforts that move in the right direction. Care should be taken that curricula are built for the full range of educational situations and institutions, including all regular
tracks, special tracks, and specialized schools (e.g., compensatory, special, and vocational education). The goal is to make certain that various kinds of ambitious curricula are available for all types of students and schools.

3. Facilitate creation of curriculum development networks. The primary change agent in a decentralized system is the curriculum development network, an agency to assist schools in the transition to a new curriculum. This would include teacher training and evaluating course content, pedagogy, and instructional products (Lieberman & McLaughlin, 1992). Because of the scope of the change, that is, the large number of schools needing change, the state must be prepared to sponsor such networks on a wide scale. To minimize cost, it will be important to design low-cost methods of facilitation, such as setting aside portions of existing state and local budgets for payment to the networks. Special attention and first priority should be given to development and change networks for low-income schools and students.

4. Create incentives for schools to initiate and sustain approved processes of curriculum upgrading. The key decision point in any version of systemic education is the time when the school adopts a new curriculum and makes the commitment to a sweeping process of school improvement and teacher training. Research on school change tells us that there can be much anxiety about and resistance to such a major change, a change that will profoundly affect every teacher in the school (Huberman & Miles, 1984). Furthermore, the change process described in this article includes new mechanisms of accountability that require thorough inspection of classroom practice and performance. For these reasons, I believe that the state must develop incentives that will induce schools to initiate and sustain the process of change. For example, it can budget for teacher training, a new school finance package (see below), and rewards for successful curriculum development.

5. Develop personnel policies that compensate skill and effort in the new curriculum. Experience suggests that an ambitious curriculum is challenging for teachers as well as students and is probably beyond the capacity of many teachers. Education in its present form would be well served by greater rewards for skill and effort, but such policies may be critical for serious curricular upgrading (Odden & Conley, 1992).

6. Develop a standardized student assessment system that tracks progress and leads practice. A central, standardized student assessment is not the primary indicator or agent of change in a decentralized system, but it serves two vital functions: (a) It permits tracking of core skills in groups of students and administrative units (including schools) (Finn, 1991), thus providing a common measure of progress and a checkpoint in a system designed to produce great diversity, and (b) it provides a model for practice by providing an authoritative operational definition of higher order skills held in common by a range of optional curricula.

7. Develop an indicator system for tracking course content, pedagogy, and instructional products. The primary indicator of change in a decentralized system is school-level information on progress toward full implementation of an ambitious upgraded curriculum. Inspection teams from curriculum networks would have the task of reporting progress (or lack thereof) in a wide variety of school curriculum practices: course content, instructional materials, pedagogy, and instructional products (tests, papers, experiments, etc.). Progress would be measured against clearly defined goals, including the level of instruction offered to different groups of students within the school. Inspection frequency would correspond to logical periods of school improvement, for example, once a year for 5 years and every 3rd year thereafter. Conventional accreditation should be abolished as completely inconsistent with the new approach.

The advantage of such a system is that a great deal of extremely reliable and valid information can be gathered and evaluated at a relatively low cost. As a rough estimate, inspection teams should be able to collect materials and visit classes in about a week and complete a report in another week. A possible weakness of the system is a conflict between the role of helper and evaluator in the curriculum network; the same organization has a stake in proving progress and remaining
objective. This conflict of interest may require establishing independent audit teams, perhaps not even associated organizationally with the networks but intimately familiar with network goals (but see the NCTM evaluation standard 14, which recommends that teachers from each school be on the evaluation team; NCTM, 1989).

Supplementing a universal school inspection system could be a centralized system for collating the results of any high-stakes, standardized student examinations voluntarily adopted by schools (e.g., the AP).³

8. Develop a new school finance system guaranteeing adequate compensatory resources and resource management. Much work remains to be done on three major problems of school finance: the basic entitlement to resources, the method of computing compensatory aid, and the management system for assuring cost-effective use of resources. Elsewhere, I have written (Clune, in press) that we are presently stranded between two systems for handling these problems: (a) a system, which I call “equity plus,” consisting of a high foundation program for almost all districts in a state, a substantial but inexactly measured amount of compensatory aid, and an external accountability system (Clune, 1992), and (b) a system, which I call “true adequacy,” that concentrates aid on children with very low outcomes, calculates the amount of aid in terms of the cost of instructional programs capable of reaching high minimum achievement standards, and manages resources through supervised local planning and implementation (Clune, in press; see also Greene, 1993). Much research remains to be done on the cost of effective education for low-income students, but a rough estimate can be implemented immediately until better information tells us how to establish new benchmarks.

9. Initiate key reforms of teacher education. Changing our system of teacher education to fit common elements of the new, ambitious curricula is a long-term enterprise involving many different institutions and policy interventions. A new system of national certification (“boards”) could serve to orient in-service training and professional development. Modification of preservice licensing exams and alternative route criteria should have a strong influence on preservice teacher training (i.e., schools of education). Universities could be given incentives for developing courses in new content and pedagogies (e.g., active learning, teacher as “coach”). In-service budgets at both the state and the school levels should be made available exclusively for approved content-oriented school improvement. In a whole range of ways, teachers should be encouraged to accept curriculum and pedagogical development as the central and permanent characteristic of the life of a professional and school.

10. Design a deregulation package to remove obstacles to curriculum upgrading. Repeal of old, inconsistent laws and regulations is easily overlooked in the excitement of building a new system, but a new approach may be difficult or impossible without a rigorous exercise of deregulation. Among the rules that must be repealed are (a) those that regulate input in any way other than approved curriculum upgrading (e.g., old-fashioned accreditation and state standards based on inputs), (b) all standardized testing other than the new student assessments, and (c) all personnel policies inconsistent with rewarding teachers for knowledge and skills (see Odden & Conley, 1992).

11. Encourage development of a broad array of outcome-based assessments of individual students and selected high stakes to provide incentives for greater learning. So far I have described a system heavy on incentives for school improvement but light on incentives for the individual student. To correct this bias, the state should develop policy designed to (a) encourage creation and adoption of evaluations of individual students that clearly report performance on meaningful outcomes (e.g., proficiency scores, student portfolios) and (b) create meaningful rewards for good performance. Examples of desirable policies might include (a) mandates or financial incentives aimed at getting curriculum networks and schools to adopt and implement useful systems of student assessment and (b) encouraging the use of such assessments in a wide variety of social and economic decisions (e.g., entrance to jobs and further education).
12. Sponsor efforts to coordinate different parts of the delivery system. A decentralized flexible system does have the problem of articulation and coordination among its various parts. Thus, the government should sponsor efforts to give students mobility and access from one part of the system to another (e.g., enabling vocational students to enter 4-year colleges).

Research Agenda

Any serious effort to produce change in practically every school in the United States implies the need for a large and sophisticated enterprise of research, development, and evaluation. One of the problems with existing research on systemic policy is the focus on policy development—how systemic the policy is, not how much change occurs in schools. Details of such a vast research agenda cannot be discussed here, but it may help to suggest some broad categories:

1. Evaluation of progress. Evaluation of change in a single organization or program is challenging enough. Evaluation of change in every school in the United States is truly a major enterprise. Factors that must be tracked include changes in instructional practice, knowledge and skills of teachers, and achievement of students. To facilitate evaluation, a tracking capacity must be built into the system itself. Analytical and reporting techniques from centralized systems of student assessments are already well developed, but an entirely new system is required to consolidate and synthesize the reports of school inspection teams about changes in practice. Tracking the knowledge and skills of teachers also could be built into the school inspection process.

In sum, a problem with a highly decentralized system is losing track of progress. To minimize the problem, the state should establish a logistical plan and monitoring system capable of tracking diffusion of the new curricula to the whole system.

2. Factors producing and impeding change at the school level. Actual change in schools does not occur simply because the system has been designed to produce change. Research tells us that change is a complex, local process with many unexpected obstacles to be avoided both inside and outside the organization. For example, curriculum networks are the primary agent of change in the system proposed here, but research on these enterprises is really just beginning and has already revealed a number of questions and tensions: how much and what kind of teacher training, at what cost; how to coordinate work of the network with policies at the school and district levels; how to balance time spent in network activities and time in the school and classroom (Lieberman & McLaughlin, 1992; Roemer, 1991). Another prime area of research about change is on obstacles encountered at the ground level, such as policies, laws, programs, and structures of governance that make change difficult or impossible (e.g., categorical programs, personnel policies, desegregation decrees, interference from various government authorities).

The research described here is not at all the same as evaluation. Evaluation is a reliable system for monitoring progress, formative or summative. This kind of research requires a much closer look at a variety of situations involving change and lack of change and would require a serious effort in research design.

3. Policy design. Perhaps logically this topic should have been discussed first: How can one measure change before designing the policies that are supposed to produce change? In fact, the design and refinement of policy must be a continuing enterprise in such a complex system (Clune, 1993). Obviously, there must be a major exercise of policy design at the beginning of the process, but the first set of policies is likely to be no more than a set of well-educated guesses about what really will work. The scale of the research needed here springs from the number of policy areas (curriculum, testing, teacher education, school change, student learning, school finance) and the complexity of the links between them. Examples of some of the most important questions are, what system of incentives is sufficient to induce schools to initiate and sustain a process of fundamental change (McDonnell & Elmore, 1987); what system of student assessments is capable of measuring the core of various ambitious curricula; what resources should be delivered to
various kinds of schools by the system of school finance; what systems of teacher testing, licensing, and training produce the needed knowledge and skills; and what kinds of student report cards and high-stakes result in higher student motivation?

4. Need for a nationally coordinated research enterprise. The need for all this research suggests a somewhat paradoxical conclusion: A centralized research function is required as the infrastructure of a highly decentralized system (Koretz et al., 1992). Although we do not need national or state governance, we really could use some kind of national research authority linked to corresponding units at the state level to provide feedback on results and guidance on technical assistance. While the authority should be nationally coordinated, it should probably be managed by some kind of independent agency reporting to the government. For such complex, expensive, and sophisticated research, what comes to mind is governing panels of highly qualified researchers along the lines of the National Institutes of Health.

Conclusion

The system proposed here should be viewed as a completion rather than a rejection of the basic insights of systemic educational policy. Instructional change certainly does require higher standards and coherent, coordinated policy from outside the school. But on close inspection, the statist-centralized version of systemic policy built around authoritative curriculum frameworks is fatally flawed on the two grounds: A common curriculum is difficult, if not impossible, to apply considering the immense diversity of American schooling, and a tolerable link between policy at the top and change at the bottom is all but unattainable. The system recommended here solves both problems by creating a system of change agents (curriculum networks) using diverse curricula approved as ambitious and chosen by local schools and by emphasizing issues of capacity-building, such as teacher training and school finance.

Indeed, I think that a centralized system would evolve into the kind of system proposed here if it were adjusted in light of the realities of implementation. I suspect that the only feasible kind of core curriculum would be a core of higher order general skills (reading, writing, reasoning, mathematics, problem solving) that cut across a range of diverse curricula (academic, applied, vocational, etc.)—something like the vision behind level 3 of the NAEP (Romberg, Smith, Smith, & Wilson, 1992). The limited scope of this core would then require development of full-fledged curricula for use in the schools. The restricted core curriculum could not be made the basis of a high-stakes exam, because some schools and classrooms would teach nothing else, while students in other schools and classrooms could pass the exam with little instructional assistance. Meanwhile, schools lacking technical and financial capacity for curriculum improvement would certainly require assistance beyond the frameworks and standardized exams, resulting in changes in school finance and some kind of diffusion network such as the curriculum networks recommended here (see the recommendation of networks and school inspectors in O'Day & Smith, 1993). The likely alternative to this reinvention of a decentralized system is not attractive: an unprecedented degree of local variation (consisting of exceptions and waivers) masked by the formal apparatus of a standardized system.

The basic soundness of the approach is confirmed by how easily one can begin to visualize the distinct components of a complete system and recognize the need for a major enterprise of research, development, and evaluation. A practical, change-oriented system built from the bottom up turns out to be less mysterious and more clearly challenging than its centralized counterpart. No system is perfect, however, and a decentralized system has its own problems, perhaps the greatest of which is the tendency to stop considerably short of systemwide change. Achieving even limited change certainly is a problem in an age of collapsing educational budgets, but it is not solved by regulating the curriculum. Good curriculum goals are readily available, but delivery of change to the school site requires scarce resources. In fact, limited budgets make it imperative to establish priorities, and the greatest need for
curriculum improvement is in low-income schools. This need for priorities suggests that a compensatory education program like Chapter 1 at the federal level should focus directly on the availability of adequate resources and high quality instructional models for low-income schools rather than requiring uniform state regulation as the unreliable instrument of local school improvement.

Notes

The research reported in this article was supported by the Consortium for Policy Research in Education (CPRE), Grant No. OERI-R117-G10007, and by the Wisconsin Center for Education Research, School of Education, University of Wisconsin–Madison. The opinions expressed in this article are those of the author and do not necessarily reflect the views of the U.S. Department of Education, Office of Educational Research and Improvement, the institutional partners of CPRE, or the Wisconsin Center for Education Research. Thanks to the students of two related seminars in the La Follette Institute of Public Affairs, University of Wisconsin–Madison. Thanks also to Andy Porter, Ken Snyder, and the reviewers of this journal for comments and to Lisa Armstrong and Alice Honeywell for timely and excellent technical support.

1In the category of “nonstate capacity,” I would include national organizations developing model curricula—NCTM (1989), 2061 (Rutherford & Ahlgren, 1990), SCANS (1991), the New Standards Project (University of Pittsburgh, 1992), and the College Board’s “Equity 2000” (Pitsch, 1992) and “Pacesetter” exam (DeWitt, 1992)—as well as networks of schools emphasizing new approaches to the curriculum. Levin’s accelerated schools (1988), Slavin’s “success for all” program (Madden et al., 1991), Sizer’s essential schools (1989), and the multitude of restructuring experiments sponsored by states and districts are examples. The fact that parts of the government are involved in some of these projects does not make them state enterprises in the sense of national curricula and educational ministries. A decentralized strategy depends on and involves various governments at various points. My general point is that there has never been such a range of efforts focused around serious curricular goals and prepared to offer high quality products of one kind or another. Further, the institutional roots of these efforts are quite durable: universities, disciplinary and teaching associations, national foundations, research organizations, and so forth.

2“Political institutions are simultaneously an affront to our sense of comprehensive rationality and a primary instrument for approximating it” (p. 17). “The long-run development of political institutions is less a product of intentions, plans, and consistent decisions than incremental adaptation to changing problems with available solutions within evolving structures of meaning” (p. 94).

3In a recent paper, Andrew Porter (in press) advises against school delivery standards and recommends reliance on outcome measures, especially student assessments. Porter’s basic objections to process (input) standards seem to be excessive prescriptiveness and lack of local ownership and enthusiasm, problems that would indeed be serious in a standardized system. But the decentralized system recommended here is not prescriptive, relies on voluntary adoption, and offers incentives for change. And while I agree on the importance of gathering information about the “bottom line” (student outcomes), information about curricular change is more concrete and informative than data on outcomes. For example, if we measure gain in achievement, how much gain should we expect and demand for different sorts of students and schools? Effective school accountability seems to require measures of both inputs and outcomes interpreted in light of each other. Porter also points out that outcome data are cheaper to collect than process data. The cost of inspection in the system I recommend could be minimal relative to the cost of school improvement (because the teams already are in the schools), and much of it probably could be absorbed in existing state and district budgets (e.g., for testing, curriculum development, and teacher training), but cost remains an important issue for research.

The most serious problem with the paper is the limited focus on accountability at the school level—this is what the paper means by “delivery standards.” Centralized process standards enforced by high stakes would certainly be troublesome, but the most troubling thing about them is holding schools accountable for performance without holding the state accountable for resources. The approach recommended here turns the issue of delivery standards on its head, holding the state accountable for delivering adequate resources, including curriculum models, and holding the school accountable for implementing the curriculum it has voluntarily adopted.

References


Decentralized Systemic Policy


Clune

12th-grade items. Unpublished manuscript, National Center for Research in Mathematical Sciences Education, Wisconsin Center for Education Research, University of Wisconsin–Madison.


Author


Received December 14, 1992
Accepted April 12, 1993