Chapter 1

Complexity and Policy Implementation

Challenges and Opportunities for the Field

Meredith I. Honig

Education policy implementation as a field of research and practice for decades has amounted to a sort of national search for two types of policies: "implementable" policies—those that in practice resemble policy designs—and "successful" policies—those that produce demonstrable improvements in students' school performance. This focus on what gets implemented and what works makes sense especially in education. After all, education has become a high-stakes, big-budget policy arena. Education commands a lion's share of state and local budgets to levels that beg hard questions about the feasibility and value added by education policies. Given its promise to serve as a significant lever of change in an institution intended to serve all children and youth, education policy affects multiple dimensions of social welfare. And given these high stakes, education policy implementation warrants careful scrutiny.

However, recent trends in education policy signal the importance of reexamining what we know about what gets implemented and what works. In practice, education policy demands arguably have become more complex. School systems now are held accountable for demonstrable improvements in the academic achievement of all students in ways barely imagined just 20 years ago. Across the country, the increasing ethnic and racial diversity of public schools (Ladson-Billings, 1999; Villegas & Lucas, 2002), a shrinking base of resources for education in many states and districts, and new systems of negative sanctions for underperforming schools (Massell, 2001; O'Day, 2002) only add to the urgency and challenge of meeting those standards. Contemporary public school systems vie for resources in competitive and contentious political arenas against projects for roads and sewers, prisons, and health care as well as school alternatives such as vouchers, charters, private schools, and home schooling. Research and experience continue to deepen knowledge about how students' experiences in school are highly dependent on conditions in their neighborhoods, families, and peer groups in ways that up the ante on school

improvement efforts to look beyond school walls for key reform partners (Honig, Kahne, & McLaughlin, 2001). The federal government, states, school districts, mayor's offices, and others each promote various educational reform agendas that typically converge on schools simultaneously (Honig & Hatch, 2004; Knapp, Bamburg, Ferguson, & Hill, 1998).

In such contentious, interconnected, and multidimensional arenas, no one policy gets implemented or is successful everywhere all the time; on the bright side, some policies are implemented and successful some of the places some of the time. For example, some research on class size reduction links smaller class sizes with increases in student performance but other research reveals no improvement (Finn & Achilles, 1990; Gilman & Kiger, 2003; Hanushek, 1999; Illig, 1997; Smith, Molnar, & Zahorik, 2002; Zahorik, Halbach, Earle, & Molnar, 2004). Linking student support services to schools has been shown not to expand students' opportunities to learn but under certain conditions it has been associated with various positive youth development and learning outcomes (Honig et al., 2001; Mathematica Policy Research & Decision Information Resources, 2005). In some districts charter schools outperform neighborhood public schools but nationwide their performance has been mixed (Mishel, 2004, September 23; Viadero, 2004a, 2004b). Single-sex schooling seems both to strengthen and to impede educational outcomes (Datnow, Hubbard, & Conchas, 2001; Lee & Bryk, 1986). Accountability policies and other central directives have limited impacts on teachers' practice in some settings but significant effects in others (Elmore & Burney, 1997; Firestone, Schorr, & Monfils, 2004; Louis, 1994a).

These realities of schooling in diverse communities nationwide suggest that those interested in improving the quality of education policy implementation should focus not simply on what's implementable and what works but rather investigate under what conditions, if any, various education policies get implemented and work. In this view, "implementability" and "success" are still essential policy outcomes, but they are not inherent properties of particular policies. Rather implementability and success are the product of interactions between policies, people, and places—the demands specific policies place on implementers; the participants in implementation and their starting beliefs, knowledge, and other orientations toward policy demands; and the places or contexts that help shape what people can and will do. Implementation research should aim to reveal the policies, people, and places that shape how implementation unfolds and provide robust, grounded explanations for how interactions among them help to explain implementation outcomes. The essential implementation question then becomes not simply "what's implementable and works," but what is implementable and what works for whom, where, when, and why?

The complexity of these policy dynamics poses a dilemma for policy analysts, policy researchers, and others who routinely produce information about

implementation. On the one hand, primary audiences for implementation information—elected officials, public managers, school principals, and others—by many reports demand clear, actionable, and reliable information that can guide their decisions especially in complex policy arenas; clear information in such arenas is often considered that which limits complexity and provides unambiguous action steps and chains of command (Cohen & Weiss, 1977; Majone, 1989; Weiss & Gruber, 1984). For example, policy recommendations in this spirit might call for the implementation of a single district-wide reading curriculum to help ensure that all schools are on the same page and offering consistent, coherent instruction. Such recommendations might urge the development of "what works" lists and clearinghouses and seek relatively unambiguous verdicts regarding program success (http://www.whatworks.ed.gov). A more complex view of education policy implementation may appear particularly unwelcome in the high-stakes accountability environments of many states and districts where short timelines for producing demonstrable improvements put a premium on swift and confident action.

But on the other hand, if such information and recommendations gloss over public school systems' complex day-to-day realities they run the risk of missing their mark and actually undermining progress. Without detailed information about the conditions under which certain interventions work, decision makers will not know if the failure of a particular reading curriculum, for example, stemmed from their choice of curriculum or poor conditions for implementation. Lists of recommended programs that "work" may obscure the resources and practices that enabled those programs to work, inadequately explain implementation results, and otherwise fail to help educational leaders understand which "successful program" might actually be successful with their own staff and students in their workplaces and communities. Recent federal emphases on scientifically based approaches to improvement arguably up the ante on researchers and practitioners alike to better understand the value and applicability of particular educational research in specific educational contexts.

This book starts from two broad premises: confronting the complexity of policy implementation is essential to building the kind of instructive knowledge base that educational decision-makers demand; and strong theoretical and empirical guides are needed to help researchers and practitioners navigate this inherently messy terrain. The time is ripe for a compilation of studies that build on this perspective. Decades of education policy implementation research and experience have been pointing to the complexity of implementation (Elmore, 1983; Sizer, 1985) and, specifically, to policy, people, and places as essential interrelated influences on how implementation unfolds (Odden & Marsh, 1988). Some scholars have developed models theoretically consistent with such descriptors (e.g., Goggin, Bowman, Lester, & O'Toole, Jr., 1990). However, arguably for the first time in education, implementation studies that

confront complexity have reached a critical mass and when examined together begin to elaborate what productively confronting complexity might entail.

This new generation of education policy implementation research is distinguished by three specific features: (1) the policies under investigation on the whole are significantly more comprehensive and varied than in previous decades; (2) the research aims to uncover the various dimensions of and interactions among policies, people, and places that help explain variations in policy results; and (3) the basic epistemological approach of the research reflects the importance of moving beyond universal truths about implementation (e.g., "you can't mandate what matters") to revealing implementation as a complex and highly contingent enterprise in which variation is the rule, rather than the exception.

This volume brings together scholars whose original empirical work contributes to this new generation of implementation research. No one chapter promises to present an overall model of policy implementation. However, when viewed together in combination with other contemporary education policy implementation studies, these chapters begin to add up to a portrait of education policy implementation as the product of the interaction among particular policies, people, and places. These studies suggest that education policy researchers and practitioners interested in improving the quality of education policy implementation should help build knowledge about what works for whom, where, when, and why.

To help elaborate the distinguishing features of these chapters and contemporary education policy implementation research more broadly, I first locate recent studies in the context of past generations of implementation research identified by many other researchers (e.g., Goggin et al., 1990; Radin, 2000; Wildavsky, 1996). My analysis of these research waves reveals that contemporary implementation research in many ways builds directly on lessons learned from the past and seeks to deepen past findings. In doing so, contemporary research breaks from the past along particular dimensions that mark it as a distinct generation—one seeking more nuanced, contingent, rigorous, theory-based explications of how implementation unfolds. I highlight throughout the next sections how each chapter in this volume illuminates this approach and I conclude with implications for implementation research and practice.

A Brief History of Education Policy Implementation

Generalizing about generations or waves of research in a multidisciplinary field such as education policy implementation surely obscures variations in the work underway at any given time. However, bodies of research during different time periods may reflect prevailing approaches and underlying assumptions that help mark distinct evolutions in knowledge. In scholarly reviews of

education policy implementation research—and implementation research in other social policy arenas—there is remarkable agreement that the field has passed through at least three stages (Odden, 1991a; see also Goggin et al., 1990; Lennon & Corbett, 2003; Radin, 2000; Wildavsky, 1996). I find that scholars generally distinguish each stage by (1) particular features of policies enacted and examined and (2) predominant approaches to implementation research. A review of these stages with attention to selected outlying studies helps highlight that contemporary education policy implementation research both builds on and departs from all three past eras in ways that mark a distinct new phase of knowledge-building about implementation.

Wave 1: A Focus on What Gets Implemented

According to Odden and others, implementation research as a formal field of inquiry emerged in the 1960s. Early implementation studies mainly focused on federal Great Society Period policies such as the Elementary and Secondary Education Act (ESEA), then newly passed in 1965 (Murphy, 1971). These policies aimed to achieve broad societal goals such as eradicating poverty but implementers were evaluated along far more modest measures: namely, the extent to which schools delivered supplemental services to low-income students (Elmore & McLaughlin, 1988). Policy designs were largely distributive, categorical, and regulatory in nature. That is, they aimed to help spread particular resources (typically funding) to groups or categories of students who met particular eligibility criteria and to ensure the appropriate use of resources as specified by policy makers. These policy designs were generally top-down in orientation—based on assumptions that policy makers *should* develop policies for implementers to carry out and monitor implementers' compliance.

   Supported in large part by federal contracts, many Great Society Period researchers conducted large-scale evaluations of these policies and were almost unanimous in their findings of implementation failure—schools and districts tended not to put programs in place in ways that faithfully resembled policy designs or, in economic terms, that could be predicted by policy designs. Researchers and others generally traced root causes of these failures to conflicts between policy makers' and implementers' interests and to implementers' overall lack of capacity and will to carry out those instructions (Murphy, 1971; see also, Derthick, 1972; Pressman & Wildavsky, 1984). Such assumptions stemmed in part from conventions of particular academic disciplines such as economics and political science—dominant in implementation research at that time—that viewed the individual implementer as the most meaningful unit of analysis and posited that these individuals were driven by individual self-interest to behave in ways not always congruous with policy designers' goals. Coalition building

among implementers, stronger incentives, and clearer instructions for imple-
mentation were heralded as important strategies for closing policy design-imple-
mentation gaps (Bardach, 1977; Sabatier & Mazmanian, 1979).

Wave 2: Attention to What Gets Implemented over Time

In the 1970s, predominant policy designs reflected some continuity and some
change. The federal Great Society Period programs persisted as a focal point
for implementation research thanks in part to ongoing federal evaluation con-
tracts. However, long-standing policies such as ESEA through multiple reau-
thorizations had come to include more specific regulations and other guidance.
By the 1970s ESEA and its signature program, Title I, had become old-hat for
many schools and districts that had been implementing its programs for
almost a decade. In addition, the types of federal policies under study
expanded to include other distributive, categorical, and regulatory policies
such as those for special education students.

Research on these federal policies also followed a pattern of continuity
and change. Research questions still probed fidelity of implementation. For
example, Kirst and Jung demonstrated that over extended periods of time, fed-
eral programs in practice did resemble initial policy designs and they con-
cluded that longitudinal approaches to policy making and policy research
would improve implementation (Kirst & Jung, 1980; see also Farrar & Milsap,
1986; Knapp, Stearns, Turnbull, David, & Peterson, 1991). However, a hand-
ful of researchers began to concern themselves with variations in implementa-
tion and to forecast the importance of policies, people, and places as mediators
of implementation.

For example, Peterson, Rabe, and Wong highlighted that policy designs dif-
fered not only in the details of their provisions but also in terms of their under-
lying mechanisms for allocating resources. They argued that the implementation
of redistributive programs (those that required government to provide more
services to certain generally underprivileged groups) led to more conflicts at var-
ious points in the policy process than developmental programs (those that made
infrastructure investments and promised benefits for wider groups) (Peterson,
Rabe, & Wong, 1986; Peterson, Rabe, & Wong, 1991. See also Lowi, 1969).
Other studies began to cast implementers in a different light—not as individu-
als who lacked the motivation to change but as engaged actors trying to cope
with the sheer number of new policy requirements that converged on the "street
level" (Weatherley & Lipsky, 1977) and to reconcile workplace demands with
their personal and professional worldviews (Radin, 1977). The importance of
attending to places or local context edged to center stage thanks in large part to
the RAND Change Agent study. This study found, among other things, that

implementation is shaped by macro- (policy-level) and micro- (implementation-level) influences; implementation unfolds as a process of "mutual adaptation" as implementers attempt to reconcile conditions in their microlevel context with macrolevel demands (Berman & McLaughlin, 1976, 1978).

These and other landmark studies began to herald *that* variations among policy, people, and places mattered to implementation. However, studies during this period seldom elaborated *how* they mattered. For example, few disagreed that local context mattered to implementation but instructions to attend to context said little about the dimensions of context that mattered, under what conditions they mattered, whether context could be attended to, and if it could, how policy makers should do so (Kirst & Jung, 1980). Furthermore, the general orientation to knowledge-building about implementation during this decade reflected persistent concerns with closing the gap between policy makers' intentions and implementers' actions and reinforcing top-down command-and-control relationships between policy makers and implementers. New, "alternative" policy models and tools such as backward mapping and decision checklists for policy makers aimed to help policy makers anticipate implementers' deviations from policy makers' plans and to take steps to avoid such implementation "pitfalls" at the point of policy design (Elmore, 1979-80; Elmore, 1983; Sabatier & Mazmanian, 1979, 1980).

Wave 3: Growing Concerns with What Works

In the 1980s, policy demands shifted again thanks in part to the publication of *A Nation at Risk* in 1983, the growing maturity of the federal Great Society Period programs, and the emergence of states as designers of broad-based policy initiatives. Policy designs during this period not only aimed to ensure full implementation but to achieve demonstrable improvements in students' school performance through new attention to curriculum and instruction and teacher professionalism (McLaughlin, 1990b). As Fuhrman and others have noted, policy making and policy research in previous eras generally:

> . . . centered on individual programs many of which were for special needs students and were more peripheral than central to core elements of schooling. They were discrete and amenable to study. By contrast, the current [Wave 3] reforms deal with central issues of who shall teach and what shall be taught and in what manner. (Fuhrman, Clune, & Elmore, 1988, p. 239)

Some of these policies stemmed from state educational agencies which emerged in many regions across the country as significant education reform

leaders. State policy development focused in part on the categorical federal programs of the prior decades but also on curriculum (Anderson et al., 1987). For example, during this decade California launched a major effort to develop curriculum frameworks and grade-level initiatives to guide school decisions and teacher professional development (Knapp et al., 1991; Marsh & Crocker, 1991; Odden & Marsh, 1988). States and districts passed and promoted prominent initiatives that called for school restructuring and school site-based management (David, 1989; Malen, Ogawa, & Kranz, 1990) and other strategies to reshape basic, usually formal school structures based in part on the theory that such restructuring would contribute to better decisions about various school operations. A related strand of policy making emerged from a host of other and relatively new "policy makers" including "whole school reform designers" such as the Coalition for Essential Schools, Accelerated Schools, and Comer Schools who aimed in part to develop schoolwide improvement strategies they could replicate across multiple schools.

Lessons from education policy implementation research during this period extended some of the past. For example, many studies of state policy implementation echoed previous waves' federal policy studies in their findings that mismatches between policy makers' and implementers' incentives impeded implementation (e.g., Anderson et al., 1987). However, more nuanced understandings of the significance of policies, people, and places also began to take shape. For example, McDonnell and Elmore expanded on the notion that differences in policy design matter to implementation by distinguishing policies by their "instruments" or tools (McDonnell & Elmore, 1987). They highlighted that policy instruments—mandates, incentives, capacity building, and systems change in particular—reflected different underlying assumptions about how to motivate implementers to change. They argued that an analysis of policy designs at this level would help reveal why policies of certain types were more or less effective. (See also Schneider & Ingram, 1990.)

A wider range of people emerged in implementation studies as consequential to implementation. For example, various researchers began to illuminate the importance of state educational agency leaders and staff as designers and implementers of policy (Cohen, 1982; Fuhrman, 1988; Fuhrman et al., 1988). While schoolteachers and principals long had been topics of study in the fields of teacher education and educational leadership, research explicitly located within the field of policy implementation began to explore how these school-based professionals shaped implementation processes and outcomes (Elmore & McLaughlin, 1988; McLaughlin, 1991a, 1991b; Rosenholtz, 1985). For example, Clune and others revealed policy implementation as a negotiated process involving at least the federal government, states, and local districts through which the terms of policy compliance were constructed (Clune III, 1983). Consistent with conventions of terminology in federal leg-

islation, studies during this period tended to refer to implementers by broad categories such as "teachers" and "state educational agencies" and not to explore how differences among individuals within these broad categories shaped implementation. Nonetheless, this research helped solidify a focus on implementers' agency as an important avenue for implementation research.

Studies in the 1980s also began to elaborate the places that mattered to implementation in several respects. For one, places included geographic locations and jurisdictions such as states that had received little attention in prior waves of reform and research. Studies revealed that these locations and jurisdictions varied in terms of their politics, culture, and histories in ways that helped to explain their differing responses to policy directives (Fuhrman et al., 1988, p. 64). Places also included new units of analysis such as "teacher networks" and "communities" and studies in this vein revealed these nonlegislated associations among implementers as powerful influences on implementers' work (Anderson et al., 1987; Fuhrman et al., 1988; Lieberman & McLaughlin, 1992; Little, 1984; Marsh & Crocker, 1991).

Some implementation researchers went so far as to make places rather than policies their main concern. That is, past decades' implementation studies generally asked whether or not a given policy was implemented. By the 1980s, a growing cadre of implementation researchers focused on high performing schools and asked: "What are the policy and other conditions in those places that explain that performance?" This approach of tracking backwards from practice to policy was a particular hallmark of the effective schools movement (Purkey & Smith, 1983; Sizer, 1986). Reinforcing this approach, McLaughlin argued that implementation researchers should move away from mainly trying to understand which policies get implemented to elaborating the various conditions that matter to enabling effective practice (McLaughlin, 1991b).

In sum, the history of education policy implementation research may be divided into three waves that correspond roughly with the decades between 1960 and 1990. These waves may be distinguished by changes in policy demands that grew progressively more varied and complex. Policy implementation research followed suit by beginning to highlight that variations in policies, people, and places matter to how implementation unfolds.

The Current State of the Field: Confronting Complexity

(W)e have learned that there are few "slam bang" policy effects. This is because policy effects necessarily are indirect, operating through and within the existing setting. Thus policy is transformed and

adapted to conditions of the implementing unit. Consequently, local manifestations of state or federal policies will differ in fundamental respects and "effective implementation" may have different meanings in different settings. (McLaughlin, 1991, p. 190)

Education policy implementation research over the past 15 years has continued to build on lessons and insights from previous waves. However, as with previous waves, two developments mark contemporary education policy implementation research as a new generation of implementation inquiry: focal policy designs that differ significantly from those of previous eras and growing attention to how policy, people, and places interact to shape how implementation unfolds. In addition, contemporary research aims to build knowledge about implementation processes in ways that mark a distinct epistemological departure from past waves' research. I discuss each of these three trends in the following subsections.

New Policy Designs

Contemporary education policies differ from those of all three previous eras in terms of their basic design. Policies with similar designs may be found in previous decades and policies with past decades' characteristics can be found throughout contemporary educational systems. However, policies with certain features have reached a critical mass in recent years and have come to constitute a distinct trend. Table 1.1 elaborates on these major policy design distinctions.

Goals
To elaborate, the goals sought by policies of various stripes have exploded to address systemic, deep, and large-scale educational improvement. To be sure, certain past federal policies aimed to tackle such ambitious problems as societal disadvantage. However, by and large, the formal goals of even those policies focused on schools' delivery of particular discrete programs, procedural changes in schools, and students meeting basic minimum standards. Title I of the Elementary and Secondary Education Act (ESEA) reflects this distinction. In its early years, Title I of ESEA aimed to reach the long-term goal of reducing poverty and disadvantage by achieving the short- and mid-term goals of encouraging schools to provide supplemental services to help low-income, low-achieving students meet basic minimum performance standards. By 2001, Title I of ESEA had become a leg in the so-called systemic or standards-based reform movement. In conjunction with other components of ESEA, now called the No Child Left Behind Act, Title I aims to help all students achieve to high-performance standards. In the short term, it focuses on helping to develop systems of schools with

TABLE 1.1
Changes in Policy Designs "Then" and "Now"

	Then (1965–1990)	Now (1990–early 2000s)
Change Goals	**TO HELP CERTAIN STUDENTS REACH BASIC MINIMUM STANDARDS through discrete programs and procedural changes** Focus mainly on adding programs on to the regular school day for specific categories of students.	**TO ENSURE ALL STUDENTS ACHIEVE HIGH STANDARDS through systemic, deep, large-scale change initiatives** Aim to change professional practice throughout schools, districts, and states and students' various communities.
Targets	**SCHOOL ACTORS** Mainly school staff.	**ACTORS SYSTEMWIDE & BEYOND** People and organizations at school, district, state, and federal levels. Actors across institutions that matter for student learning, such as families, neighborhood services organizations, and youth agencies.
Tools	**LIMITED** Mainly federal mandates and incentives and other instruments that assert top-down command-and-control relationships in hierarchical education systems.	**SIGNIFICANTLY EXPANDED** New tools reinforce traditional control models (e.g., threat, sanctions, high stakes) but also depart from them (e.g., capacity building, systems change, learning, and community). **MULTIPLE** Tools with conflicting logics or theories about how to effect change now more likely operate within single omnibus policies or otherwise converge on schools.

aligned content and performance standards, student performance assessments, data-driven decision making, penalties for failure to meet adequate yearly progress, major investments in supplemental services (many of which are provided outside public school systems), and school choice provisions.

Likewise, contemporary education policies aimed at fostering teacher professional learning communities and improved instruction in math, science, and reading move beyond past decades' effort to distribute programs and seek fundamental or core changes in the beliefs and practices of schoolteachers and district central office and state administrators (Coburn, 2003; Cohen & Ball, 1990; Elmore, 1996). Education policy goals sometimes extend to a scale beyond formal school systems to address the quality of learning opportunities in students' families and communities (Honig & Jehl, 2000).

Targets

The targets—people and organizations named in policy designs as those slated for change—once resided almost exclusively in schools. Contemporary policy designs now more routinely include targets who sit throughout and beyond formal educational systems. For example, as noted above, systemic reform initiatives focus on the decisions of leaders in schools, school district central offices, and state educational agencies consequential to the alignment of curricular content, instruction, and assessments. Likewise, state agencies and district central offices have launched accountability policies that place various demands on schools but that also call for the marshalling of their own staff to participate in implementation (Massell & Goertz, 2002). In other words, those targeted to implement educational policies may very well be the policy makers themselves.

Many policy designs also no longer exclusively focus on schools but rather now target various organizational actors across institutions that seem to matter for improved school performance including those in families, neighborhoods, businesses, community organizations, the courts, and service systems (Crowson & Boyd, 1993; Mawhinney & Smrekar, 1996). For example, federal policies related to science programs, bilingual education programs, after-school partnerships, and parental involvement ask schools to collaborate with families and various community-based organizations to develop and implement reform efforts (Honig & Jehl, 2000). Statewide school restructuring in Kentucky and districtwide reform in Philadelphia include community service organizations as key education policy implementers. In Chicago, low-performing schools on probation must utilize a school-level support system that includes "an external partner" such as a university or professional development organization (Burch, 2002; Finnigan & O'Day, 2003). For almost a decade, the Comprehensive School Reform Demonstration Program has required that schools work with school reform support providers in the implementation of comprehensive school reform designs (Berends, Bodilly, & Kirby, 2002; Bodilly, 1998).

Tools

Perhaps not surprisingly given these changes in goals and targets, the policy tools or underlying levers of change employed by policy makers have expanded significantly (McDonnell & Elmore, 1991). Prior to the 1990s federal, state, and other policy makers primarily relied on mandates and incentives —tools consistent with top-down command-and-control models of decision making. The 1990s and beyond have witnessed a broader set of policy tools in practice. For example, systems change tools (instruments that aim to effect change primarily by shifting authority among various parties) have appeared more prominently in recent years as part of school choice and site-based decision-making

policies (e.g, Bryk, Kerbow, Rollow, & Sebring, 1993), school-linked services efforts (Smithmier, 1996), and new small autonomous schools initiatives. Teacher professional development initiatives throughout educational systems often rely on capacity building tools (instruments that aim to build resources and capabilities for future use) to increase teacher supply and quality.

Still other contemporary policies seem to represent policy instruments not specifically identified in previous catalogs of education policy tools. Various school improvement planning and waiver processes support schools' collection and use of student performance and other data to develop strategies for educational improvement (United States Department of Education, 1998b; United States General Accounting Office, 1998). Accordingly, these policies employ what Schneider and Ingram referred to as "learning tools"—tools that reflect that policy makers do not necessarily know which strategy will improve outcomes and their willingness to fund implementers to invent such improvement strategies (Schneider & Ingram, 1990). The bully pulpit or hortatory policy tools—tools that rely on the sheer power of argument and persuasion—have grown in prominence especially in the context of the 1990s standards-based reform efforts that used such slogans as "all students can learn" and "it takes a village to raise a child" as primary levers of change (Smith, Levin, & Cianci, 1997). Arguably, standard-setting and credentialing constitute a distinct policy tool, particularly prominent in the 1990s and beyond, that relies on performance targets themselves as the main lever for change (Mitchell & Encarnation, 1984). Some high-stakes testing, reconstitution, and accountability policies stem from the premise that schools will reform and improve under threat of penalty and takeover to lengths that seem quite distinct from negative incentives (Mintrop, 2003; O'Day, 2002). Still other policies use the formation of "communities" or partnerships to leverage change in various settings (Rochefort, Rosenberg, & White, 1998).

Adding to the complexity of contemporary policy designs, policies with different tools or theories of change converge on schools, districts, and states (Chrispeels, 1997; Hatch, 2002; Honig & Hatch, 2004; Knapp et al., 1998; Newmann, Smith, Allensworth, & Bryk, 2001). Schools and other implementing organizations always have managed multiple policies at the same time. However, the diversity of policy tools simultaneously at play in contemporary public educational systems means that implementers now juggle an arguably unprecedented variety of strategies, logics, and underlying assumptions about how to improve school performance in ways that significantly complicate implementation (Hatch, 2002). For example, some schools participate in new small autonomous schools or site-based management initiatives that promise considerable new autonomy for schools over curriculum, instruction, and other operations and accordingly rely on systems change or learning tools as the main levers of change. These same schools may also face mandates to participate in dis-

trictwide reading programs or professional development opportunities. Accordingly, such schools face demands both to set their own curriculum *and* follow others' decisions about curriculum. The convergence of such demands that rest on fundamentally different theories of change can create significant confusion regarding who ultimately decides and, if not managed strategically, can frustrate educational improvement goals (Honig & Hatch, 2004).

New Approaches to Implementation Research

Whereas past implementation research generally revealed *that* policy, people, and places affected implementation, contemporary implementation research specifically aims to uncover their various dimensions and *how and why interactions among these dimensions* shape implementation in particular ways. I outline these dimensions in Figure 1.1.

Policy
Contemporary studies generally suggest that policy designs have three key dimensions—goals, targets, and tools—and aim to uncover how differences at this analytic level influence implementation. For example, researchers now commonly highlight that policies with *goals* related to the core of schooling— teachers relationships with students, their subject matter, and their work- places—pose fundamentally different implementation challenges than policies that seek more peripheral changes such as new course schedules or classroom
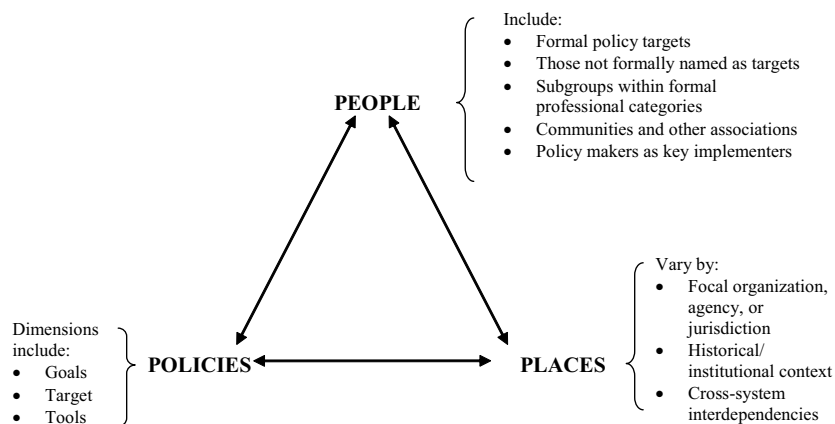

FIGURE 1.1.   Dimensions of contemporary education policy implementation in practice and research.

seating arrangements (Cuban & Tyack, 1995; Elmore, 1996; Lee & Smith, 1995; McLaughlin & Talbert, 2001; Siskin, 1994). Goals also differ by scope: Policies that aim to impact schools districtwide require a different degree of engagement by district central offices than policies that focus on a limited number of schools. Policies that focus on changes in the short term have different consequences in implementation than those that allow for a longer implementation horizon. For example, Hess highlights that accountability policies in general aim to deliver diffuse benefits over the long term; but because costs in the short term are so high, implementation of such policies typically meets strong immediate resistance, particularly among the communities the policies aim to benefit over the long term (Hess, 2002). Failure to attend to the different challenges and opportunities such policies present in short and long terms may significantly curb implementation.

Certain goals also are more or less attainable depending on implementers' starting capacity or current performance relative to the goal. For example, schools that are labeled as low-performing but that are on the cusp of meeting required performance standards face different challenges in reaching those standards than schools performing at lower levels (Mintrop, 2003). Similarly, adopting particular standards for the teaching of mathematics for some teachers may constitute a core change in their practice but for others such adoption may involve a more peripheral change (Hill, this volume).

Policy designers' choices of policy *targets* appear in implementation research as influences on implementation in their own right. For example, Malen in this volume highlights that those who stand to win or lose from particular policies significantly shape the mobilization of groups either in support of or against implementation. (See also Hess, 2002; Stone, 1998.) How various groups are named or labeled in policy designs sends signals about the targets' value in ways that significantly influence policy outcomes (Mintrop, 2003; Schneider & Ingram, 1993; Stein, 2004). For example, Pillow argues that the social construction of teen mothers as a target group within education policies has systematically denied them access to educational opportunity despite the provision of other resources (Pillow, 2004. See also Schram, 1995). Stein has revealed implementers' themselves as significant creators and reinforcers of group labels and has demonstrated how such labels may function to frustrate precisely the equity and other policy goals that implementers aim to advance (Stein, 2004; see also Datnow et al., 2001).

Tools also exert their own influences on implementation and have differential benefits depending on other implementation conditions. For example, in this volume, Coburn and Stein demonstrate how in some settings the implementation of teacher professional communities may be reinforced by central mandates whereas in other districts such mandates are unnecessary or prohibitive. (See also McLaughlin & Talbert, 2001.) The same accountability

policies are met with different degrees of resistance within states depending on teacher motivation, capacity, incentives, and other factors (Mintrop, 2003). Honig has highlighted that bottom-up reform initiatives as designed generally rely on systems change and learning tools and at a minimum can spark a rethinking of relationships between school district central offices and schools; however, their strength as levers of change seems to depend on supportive contextual conditions, the starting capacity of district central offices and schools, and the assistance of intermediary organizations (Honig, 2001).

People
The people who ultimately implement policy significantly mediate implementation in a wide variety of ways that have begun to take center stage in contemporary implementation studies. First, to be sure, researchers continue to examine how those targets formally named in policy designs respond to policy demands. But given the expansion of the types of formal policy targets noted above, a focus on targets now means that contemporary studies are more likely than those of past decades to consider a host of individuals both inside and outside the formal education system including parents, youth workers, health and human service providers, and comprehensive school reform designers to name a few (e.g., Honig et al., 2001). A battery of new policy initiatives and related research highlight school district central office administrators as key mediators of policy outcomes (e.g., Burch & Spillane, 2004; Hightower, Knapp, Marsh, & McLaughlin, 2002; Honig, 2003; Spillane, 1996).

Second, researchers also focus on individuals who are not formally named as targets in policy designs but who nonetheless participate in and otherwise influence implementation. For example, Shipps has shown how business leaders had a profound effect on the implementation of various Chicago reform efforts over the course of nearly two decades even though they were not named specifically in policy designs as targets (Shipps, 1997). City mayors play increasingly prominent roles in education policy implementation not only as policy makers but as primary influences on the implementation of policies passed by state and local school boards (Cuban & Usdan, 2003; Katz, Fine, & Simon, 1997; Kirst & Bulkley, 2000).

Third, past decades' research tended to focus on groups of implementers based on their formal professional affiliations (e.g., "teachers," "central office administrators") and to assume that such groups on the whole held certain interests, beliefs, values, ideas, knowledge, and other orientations that shaped their participation in implementation. Contemporary studies are more likely to probe differences among sub-groups within these broad categories. Louis has demonstrated that implementers' functional roles—such as stimulator, storyteller, networker, and coper—reveal important implementation dynamics not always obvious from formal employment categories such as "teacher"

(Louis, 1994a). Likewise, school district central offices are less likely to appear, in Spillane's words, as "monolithic" but rather to be viewed as consisting of frontline and midlevel administrators among others who each face different demands, opportunities, and constraints in implementation (Spillane, 1998b; see also Burch & Spillane, 2004; Hannaway, 1989; Honig, 2003).

Research in this vein sometimes highlights the processes whereby various individual and group orientations shape implementation. For example, Spillane, Gomez, and Reiser in this volume build on their previous publications to reveal that implementers' identities and experiences extend well beyond their formal professional positions; these authors elaborate the individual, group, and distributed sense-making processes through which implementers draw on various identities and experiences to shape their choices during implementation. In this view, opportunities for people—policy makers and implementers alike—to learn about policy problems, policy designs, and implementation progress essentially shape how implementation unfolds (Cohen & Hill, 2000; Louis, 1994a; O'Day, 2002; see also Honig, this volume).

Fourth, researchers have come to reveal that people's participation in various communities and relationships is essential to implementation. For example, researchers have shown that teachers within schools and districts are situated in professional communities that help shape their beliefs and worldviews and ultimately their interpretations of policy messages (Cobb, McClain, Lamberg, & Dean, 2003; Coburn, 2001a; McLaughlin & Talbert, 2001; see also Coburn & Stein, this volume). Hill in this volume argues that teachers and others belong to different discourse communities that significantly shape their responses to ambitious standards-based reform demands. Pollocks' school-based ethnography examines in part how such communities in schools influence how teachers talk about and act on potentially racially charged issues in their educational improvement efforts; she recounts multiple instances in which teachers aimed to use race-neutral language in an effort not to generate negative racial stereotypes of students but in the process actually reinforced the very categories they sought to avoid (Pollock, 2001). Smylie and Evans (this volume) discuss how other forms of social interactions and trusting relationships both fuel and frustrate reform (see also, Knapp, 1997).

Importantly, contemporary education policy implementation research also continues to move beyond traditional distinctions between policy makers and implementers and teaches that both are consequential sets of people who shape how a policy is designed and implemented. For example, in this volume, Malen reveals how policy design and implementation are overlapping processes that unfold in a series of games through which those in formal designated policy-making roles and those in implementation roles shape both processes—even in cases such as school reconstitution in which traditional top-down control dynamics are a fundamental aspect of policy design. Also in

this volume, Datnow elaborates on her prior work to reveal education policy systems as nested systems in which local, state, and federal actors play key roles in co-constructing policy design and implementation. Honig in this volume and elsewhere flips these traditional policy-making and implementing roles on their heads to reveal implementers as significant drivers of policy and policy makers as key implementers.

Places

Contemporary researchers build on Wave 3 studies in elaborating dimensions of places as fundamental to implementation outcomes in several respects. First, as in the past, these researchers find that governmental organizations such as state educational agencies are important settings where implementation unfolds (Hamann & Lane, 2004; Lusi, 1997). As noted above, a growing cadre of researchers explores school district central offices as central implementation sites. However contemporary studies are more likely than those of the past to probe differences in how formal organizational systems operate in the implementation process. For example O'Day revealed how district central offices function as complex systems in ways that lead to particular implementation processes (O'Day, 2002). Similarly, an emerging literature on *urban* districts teaches that these districts have particular political and institutional resources for implementation that mark them as a distinct subset of districts (Kirst & Bulkley, 2000; Orr, 1998; Stone, 1998).

Second, many contemporary researchers name their districts and states in their studies in an effort to build a body of knowledge about how implementation unfolds in these locations and to call attention to how deep-seated historical institutional patterns shape implementation outcomes. For example, the implementation literature now includes a substantial substrand concerning Chicago school reform (e.g., Bryk & Sebring, 1991; Bryk et al., 1993; Katz et al., 1997; O'Day 2002). In this volume, Dumas and Anyon reveal implementation as significantly shaped by race- and class-based tensions that may seem familiar nationwide but that are deeply rooted in the local educational, economic, and political institutions of New Jersey's cities (see also Anyon, 1997).

Some contemporary research—what I call "place-based studies"—builds on the effective schools tradition in that it focuses on particular geographic locations as a main concern and asks which policies and other conditions account for education outcomes in those settings. For example, Anyon's research for *Ghetto Schooling* began as an investigation into the school experiences of students in Newark, New Jersey's elementary schools and ultimately revealed important lessons about Marcy School's main policy implementation challenge at that time—school restructuring (Anyon, 1997). Orr's examination of Baltimore addressed the trajectory of multiple policies in which the city was engaged over more than a decade to highlight not only how particular policy

initiatives fared but also more broadly how Baltimore as a community and urban system managed change (Orr, 1998, 1999).

Some of this "place-based" research has shown that schools are inextricably linked to other places—namely the urban institutions they operate within and alongside—despite Progressive Era reforms and other efforts to separate "school" from "city" (Bartelt, 1995; Yancey & Saporito, 1995). Such interdependencies mean policy makers, researchers, and others should cast a broad net when considering which places matter to implementation. On the flipside, they should aim to uncover the educational impact of policies in other sectors such as community development, health care, and social services (Anyon, 2005).

Overall, these three dimensions of implementation—policy, people, and places—come together to form a conception of implementation as a highly contingent and situated process. In this view, the benefits or limitations of one dimension cannot be adequately understood separate from the other. For example, as noted above and in the chapters by Hill and Spillane and associates, how teachers and central office administrators make sense of standards-based curricular reform depends on the policy tool employed within the given policy design, their own prior knowledge and experiences, and the broader institutional setting in which they operate. Maryland superintendents such as the one featured by Malen (this volume) seem to face different obstacles to implementing high-stakes accountability policies than the Chief Executive Officer in Chicago (Finnigan & O'Day, 2003) or state-level leaders (Massell, 2001). In these studies, variation in implementation outcomes is not the exception but the rule and researchers aim to understand how different dimensions of policies, people, and places combine to shape implementation processes and outcomes.

In the process, this generation of research is beginning to move from a static to a dynamic and contingent view of implementation capacity. That is, researchers are less likely than in the past to view capacity as a set of resources with a fixed value and more likely to conceptualize capacity as including a variety of supports whose value depends on what particular people in certain places are trying to and are currently able to accomplish. For example, past generations of researchers tended to suggest that resources such as strong leadership or increased funding were universally important to implementation and aimed to identify the specific dimensions of leadership and the total amount of funding necessary for implementation across sites. Contemporary researchers more often highlight that the importance of such resources varies depending on many factors, including what people already know and can do, the historical patterns of opportunity in particular jurisdictions, and the stakes associated with implementation outcomes. As Smylie and Evans demonstrate in the context of the Chicago Annenberg Challenge (this volume), social

capital—defined in part as trusting relationships among key actors—is often highlighted as essential to implementation; however, in their case, social capital did not support implementation across the board and in some instances was actually marshaled against policy goals. Loeb and McEwan (this volume) cite numerous examples of the relative value of various resources in the context of implementing school choice and accountability policies.

In part by broadening their view of capacity, contemporary researchers have revealed a range of previously hidden or unnoticed resources necessary for implementation including various forms of cultural, political, social, technical, and institutional capital (Honig, this volume; Orr, 1998; Schram, 1995; Spillane & Thompson, 1997). The ability to learn new ideas appears as a distinct resource essential to implementation in complex policy environments; some researchers argue that this ability is so important that policies should be evaluated by the extent to which they build policy actors' capacity to learn (Cohen & Hill, 2000; Elmore, 1983; McLaughlin, this volume) and otherwise to evaluate, negotiate, and craft the fit between policies, people, and places over time (Honig & Hatch, 2004).

Distinct Approach to Knowledge-Building

The discussion above reveals that contemporary education policy implementation research also may be distinguished epistemologically—by its orientation to the nature of knowledge and knowledge-building about implementation. First, as already noted, contemporary researchers are less likely than those in past decades to seek universal truths about implementation. Rather, they aim to uncover how particular policies, people, and places interact to produce results and they seek to accumulate knowledge about these contingencies. These researchers seem to take to heart McLaughlin's admonition that "generalizations decay"—few if any findings hold true across all contexts or across all time. For example, in the early years of education policy implementation research some researchers argued that policy makers could not mandate what matters to educational improvement—that mandates were insufficient instruments for changing teachers underlying beliefs about and engagement in their work; however, over time researchers have shown that sometimes policy makers could mandate what mattered—that given certain conditions mandates did in fact leverage core changes in schools (McLaughlin, 1991b). This orientation to uncovering contingencies—what I refer to here as confronting complexity—stems not from a lack of rigor or scientific-basis for educational research but rather from the basic operational realities of complex systems in education and many other arenas. In light of these realities, education policy implementation researchers aim to uncover the various factors that combine to produce

implementation results and to accumulate enough cases over time to reveal potentially predictable patterns (Majone, 1989).

Second, contemporary research increasingly reflects the orientation that variation in implementation is not a problem to be avoided but part and parcel of the basic operation of complex systems; variation should be better understood and harnessed to enhance the "capacity of program participants to produce desired results" (Elmore, 1983, p. 350; see also Honig, 2003; O'Day 2002). This view stems in part from contemporary research on student and teacher learning that suggests one size does not fit all when it comes to educational improvement, especially in diverse urban school systems; supports provided to students should vary depending on what students, teachers, and other implementers already know and can do (Darling-Hammond, 1998; Villegas & Lucas, 2002). This orientation also reflects relatively recent education policy implementation findings about sense making, interpretation, and learning as unavoidable dimensions of implementation processes. Studies in this vein uncover how individual, group, and cognitive processes contribute to implementers' variable policy responses and, for certain implementers in some settings, the achievement of policy goals (Spillane et al., 2002).

Third, contemporary research is more deeply theoretical than in past decades in keeping with conventions that define rigorous quantitative and qualitative research. As the subsequent chapters demonstrate, theory provides criteria for site selection, guides data collection and analysis, and, importantly, helps explain why certain interactions among policy, people, and places contribute to particular implementation outcomes. Such research aims not to develop a universal theory about implementation as an overall enterprise but to use theory to illuminate how particular dimensions of policies, people, and places come together to shape how implementation unfolds.

The theories on which education policy implementation researchers now draw come in part from disciplines familiar in implementation arenas such as political science and economics but researchers have begun to apply those theories in new ways. For example, economic analyses of policy implementation have tended to assume a singular implementation actor with certain almost automatic responses to policy demands. But as Loeb and McEwan reveal in their chapter, many contemporary applications of economic theory to implementation include significantly deeper explorations of how implementers' agency and context matter to implementation. At the same time, the field of education policy implementation research has expanded to embrace theoretical constructs from disciplines not traditionally applied to implementation such as those from anthropology, cognitive science, psychology, and learning theory. Critical and sociocultural theories have contributed to particularly vibrant lines of analysis (Anyon, 1997; Lipman, 2004; Stein, 2004; Sutton & Levinson, 2001).

Perhaps in a related development, qualitative research designs and methods have become important sources of knowledge for implementation researchers. In particular, strategic qualitative cases—cases that provide special opportunities to build knowledge about little understood and often complex phenomena—have long informed implementation in other fields and seem to be becoming more standard fare within education. Such methods and research designs, especially when well grounded in theory, have allowed contemporary researchers to elaborate the dimensions of and interactions among policy, people, and places that comprise implementation in contemporary educational systems. In fact, the more complex portrait of implementation processes advanced here may have become possible only recently thanks in part to the use of theoretically grounded qualitative methods for capturing such complexity.

PROGRESS OR POOR GUIDES FOR PRACTICE?

Critics of contemporary education policy implementation research and education research more broadly occasionally argue that educational researchers paint such complex portraits of the world that they obscure what is implementable and what works and serve up poor guides for practice. To the contrary, the chapters in this book suggest a deep interest on the part of researchers, their funders, and others in building a base of knowledge that can guide practice in informed, responsible, and productive ways. These chapters reflect the orientation that those who aim to produce "usable" knowledge should seek to highlight and sort through the complexity that is fundamental to implementation in contemporary education policy arenas—that implementation studies should keep pace with and reflect, not minimize, or ignore the complexity of contemporary policy demands and implementation processes. Accordingly, critics might consider that confronting complexity has been a positive development. As organizational theorists have long shown, complexity in its various forms—including variation—can serve as a stimulus for innovation and improvement especially given the diverse and sometimes unpredictable circumstances under which educational leaders routinely operate.

This volume represents a small but important sampling of how, over the course of the past four decades, education policy implementation researchers increasingly have confronted the complexity of implementation. In her important concluding chapter to this volume, McLaughlin elaborates how researchers can continue to build on this tradition. She emphasizes in particular that in confronting complexity future education policy implementation research should aim to deepen knowledge about how entire systems—those within and beyond the formal educational system—matter to educational outcomes and about the learning processes that are and should be fundamental to

how implementation unfolds. In short, future researchers should delve deeper into the complexity.

The view of implementation advanced here increases the urgency for educational leaders to become more savvy consumers of research. It suggests that at this stage of knowledge development about education policy implementation leaders should look to research not for prescriptions or to light a direct path to improvement in their own local communities. Rather, they should mine the research for ideas, evidence, and other guides to inform their deliberations and decisions about how lessons from implementation research may apply to their own policies, people, and places. For example, leaders could ask whether the transformation of large high schools into multiple smaller schools contributes to educational improvement. But a search for answers would yield many studies that reveal positive results and also some important confounding cases. Rather, a more productive question might be: under what conditions within my own district, school, or organization might small schools yield positive results for my particular students? When viewed in this way contemporary education policy implementation research seems to reflect an underlying belief that educational leaders and other practitioners are willing and able (or can become willing and able) to confront complexity and to draw on a variety of increasingly nuanced research findings and other evidence to help inform their own work. This then is an optimistic book.[1]

NOTE

[1]This line refers in part to Bardach's (1980) seminal work, *The implementation game*. The introductory chapter of Bardach's book summarizes and generally laments the various ways implementation deviates from policy design and concludes: "This is not an optimistic book" (p. 6).