care are met; and that provide education so all people can realize their potential. It is both unjust and unwise not to address gross inequality of opportunity and wealth. To neglect these issues deepens poverty, social conflict, and crime; it does not advance prosperity.

These brief descriptions inadequately capture the sophistication of the intellectual arguments in the literature on these thorny issues and the range and nuances of view and opinion between these poles that exist in the thinking of individuals. But they may serve the purpose of illustrating how different views of the human condition and of the causes and "cures" of human problems come into play in the policy and political arena.

Although the thinking of some people may be well-described by the "left" or "right" viewpoints presented previously, many people, perhaps most, acknowledge that each perspective contains some valid points. Governmental policy in democracies typically reflects compromises forged through the resolution of conflicts among particular interests; a blend of views concerning inequality and social justice, cultural issues; and the effectiveness of markets, bureaucracy, and hierarchy in solving social problems.

What Education Studies Are Influential in Political Environments?

In "Education Research That Matters," the concluding chapter of *Education Research on Trial,* Pamela Barnhouse Walters and Annette Lareau[22] assess the influence of different contributions to education research by employing a number of perspectives and indicators. Unsurprisingly, the values and purposes of different users of research matter greatly.

In their own research and writing, educational researchers most frequently cite scholars who have made important theoretical or methodological contributions. The top three scholars cited in a review of 129 education journals over 20 years were Jean Piaget (3,816 citations of work emphasizing stages of cognitive development in young children), Albert Bandura (3,338 citations of work on social cognitive theory and self-efficacy), and John Dewey (2,847 citations on theory, democracy, and education). Others in the top 13 most frequently cited scholars (1,500 citations or more) made contributions illuminating issues such as learning styles, impact of social class, theories and measurements of intelligence, self-concept, motivation, and other factors.[23]

Another indicator of influence, the Distinguished Contributions to Research in Education Award of the American Educational Research Association (AERA), similarly reflects the inclination of professional education researchers to value broad theoretical and methodological contributions. Of some 44 awardees from 1964 to 2007, Donald Campbell and Julian Stanley in 1980 and 1981 were honored for their contributions to the theory and methodology of experimental and quasi-experimental research. But Walters and Lareau suggest than none of the awardees were "primarily known for *conducting* random-assignment evaluation studies in education or other kinds of assessments of the efficacy of particular educational programs or interventions." Instead, they suggest that the AERA awardees were honored for focusing on core issues of broad significance that "modified, challenged, or advanced the literature" of the field.[24] Walters and Lareau argue that, even though its direct relevance to policy may not be immediately evident, such scholarly research has been and remains consequential for policy and practice.[25]

Policy professionals, however, may not readily recognize the influence of the scholarly contributions highly valued by the research community. A 2006 study by the Editorial Projects in Education (EPE) sought to identify studies, organizations, people, and information sources that were most influential in educational policy during the prior decade.[26] The studies were rated using three sources of information, a survey of expert ratings, news coverage, and scholarly citations. The EPE study (also analyzed by Walters and Lareau) found that the nature of influential studies varied substantially.

By a wide margin, the two most influential studies (according to all three indicators, expert survey, news citations, and scholarly citations) were not discrete research projects but assessment reports, the *National Assessment of Educational Progress (NAEP)*[27] and *Trends in International Mathematics and Science Study (TIMSS)*.[28] Another category of influential studies included highly publicized commission reports or collaborative national panel studies: *Report of the National Reading Panel; Preventing Reading Difficulties in Young Children; How People Learn: Brain, Mind, Experience, and School; What Matters Most: Teaching for America's Future;* and *Ready or Not: Creating a High School Diploma That Counts.*[29] One experimental study, Tennessee's *Student/Teacher Achievement Ratio (STAR),*[30] made the list of 13 influential studies. The remaining influential studies on the list consisted of a body of work of individual scholars (and one organization) on particular topics: William L. Sanders on value-added methodology; the Education Trust on teacher quality; Jay P. Greene on high school graduation rates; Paul E. Peterson on school choice and vouchers; and Richard F. Elmore on school reform.

All of these studies have attracted the interest of policy makers and educators, and in different ways have shaped policy and practice. From the vantage point of a single observer, however, the impact of these studies on the behavior of policy makers has been quite different among the categories.

Over an extended period of time the assessment projects, NAEP and TIMSS, have attracted the attention of policy makers and practitioners and prompted them to take action to improve outcomes. No assessment is technically perfect, of course, so they have not been accepted uncritically. But these assessments are technically robust enough to be taken seriously as a continuing benchmark of progress or a reason for concern. Because they do not in themselves provide much guidance for what policy makers or practitioners should do to improve outcomes, the responses to these assessments have varied greatly. But they have provoked action.

As one might expect, the commission reports and collaborative panel studies generally were quite impressive, thoughtful, wide ranging, and balanced. These efforts typically draw on the best thinking and research in the field, and they have likely shaped the choices policy makers and practitioners make in their efforts to improve outcomes. The contributions of such efforts should not be underestimated, but breadth and balance in a report is not conducive to crisp, causal impact. It is common for both practitioners and policy makers to use the phrase "another commission report to gather dust on the shelf," not because the reports themselves are poor quality, but because it is not easy to implement broad, multifaceted recommendations.

The impacts of the remaining "influential studies" vary in interesting ways. Although its findings were not universally accepted (perhaps largely due to the cost implications), the Tennessee STAR experiment prompted policy interventions to reduce class sizes, particularly in California. The unanticipated effects of creating a teacher shortage and reducing the average quality of the teaching force apparently diminished the positive effects of the intervention.[31] Class size remains a controversial policy issue.

The nature and extent of influence of the individual scholars (plus the Education Trust) on particular issues seems to vary according to the topic and the nature of the research. Sanders's work on the educational gains associated with particular teachers (value-added) attracted and still attracts great attention. The use of student learning outcomes as a factor in teacher evaluation and performance continues to be on the policy agenda, but the implementation of this approach has been slowed by persistent questions about the validity and reliability of "value-added" assessments for particular teachers. The Education Trust's work documenting the tendency for teacher placement policies to compound the disadvantage of underprivileged students continues to generate policy attention, even though the obstacles to change remain substantial.

Greene's work on high school graduation rates focused on particularly egregious inconsistencies in data collection and reporting. This, in comparison to other issues, was a relatively simple problem, and it has largely been solved. Greene was not alone in focusing on the problem, but he deserves a good share of the credit for stimulating a solution. Peterson's work on school vouchers has played a role in the debate concerning the role of the marketplace in public education, but the debate continues without a clear resolution in sight. Elmore's sophisticated work on school reform (especially, but not limited to, the role of instructional leadership) has certainly caught the attention of policy makers and practitioners but evidently not in entirely satisfying ways. His most recent writing reflects deep skepticism about the potential for the reform of established schools and institutions to improve student learning.[32]

Walters and Lareau conclude their essay by describing educational research in two categories—research that has had an impact on future research and research that has had an impact on policy. The Coleman Report on the effects of socioeconomic status on school outcomes and Bandura's work on social cognition are offered as examples to illustrate the research-influential category. To illustrate the policy-influential category, they suggest the STAR class size study and Peterson's work on vouchers. The policy-influential category is, in their words, "often very focused and narrow, often a theoretical, timed to mesh with a new policy interest, consistent with the political agenda of existing interest groups, [and] sometimes seen as high quality and sometimes not."[33] In contrast, they characterize the research-influential category of research as broader, informed by theory, timed to respond to developments in the scholarly field, rarely connected with a political agenda, and widely viewed as high quality.

I do not quarrel with the general description of Walters and Lareau's two categories; they are useful in the same ways my previous description of the political left and right may be useful, but they risk being too categorical. In a field such as education, where the primary purpose is not to gain abstract knowledge but to transmit knowledge and skill, it is a problem if the research community is, or even appears to be, excessively self-referential on the one hand or driven by political agendas on the other hand. The remainder of this chapter explores how research and evidence have and might further engage directly and improve the outcomes of policy and practice.

Consequential and Less-Consequential Research

Walters and Lareau's chapter assessing the influence of various research efforts on policy makers and other researchers is a good point of departure for considering how research and evidence can be *more* than influential. How can research and evidence help policy and practice actually improve? This section of the chapter will explore the difficult task of using research and evidence to help policy makers become more successful in reaching their goals. It will consider three challenges: (a) resistance to dissonant information; (b) the difficulty of sorting through cause and effect connections due to complexity and change over time; (c) the difficulty of tailoring inquiry to the structure, dynamics, and power distributions of the decision-making system.

Resistance to Dissonant Information

The resistance of policy makers (and humans generally) to dissonant information is a significant obstacle to the improvement of policy through inquiry and evidence. People commonly have relatively settled and divergent views on the major issues involved in making policy. While "research" is frequently employed as a tool of argument and advocacy, in politically competitive environments the goal of a "researcher duelist" often is to design and implement a "killer" study.

In organizations with an overriding, straightforward purpose such as generating profits, a "killer" study may be feasible under some circumstances. In government and nonprofit organizations with complex purposes and competing worldviews, debates over purposes and strategies for achieving purposes have rarely been settled by a research study. Typically, a sizable fraction of decision makers on both sides of an issue will be quite settled in their views and resistant to all but overwhelmingly conclusive research. For such decision makers, the most persuasive studies supporting the opposing side of an issue are more likely to generate energetic rebuttals rather than the modification of their views. The issues of school vouchers and class size reduction offer many examples of this dynamic.

More fertile ground for research contributions lies in the middle of the political spectrum where decision makers are more open to dissonant information. Studies that are visibly "advocacy research," tilting toward either end of the political spectrum or one side of an issue, have no chance of persuading those who are firmly opposed to the arguments of the studies, and they even have difficulty influencing those in the middle ground. In my view rigorous studies that carefully avoid the appearance of predisposition and bias have more power to shape opinions in the middle of the political spectrum.

Part of the solution to this problem lies in the framing of a study. Studies that focus narrowly on a politically contentious intervention or policy issue will inevitably be viewed as advocacy research, welcomed and applauded by the already converted, and attacked by established opponents. Broader studies that are visibly exploratory are more likely to have an educational impact across the political spectrum, or at least the central part of it. Such studies may not have a dramatic immediate impact, but they are more likely to move policy over time than research focused on polarizing issues.

A good example of such research is *Crossing the Finish Line: Completing College at America's Public Universities* by Bowen, Chingos, and McPherson.[34] This study examined previously untapped institutional data to explore the factors associated with successful completion (or noncompletion) of degrees by students in public universities. Its database was large enough to be persuasively representative of the entire sector, and it rigorously explored the role of academic preparation, financial assistance, and other factors in predicting successful college completion. Perhaps unsurprisingly, but convincingly, it found that financial assistance has the greatest impact on the success rates of lower-income students. More surprisingly to some, it found that the completion rate of low-income students is lower when they attend institutions that are less selective than other institutions to which they could have been admitted. "Undermatched" students do worse than similar students attending a more selective, perhaps more demanding, institution.

This study contributed to policy and practice by confirming existing evidence that financial assistance is most powerful in advancing completion when it address genuine financial need and that academic preparation is a necessary complement to financial assistance. Its finding regarding undermatched students suggests both that the educational culture of less-selective institutions might usefully become more demanding and that it is harmful, not helpful to funnel able low-income students into less demanding, less-selective institutions. Finally, it demonstrated the utility of mining administrative, student-level databases to learn ways of improving policy and practice.

Resistance to dissonant information is also less of a problem when information is presented that adds fresh, policy-relevant perspective on issues that are not politically fraught. For example, challenging, time-consuming paperwork is unpopular all across the political spectrum. An experiment using H&R Block tax preparers to help low-income families complete the federal financial assistance application (previously cited in Chapter 2)[35] confirmed that the form itself is a barrier to participation in higher education. This research surely has contributed momentum to efforts to simplify the application process for financial assistance.

Advocacy research, of course, will always exist, and it has its place. But researchers who want to move policy are likely to be more successful if they find ways of working around resistance to dissonant information.

Sorting Through Elusive Connections Between Cause and Effect

The complexity of social problems and social interventions typically makes it quite difficult for social research to be conclusive. Ray Pawson's discussion of mentoring programs summarized in Chapter 2 is a good example of a situation in which a constellation of requirements must be addressed in order to achieve the desired objective of an intervention. Different programs might be quite successful in meeting one, two, or three such needs while failing to meet a fourth or a fifth. When programs are evaluated on "bottom line" results, significant achievements could be masked by partial failure.

The STAR research focused on class size discussed previously is an example of a slightly different sort. In this case, evidence suggests that smaller class sizes help improve teaching and learning, but the experimental condition of "other things being equal" cannot be assured in the real world. More teachers to staff smaller classes might not be as skillful as a smaller number of teachers with larger classes, or the financial costs of smaller classes might result in trade-offs that reduce the quality of other important factors, including academic supports or even adequate compensation to attract and retain strong teachers.

Certainly "bottom line" results are what ultimately matter, but for most complex problems, simple interventions rarely yield impressive results. This is why the simple intervention of charter schools or other silver bullet "solutions" fail to yield the results that would gratify their advocates or silence their critics. Charter schools alone cannot solve all of the problems that impede the success of disadvantaged students, even though they might solve some of them. Many other factors that vary among charter schools (as well as among non-charter public schools) play a role in their success or failure.

A good example of the challenges complexity offers to public policy analysis can be found in a retrospective analysis of Lyndon Johnson's War on Poverty. Ronald Reagan famously quipped, "Lyndon Johnson declared war on poverty, and poverty won." Reagan was clearly right if the measure of victory is the total eradication of poverty. But a recent analysis of the full spectrum of issues tackled by the War on Poverty suggests it achieved some significant victories as well as suffered some failures. Also, changes in the underlying challenges facing poor people partially thwarted the strategies adopted by the Johnson administration.[36] The difficulty of assessing social policy is compounded both by complexity and by changes over time in the surrounding environment.

The War on Poverty included multiple campaigns. One campaign, reducing poverty among the elderly, was clearly a success. Comparing the poverty rate in 1962 to the rate in 2012, Bailey and Danziger found the changes displayed on Table 6.1 for different groups within the national population.[37] (Because this analysis excludes food stamps, housing allowances, and the earned-income tax credit, it actually understates poverty reduction during this 50-year period.) Although the general direction of change is

TABLE 6.1
Results of the War on Poverty (Based on Income Below the Poverty Line)

Population Component	*Percentage in Poverty, 1962*	*Percentage in Poverty, 2012*
All groups	20%	14%
Elderly	30%	9%
Blacks	40%	22%
Children under the age of 18	22%	20%

positive, a number of contextual changes over this period worked to increase poverty despite the governmental interventions:

- Wage inequality increased from 1963 to 2009.
- Male incarceration rates, nonmarital childbearing, divorce rates, and female-headed households all grew.
- The non-elderly poor are increasingly Latino; the percentage of non-elderly poor who are Black or White has dropped significantly.[38]

These environmental changes would seem to be the major factors explaining why childhood poverty has decreased so little, if at all.

When individual programs are examined, it is typically difficult to pin down their effectiveness due to the number of external variables with an impact on their intended outcomes. The initial academic gains associated with Head Start seem to fade over time. Is this due to weakness in the program, or due to the quality of schools serving poor children as they progress through the elementary grades?[39]

Significant need-based aid for higher education has contributed to higher participation, but college costs have risen faster than Pell Grants and large disparities by race and income remain.[40] Title I support for the K–12 education of disadvantaged children has reduced the gaps between inputs and outputs in rich and poor states and schools, but the gaps are still quite large.[41] (In *The Ordeal of Equality*, David Cohen and Susan Moffitt suggest that the failure to increase the capabilities of teachers and school leaders for educating disadvantaged children and the fragmentation of school policy due to local control largely explain the failures of Title I.)[42] Finally, workforce development programs had some positive effects, which varied for different populations, but rapid increases in skill requirements in the job market made it much more difficult to train and place in jobs the hard to employ.[43]

Since launching the War on Poverty, the nation has reduced elderly poverty and infant mortality while substantially increasing access to medical care, housing assistance, food assistance, and life expectancy.[44] It seems reasonable to attribute at least some of these gains to the interventions of the War on Poverty, but it is not possible to trace a causal chain that would convince all skeptics of the benefits of the effort and investment. And despite all these efforts, childhood poverty apparently has not improved at all.

What are the implications of complexity for researchers and policy makers? First, it is unrealistic to expect simple, one-dimensional interventions to yield dramatic solutions for complex problems. Second, all the elements of complex, multifaceted interventions can rarely be implemented consistently with necessary quality in single locations, and, perhaps, never with fidelity across locations. Third, both time and the capability of adapting in response to learning or change are required to make progress on social problems. The expectation or requirement that all social programs must be proven to "work" through experimental studies (as recently suggested by policy analysts Peter Orszag and John Bridgeland, each of whom, respectively, had major roles in the Obama and George W. Bush administrations)[45] inevitably leads toward the conclusion that nothing "works."

This is not to suggest, of course, that "everything works" or that policy should not be tested by analysis. It would be more accurate to say that no policies "work" all by themselves. Policy makers frequently have control over part of the solution to social problems, typically the allocation of financial resources, but they rarely can control or even materially influence many of the other factors important to the outcomes they seek. Effective policy is likely to require a sensitive balance between necessary interventions into local conditions along with respect for practitioners and the challenges of practice. Policy should avoid regulations that substantially constrain the ability of practitioners to respond to local context and complexity. Research and analysis can and should play a role informing both the macro levels of policy and the micro levels of practice, but it takes recognition of complexity and deliberate, wise decisions to well synchronize policy and practice.

College student financial assistance is a good example of the failure to meet this challenge. The U.S. government has made a significant investment in Pell Grants and in guaranteed student loans to enable and encourage widespread participation in postsecondary education. The policy objective has been to enable all qualified and motivated low- and moderate-income students to obtain a postsecondary education. The actual work of enrolling and educating them occurs in states and institutions.

The federal programs (supplemented by state and institutional programs) have clearly made a significant impact on postsecondary participation and completion; however, in too many cases the policies have been inefficient and ineffective on two related dimensions: adequately meeting financial need and avoiding the wasteful expenditure of funds. In many cases financial aid is insufficient to meet the need, and in other cases aid is provided where government support is not needed or where students and institutions are not prepared to achieve the program's objectives. Let me elaborate.

Despite a large federal investment, it is evident that the lowest-income students rarely have adequate aid to attend college full time. A full Pell Grant, added to the earnings from substantial part-time work during the academic year and full-time work over the summer, is just adequate to cover living and other nontuition expenses, which range from $12,000 to $18,000. Unless the states or institutions themselves provide enough assistance to cover tuition costs (a few do, but most do not), very low-income students have two unattractive choices: work full time while attending school part time, or borrow all the funding required for tuition expenses. A common choice has been to work full time and attend school part time; failure to complete is frequently the result.

Part of the reason financial assistance is inadequate for so many low-income students is that a substantial portion of the public investment in student assistance is unproductive in terms of the main program objective. Some states emphasize merit aid, allocated without regard to financial need, and even much of the federal aid is allocated to tax credits and loans available to higher-income families. These investments are defended because they reduce the cost of college to middle-income families and, in some cases, induce strong students to go to college in their home state. But on the margin, merit aid to middle- and upper-income families influences where students go to college, not whether they will go. Virtually all of these students would enroll without the aid.

Another reason funds are inadequate (and a political weakness of the program) is that too many grant recipients, poorly prepared to succeed in postsecondary education, fail to succeed after enrolling. Some institutions (both nonprofit and for-profit) are complicit in this failure by taking money allocated for the education of low-income, "high-risk" students without the ability and commitment necessary to help them succeed. During the George W. Bush administration, Congress tried to address the academic preparation problem by providing federal financial incentives for taking a rigorous college preparatory curriculum in high school. In principle this was a good idea, but it proved cumbersome and unworkable at the federal level. The Obama administration considered dealing with the complicit institution problem through a federal system to rate institutional performance, but eventually decided simply to expand the information available to the public without formally ranking institutions.

In this case, as for many social problems, policy and practice are interdependent. A policy focusing resources on a problem cannot succeed without effective "on the ground" practices to address the problem. At the level of practice, schools and colleges (and their self-regulating accreditors and professional associations) need to improve student preparation and institutional effectiveness. Both resources to support practice and resources and policies designed to improve practitioner capability are required.

The American student financial assistance system has, by and large, been successful, but it has not fully met its objectives. At the policy level, it is quite clear that substantial state and federal sums have been allocated to make college more affordable (often for middle-income students) without adequately enabling and increasing postsecondary educational attainment for large numbers of lower-income students. At the level of practice, educators have failed to assure that a large number of students who might benefit from the program have the right level of preparation and support to succeed. The analysis of data and evidence can best increase the success of policy and practice when the complexity of problems and the mutual interdependence of policy and practice are recognized.

Tailoring Inquiry to the Decision-Making System

It has often been frustrating to be a policy professional committed to using evidence and analysis to inform and improve policy. My greatest disappointment is not the resistance of policy makers to dissonant or challenging information; that is natural and inevitable, simply a problem to solve. My greatest disappointment is the number of missed opportunities to make important contributions by talented, caring scholars. These missed opportunities most often come from a failure to recognize and acknowledge the limitations of research, especially traditional research seeking to establish general knowledge, and a failure to understand how policy decisions are made.

Research or evidence has the greatest power to influence decisions in government or organizations when it is tailored to the structure, dynamics, and power distributions of the decision-making system. Let me suggest five characteristics of research efforts that are successfully "tailored" to be consequential. Consequential research and evidence is more likely to be

1. focused on realistic, potentially consequential operational options in decision making;
2. focused on salient, significant issues, where the demand for improvement is high, and the consequences of inaction are serious;
3. reasonably accessible to intelligent, educated decision makers who may lack sophisticated knowledge of statistical analytical techniques;
4. based on credible *local* data about the places to be affected by the decision or directly and clearly relevant to these places; and

5. studies where the likely effect sizes are neither trivial in size nor a close call in terms of differences among alternatives if causal effects are inferred.

Let me elaborate on each of these in turn.

Realistic, operational decision options. Scholars, too often graduate students looking for a dissertation topic, frequently seek to discover whether some broad, general characteristic of a state (the form of governance employed, the use or nonuse of formulas in the budgeting process, the allocation of funds to student assistance, the existing level of educational attainment, etc.) can explain an important part of the variance in a valued outcome, such as the generosity of state support, student enrollment, or student attainment. I have read many such studies (including one chapter in my own dissertation), and I have never observed, nor can I imagine one of these studies influencing a policy decision.

The first problem is that studies of the impact of general characteristics, no matter how Herculean the effort, cannot control for enough of the relevant variables to produce a sizable effect. Even if a statistically significant result is obtained, it rarely (if ever) explains enough of the variance to persuade most scholars of its importance or enough variance to be a decision-influencing factor, even for those policy makers who can understand the statistics.

The second problem is that the broad general characteristics examined in such studies are either not realistically under the control of policy makers, or that decisions are highly unlikely to be made based on the logical relationships examined (especially with weak results) in such studies. Take the case of governance. I have observed a number of cases where states have changed in some ways their system for governing or coordinating higher education. These decisions happen infrequently, they are always controversial (requiring the expenditure of considerable political capital), and they typically are made because of frustration or annoyance with the way the people involved in the current system are working with policy makers. While to some extent they change the dynamics of the decision-making process, such governance changes rarely result in fundamental change in the results generated by the higher education system. Many other factors are more influential in determining outcomes.

More consequential research is focused on realistic operational options and on decisions actually under the control of decision makers. Studies to evaluate whether a *particular* policy decision has yielded the intended effects such as Susan Dynarski's study of the Georgia HOPE scholarship program[46] and Joseph Burke's work on performance funding[47] are good examples of useful and consequential research. Such studies may not cause policy makers to pursue or avoid a particular course of action, but they add directly relevant evidence to the policy debate.

Focus on salient, significant issues. A study is most likely to get the attention of policy makers and make a contribution to policy when it addresses a question of immediate policy concern. This observation may be too obvious to deserve mention, but scholars and policy analysts who focus on issues that are already on the minds of policy makers will capture their attention more readily than scholars whose first priority is to shape what policy makers think about.

The budget director of a large state (serving a Democratic governor with a Democratic legislature) was once heard to say, "We have three big problems in this state: K–12 education, health care, and pensions. Higher education isn't doing much to solve any of them." No doubt his opinion was unfair and inaccurate (at least to some extent), but if and where it exists, the perception of disengagement from important public issues is a problem for scholars. It can be a challenge for researchers and policy analysts to be relevant to policy debates without becoming a tool for one side or another in a philosophical disagreement, but the rewards of relevance will grow over time.

The Consortium on Chicago School Research, founded in 1990, is an outstanding example of a successful research effort focused on a salient public policy issue. The Consortium was created when improving the quality of public education in Chicago was an urgent priority. Reflecting the public mood and visible problems, not any sort of empirical review, then U.S. Secretary of Education William Bennett suggested in 1987 that the Chicago public schools were the "worst in the nation."[48] The state passed legislation that dramatically decentralized Chicago public schools in 1988, creating individual local school councils and reducing the power of the school board. Then in 1995, legislation was passed to give the mayor of Chicago power to appoint both the school board and the CEO of the Chicago schools. (These somewhat contradictory changes in governance still exist in uneasy tension, suggesting that governance changes of any kind are incomplete solutions to problems.)

The Consortium, then led by University of Chicago sociology professor Tony Bryk, drew on the resources of many of the universities in Chicago as well as the data resources of the Chicago public schools to employ research and data analysis to help "school reform" succeed. The complexity of the system and the depth of the interrelated problems—urban poverty, bureaucratic dysfunction, multigenerational educational disadvantage, difficult labor relations, and many more—defied silver bullet solutions.

Over many years Consortium-affiliated researchers sought to understand how the schools succeeded and failed to succeed by employing multiple research methods. (One dimension of the complexity of the problems was illustrated to me in an early meeting in which data concerning student mobility were presented—a shockingly high percentage of students attended different schools both during and between school years.) Consortium research projects ranged widely, including, for example: classroom observation to identify variance in the amount of time actually spent in instruction, interaction with students, and administrative activities; surveys of teachers, students, and parents in an effort to understand school culture and the community environment; and finding a way to estimate and compare the average contribution of individual schools to student progress, which required a painstaking analysis of the differences between versions of the Iowa Basic Skills test and controlling for student mobility among schools.

The breadth, depth, and persistence over time of this research program has yielded important general insights about elementary education, such as the relationship between higher levels of trust among the adult staff of a school and student learning, and particular insights about policies and practices in Chicago. A particularly dramatic contribution came from studies demonstrating that failure to graduate from high school frequently can be predicted by getting off track in eighth or ninth grade. Chicago's high school graduation rate has steadily increased to reach 69% in the 2013–2014 academic year,[49] apparently due in part to aggressive efforts to take corrective action in eighth or ninth grade when teachers and counselors observe a "getting off-track" incident such as failing a course or a suspension.[50] A thorough discussion of the Consortium's contributions can be found at its website and in *Organizing Schools for Improvement: Lessons from Chicago*, published in 2009.[51] Chapter 7 describes in more detail its initial work focused on elementary education.

Reasonable accessibility. Rigorous scholarship typically involves exhaustive reviews of the literature; scrupulous attention to detail; complete documentation of research methods; and often elaborate, fully documented statistical analysis of data. The skills and requirements of doing and presenting rigorous scholarship to an audience of scholars are often counterproductive in a policy environment. Outstanding PhDs who find themselves working in policy often have to learn new skills in order to design and present their work effectively.[52]

Scholars, and especially recent graduate students, frequently feel compelled to demonstrate a comprehensive grasp of a topic—a compulsion that is normally counterproductive, possibly fatal in a policy environment. *Not* presenting everything known about a topic is not "dumbing down." To the contrary, it takes considerable skill and intelligence to know what is essential to make a case and to understand how to capture and hold the attention of an audience managing information overload and many demands on its attention.

Similarly, elaborate statistics may be necessary for some analytical tasks, but a non-technical audience will be more persuaded by the face validity of a study than the statistics. If complex statistics are essential, they must be done well and stand up to potentially competitive (if not hostile) scrutiny. But statistical display to a policy audience is more likely to arouse suspicion than trust—if the message is hard to understand, then the messenger is less likely to be trusted and more likely to be suspected of deception or subterfuge.

When framing, conducting, and presenting research to a policy audience, simple, straightforward, and easy-to-grasp arguments add power and persuasiveness. Unnecessary words, potentially distracting information or side arguments, and difficult-to-understand mathematics subtract from the impact of the message. Achieving clarity and simplicity in research is challenging when the realities of the life and policy and practice themselves are inherently complex. But powerful policy research, just like a powerful scientific theory, is parsimonious.

Locally focused data and analysis. Typically, the goal of science is generalizable knowledge. How can inferences from a sample of data, from experiments in a few places, be generalized to other settings or to the population as a whole? Social science knowledge that can be generalized across settings is quite valuable, of course, but it is also rare and very difficult to acquire.

Researchers and policy analysts can dramatically increase their contributions by obtaining and analyzing *local data* that by definition are directly relevant to the interests and concerns of local policy makers. When the data studied includes the entire population, the confusing issues of statistical significance are side stepped, and sampling error is not an issue. Policy makers can quickly decide for themselves whether substantive significance is present. For example, the local focus of the Consortium on Chicago School Research discussed previously guaranteed that local decision makers would pay attention.

Let me illustrate the difference that a local focus might make with an example of findings from the National Center for Education Statistics Longitudinal Study, 1988 to 2000. Figure 6.3, first published by the National Advisory Committee on Student Financial Assistance demonstrates a very strong relationship between socioeconomic status (SES) and college participation, despite academic ability.

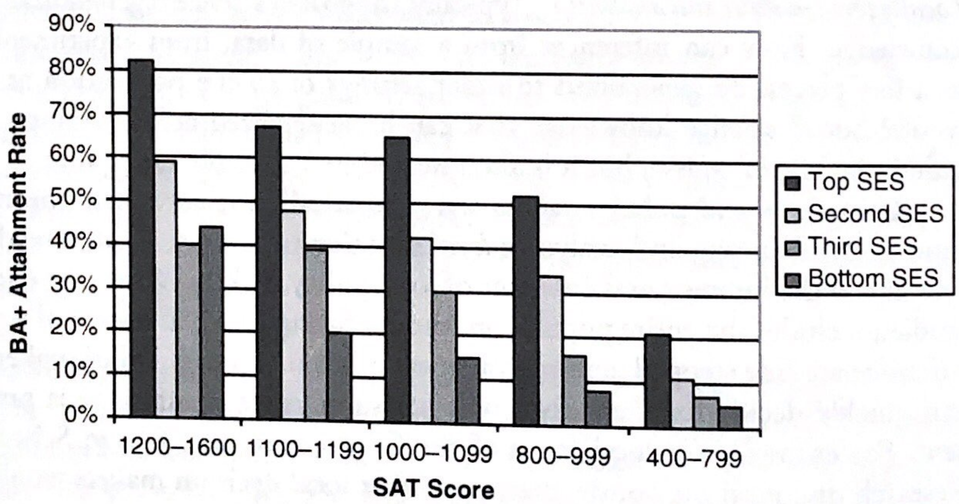
This chart has been frequently shown to policy audiences with the powerful message that high achievement, low SES students enroll in college at a rate 20 percentage points lower than high achievement, high SES students, and at essentially the same rate as low achievement, high SES students. Figure 6.4, taken from the same longitudinal study, compares baccalaureate degree completion

Figure 6.3 The Effects of SES on College Participation for High- and Low-Ability Students

College Participation by Achievement Test and SES Quartile		SES Quartile	
		Lowest	Highest
Achievement Quartile	Highest	78%	97%
	Lowest	36%	77%

Source: U.S. Department of Education Advisory Committee for Student Financial Assistance, February 2001.

Figure 6.4 BA Completion by Age 28 by SES Quartile and Estimated SAT Scores



Source: Anthony Carnevale, *Liberal Education*, Fall 2008, p. 58.

by age 28 for students in each SES quartile when controlling for estimated SAT score. This analysis shows that roughly two thirds of top quartile SES students with average ability (SAT scores in the 1000–1099 range) complete BA degrees by the age of 28. Only about 15% of similarly academically qualified students in the lowest quartile of SES had completed a BA degree by the age of 28.

These analyses have been shared with numerous audiences, and they quite possibly have had an impact on state and national student assistance policies in the past decade. But I don't believe they have galvanized many policy makers into action. I've often wondered what the policy response might be if similar data were available at the state and local level. How many high ability (or average ability) low SES students are not receiving a postsecondary education in our city or state? How does our success rate compare to other states and cities?

It is fairly common to see data that rank states and cities along various single dimensions of wealth, business activity, unemployment, or educational attainment. But such descriptive, one-dimensional rankings can be dismissed as simply describing "fate," the hand that has been dealt to a city or state, not an assessment of performance or possible neglect. If the analysis of these figures were commonly available at a local level, they could be used to stimulate corrective interventions and to monitor progress over time. Local evidence and local research is much more powerful in the policy arena than "general" research.

Focus on nontrivial effects. While thinking about writing this book, I asked a respected economist who has studied educational issues to share his thoughts on the many well-executed studies with disappointingly small or inconclusive findings. He sent me a copy of one of his studies, which found statistically significant, but relatively small positive effects of an after-school support program, accompanied by the plaintive question, "What size effect do people expect to get?"

Given the complex lives of poor urban youth, the relatively modest footprint of an after-school program in those lives, and the difficulty of designing and implementing a robust experimental study on such programs, I can understand why he believed that his study and the effect it found was an achievement. But elaborately conceived and executed experimental research with mixed, or even consistently positive results of small size do not move policy.

The experience of the past several decades suggests that very few if any single interventions or programs can consistently generate sizable effects. The What Works Clearinghouse in education has examined thousands of such studies, and the research community continues to generate dueling studies finding trivial differences on controversial interventions and policy strategies. These do not add up to positive change. It is difficult to argue for significantly increasing investments in similar research studies.

Sizable effects typically require multiple interventions in a multifaceted campaign, focused on clear goals and sustained (with intelligent adaptation) over time. Research and analysis can play important roles in such campaigns, as it has done in improvements in manufacturing, in health care, in education, and in social welfare through the still controversial War on Poverty and many other policy initiatives over time.

Concluding Thoughts on Research, Evidence, and the Politicization of Education

This chapter has ranged over a lot of territory. I have sought to make sense of political life as I have studied and experienced it, and then consider how research and evidence can help improve policy. It is evident that policy leaders and researchers of virtually every persuasion seek to employ data and evidence to support and advance their visions of what ought to be the social order. It is equally evident that research and evidence are unlikely ever to resolve once and for all the conflicts between dueling views of human nature and the world.

Political competition, however, is made more productive when research and evidence are employed in political debate. Evidence can be a mediator between conflicting worldviews, and in some cases it may even change worldviews. But evidence seems to be most powerful in the policy arena when it is modest in its aspirations—seeking to establish a shared view of the facts through commonly accepted measurements, focusing on local conditions and shared goals for improvement, and encouraging a frank and honest discussion of problems, limitations, and failures, as well as aspirations for improvement.