



Evidence Use and the Common Core State Standards Movement: From Problem Definition to Policy Adoption

Author(s): Lorraine M. McDonnell and M. Stephen Weatherford

Source: *American Journal of Education*, (-Not available-), p. 000

Published by: [The University of Chicago Press](#)

Stable URL: <http://www.jstor.org/stable/10.1086/673163>

Accessed: 02/09/2013 14:10

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *American Journal of Education*.

<http://www.jstor.org>

Evidence Use and the Common Core State Standards Movement: From Problem Definition to Policy Adoption

LORRAINE M. MCDONNELL

University of California, Santa Barbara

M. STEPHEN WEATHERFORD

University of California, Santa Barbara

Despite calls for research-based policies, other types of evidence also influence education policy, including personal experience, professional expertise, and normative values. This article focuses on the Common Core State Standards (CCSS) initiative, examining how research use varied over stages of the process and how it was integrated with other types of evidence. By drawing on elite interviews, we find that CCSS promoters and developers used evidence in much the way that policy analysis research would predict and that while research evidence was a major resource, it was combined with other types of evidence depending on political and policy goals at different stages of the CCSS process.

From the provisions of No Child Left Behind (NCLB) to the Investing in Innovation grants competition, the expectation in recent federal policies is that states, local districts, and schools will design their programs based on research. Increasingly, state and local initiatives echo the federal call for research-based policy. Yet decades of policy analysis suggest that even officials predisposed to acting on research knowledge do not rely solely on it to inform their decisions or to build a case in support of particular policy options.¹ That validated research is not the sole ground for policy choice comes as no surprise to students of politics. Majone (1989) emphasizes that evidence is intended not only to inform but also to persuade, and he casts the net widely, including “[any] information selected from the available stock and introduced at a specific point in the argument in order to persuade a particular audience of the truth or falsity of a statement” (10). Evidence, then, is a much broader category than research, ranging from the results of formal research studies to statistical data, judgments

Electronically published August 23, 2013

American Journal of Education 120 (November 2013)

© 2013 by The University of Chicago. All rights reserved.

0195-6744/2013/12001-0001\$10.00

NOVEMBER 2013 000

based on professional expertise, the personal experience of practitioners, and appeals to values articulated through stories and symbols.

Research is an important but variable component of this mix and, as such, raises two questions: How does the use of research vary over stages of the policy process? How is it integrated with other types of evidence in policy deliberations? Findings from policy analysis suggest that several factors influence decisions about whether and how to use research, including its availability and perceived usefulness, along with the incentives policy makers have for guiding and justifying decisions using alternative warrants such as constituent preferences and political ideology. However, because political and policy goals differ over the course of policy development, the factors shaping research use also vary depending on the stage of the policy process.

This article explores the question of how the use of research and other types of evidence differs as policy evolves from an idea to a set of prescriptions or incentives that are considered for formal enactment. The article's focus is the use of research in the development of the Common Core State Standards (CCSS) in K–12 mathematics and English-language arts (ELA) and their subsequent adoption in 45 states. In the first two sections, we describe the special role of policy entrepreneurs and then summarize assumptions drawn from research about evidence use during different stages of the policy process. The following section briefly describes our research. We then assess the extent to which evidence use in the Common Core is consistent with these patterns.²

The Role of Policy Entrepreneurs in Evidence Use

Policy research has consistently demonstrated the critical role that policy entrepreneurs play in bringing new ideas into different policy arenas and in advancing those ideas, sometimes for many years before they reach decision-makers' agendas (Baumgartner and Jones 1993; Kingdon 1995; Mintrom 2000; Sheingate 2003). Policy entrepreneurs are “advocates who are willing to invest their resources—time, energy, reputation, money—to promote a position in return for anticipated future gain in the form of material, purposive,

LORRAINE M. McDONNELL is a professor of political science at the University of California, Santa Barbara. Her research focuses on the politics of education policies, their design and implementation, and effects on school practice. M. STEPHEN WEATHERFORD is a professor of political science at the University of California, Santa Barbara. His research focuses on policy history and political leadership, including studies of how local activists and parents mobilized in response to school desegregation and citizen deliberative forums in school politics.

or solidary benefits” (Kingdon 1995, 179). They can occupy a variety of formal and informal policy-making roles, but the most effective ones typically have a claim to a public hearing, are known for their political connections or negotiating skills, and are persistent (Kingdon 1995).

One of the resources that policy entrepreneurs draw upon is research. Entrepreneurs are advocates, however, and they use research strategically in defining a policy problem, framing it, and then shaping and promoting a particular solution. Although these strategic and persuasive uses of research are at odds with the rational choice models that dominated policy analysis in its early years, they represent a more valid depiction of the actual use of research in policy-making venues (Majone 1989; Stone 2012). A simplistic version of this assumption is that policy entrepreneurs use research results selectively and even distort findings to advance their case. Entrepreneurs may certainly misuse research, but the competitive nature of the political process and the fact that credibility is the policy advocate’s most important resource minimizes outright distortion. At the same time, the inevitable complexity of real-world policy making allows ample room for uncertainty and interpretation. For instance, in some cases, such as school choice, research results are inconsistent and contested. Even in cases where the research base is solid, however, as with early literacy acquisition or the determinants of student retention, its application to problem definition and policy solution is open to interpretation and framing, depending on the context and feasibility factors such as political, organizational, and resource conditions.

How policy entrepreneurs use research knowledge also depends on who the entrepreneurs are. In some instances, they may be knowledge producers such as university researchers or think-tank staff who choose to disseminate their work directly as it relates to particular policy issues and venues. Alternatively, they may be intermediary organizations that act as research brokers and translators in the course of pursuing their organizational mission (e.g., foundations, nonprofits, interest groups). Policy entrepreneurs may be current or former elected officials, whose political status lends their advocacy an additional claim to the attention of their colleagues. Regardless of whether they function from an official position inside government or an unofficial one outside, policy entrepreneurs have to be sensitive to the differing types of evidence needed throughout the policy development cycle.

Stages of the Policy Process and Evidence Use: Assumptions from Policy Analysis

Policy making is rarely linear—moving from problem definition to design, agenda setting, enactment, and finally implementation. Rather, policy options

Evidence Use and the Common Core State Standards Movement

may exist before a problem is defined, and during implementation, policies may cycle back for new enactments or fall off the agenda altogether. Nevertheless, despite nonlinearity and limited predictability about direction and duration, most policies move through well-defined stages, and past research suggests that evidence use varies across phases of the policy cycle.³ The current analysis focuses on three distinct stages: (1) problem definition and promoting a solution, (2) policy design, and 3) policy enactment.⁴

Problem Definition and Promotion of a Solution

In his seminal work on agenda setting, John Kingdon (1995) details the interplay between disparate facts and their interpretation in the process of defining a given situation as a policy problem. Indicator data or a body of research may, for instance, identify a problem—for example, highway deaths are rising, childhood immunizations rates are declining, program costs are increasing—or research might indicate that a current policy is not producing its intended effects. Typically, most members of a policy network will accept such research-based information as valid. At the same time, “there is a perceptual interpretative element” in defining policy problems (Kingdon 1995, 110). That element is manifested in what data are highlighted, how they are interpreted, and in what factors are identified as causing the problem.

As Stone (2012) notes, problem definition is the strategic representation of situations. It is strategic and interpretative for several reasons. First, whether a given situation constitutes a policy problem depends on how participants perceive the discrepancy between it and some ideal state or social goal. What that goal is and the distance between its attainment and the status quo depend on their value preferences. Second, part of problem definition is identifying causes, and “to identify a cause in the polis is to place burden on one set of people instead of another” (Stone 2012, 207). How the cause is defined has direct implications for who will bear the costs of a policy and who will benefit from it. Therefore, those likely to be affected by a policy have a strong incentive to influence the selection and framing of relevant evidence. Finally, how a problem is defined shapes the policy solution proposed. So policy entrepreneurs promoting a particular policy option will select evidence that allows them to define a problem in such a way that their policy proposal becomes the preferred solution.

The nature of problem definition, then, suggests that although research results and indicator data play a role, other types of evidence may be equally important. These include appeals to values such as equality, liberty, and economic security. Metaphors may also be used to evoke strong political and cultural symbols (e.g., bureaucratic red tape, invasion of privacy; Stone 2012).

Policy entrepreneurs also often try to put a human face on a problem through the use of anecdotes and other narrative devices.

The likelihood that a range of evidence will be used at this stage of the process increases for major policies that embody significant ideas. Ideas in the policy arena are specific policy alternatives (test-based accountability, school finance reform) as well as the organized principles, values, and causal beliefs in which policy proposals are embedded (democratic representation, distributional equity; Beland 2005, 2). Because ideas typically embody a set of desired ends and the strategy or theory about how to produce those results, they capture normative and instrumental dimensions of policy. In promoting their ideas, policy entrepreneurs frame arguments that combine normative theories and empirical evidence from a variety of sources—for instance, about what education ought to achieve and how it ought to be delivered.

Significant policy ideas grounded in a combination of research and other evidence are critical in altering well-established policy regimes. Sometimes called “policy monopolies,” these institutional arrangements have several advantages that allow them to resist changes even when they have outlived their usefulness. They (*a*) are supported by powerful policy ideas that are connected to core political values, (*b*) combine empirical information and emotive appeals, (*c*) can be easily understood, and (*d*) are communicated directly and simply through image and rhetoric. Policy entrepreneurs can disrupt a policy monopoly and effect major change, but to do so they need to redefine the dominant policy image using ideas that challenge it and capture the imagination of the media, policy makers, and the public. Entrepreneurs provide new understandings of policy problems and new ways of conceptualizing solutions (Baumgartner and Jones 1993).

In sum, the policy analysis literature suggests that during the problem definition and solution identification stage, research-based evidence is likely to be combined with evidence appealing to elected officials’ and their constituents’ core values and that evokes positive emotional responses. Normative evidence is especially important when the option being advocated embodies a significant idea and requires a major policy change. However, even in these cases, some evidence is necessary that demonstrates there is a reasonable likelihood the proposed policy option will result in the expected outcome.⁵ That evidence may be selectively framed and interpreted, but it needs to be derived from systematically collected and analyzed data generally viewed as reliable and valid.

Policy Design

This stage of the policy process consists of technical tasks undertaken in a politicized context. A policy option has now moved to the decision agenda

Evidence Use and the Common Core State Standards Movement

and is being considered for formal enactment. At this point, legislative language is drafted, specifying such details as the funding mechanism and administrative arrangements. Legislative staff are primarily responsible for the work, and they draw on legal and fiscal expertise to ensure that the resulting language is consistent with existing policy and legislative authority. Related policies often serve as templates for how to structure the new policy. Staff may also draw on evaluations of past policies and on research documenting the relationship between different types of intervention strategies and educational outcomes. Major sources of evidence at this stage are research, past policies, and the professional judgment of legislative and executive branch staff (Quirk and Nesmith 2011).

The political context in which the details of a policy are developed has several implications for what additional evidence is used. First, those advancing the policy need to maintain support and blunt opposition. They also need to consider the interests and concerns of the agency staff and street-level bureaucrats who will be implementing the policy, and thus other evidence will come from these individuals and the groups that represent them. The evidence might be research findings and other data, if available, as well as professional judgment and personal experience. It can be presented formally in legislative hearings and also discussed in informal meetings with elected officials and their staffs. Often the most persuasive evidence will be local knowledge from an official's own constituents—for example, conversations with hometown educators about their classroom experience. A second political dimension involves timing. The opportunities for action on a policy proposal—the open policy window—may not be available for long (Kingdon 1995). Policy entrepreneurs, forced to move quickly before the window closes, may bring forward policy proposals that are not fully developed and that miss or ignore evidence suggesting problems or unintended effects.

The evidence used at this stage, then, is less likely to include the type of normative and emotional arguments that are often central in the first stage. Evidence grounded in legal analysis, evaluation studies, and basic or applied research on the relationship between policy interventions and learning outcomes is likely to play a central role in drafting specific policy provisions. However, the political context means that less systematic sources of evidence are also considered and that some research may be ignored.

Policy Enactment

The imperative at this stage is to build a minimal winning coalition in favor of the policy. The evidence brought to bear is basically a reprise of what was used in the first stage to define the policy problem and promote a particular solution. However, because the policy has subsequently been developed from

a general idea into a detailed legislative proposal, the evidence is likely to be more specific and to be targeted to address the questions and concerns of individual legislators, especially those who are still undecided about the policy (Mucciaroni and Quirk 2006, 24). At this point, legislators will continue to consult with relevant interest groups and constituents. But they will also look to colleagues who are experts in the policy domain (e.g., committee chairs) and to leaders of their partisan caucuses for cues about how to vote and for evidence to justify the vote (Kelman 1987).

Rationale for Focusing on the Common Core and Study Data

The CCSS is an especially appropriate research site for observing the use of evidence for three reasons. First, it has the potential to become one of the most significant policy shifts in American education in more than a century, for it fundamentally alters a long-standing policy regime of academic content standards unique to each state. Second, the Common Core's advocates explicitly promoted it as "research and evidence-based" and established procedures to encourage the use of research in drafting and validating the standards. Finally, the Common Core is an ongoing policy that moved from the idea stage to adoption within 5 years. As such, it provides the opportunity to examine research use in real time rather than *ex post*, as has been typical of earlier studies. Because our primary interest is in mapping the use of evidence over the development of this policy, we employ the predominant approach to within-case qualitative causal analysis in political science—process tracing (Bennett and Elman 2006; George and Bennett 2005; Mahoney 2000). The goal of process tracing is to gather information about specific events and outcomes, especially through a close examination of the intervening processes that link the multiple features and individual actors involved in the case. For instance, process tracing has been used to explore causal relationships underlying key foreign and domestic policy decisions (Collier et al. 2005; George 1979).

The analysis draws on extensive interviews with leaders of the Common Core movement and their supporters, members of the work groups and committees charged with writing and validating the CCSS, national and state education policy makers and researchers, as well as groups critical of the CCSS.⁶ Over the past 2 years, 111 interviews have been conducted at the national level and in four states. These structured interviews focus on the politics and process of Common Core promotion, development, and adoption; why participants chose to use certain types of evidence; and what other types were either unavailable or not used.⁷ Potential interviewees were identified through multiple sources, including a large database of documents related to the CCSS and its

Evidence Use and the Common Core State Standards Movement

development;⁸ consultation with staff at the James B. Hunt Jr. Institute for Educational Leadership and Policy, who were active participants in the CCSS process;⁹ and additional names of potential informants offered by interviewees. Our primary method for analyzing the interview data was to identify the stages of the policy cycle in which respondents participated and then to compare their reports of the types of evidence used with hypotheses derived from the policy analysis literature. We found few inconsistencies among respondents about evidence use by stage of the policy cycle. Differences in the interpretation of evidence or the weight given evidence types and sources were largely explained by role position (e.g., CCSS critics interpreted the relationship between standards and achievement differently than proponents; academic curriculum experts gave greater weight to research results and less to judgments about feasibility based on administrative experience than state agency staff).¹⁰

The Stages of the Common Core Initiative and Evidence Use

Defining the Problem and Promoting Common Standards as the Solution

Although more than 20 national organizations and their members would eventually become involved in promoting adoption and implementation of the Common Core, a small group of policy entrepreneurs initially promoted the idea of similar academic content standards across multiple states. This group included several former governors, education advocacy groups, and organizations representing state and local officials. The origins of the Common Core movement can be traced back to several unsuccessful attempts 2 decades earlier, including efforts spearheaded by organizations representing subject-matter specialists and the National Council on Education Standards and Tests (NCEST), chaired by Roy Romer, then Governor of Colorado.¹¹ These attempts, introduced in both the G. H. W. Bush and Clinton administrations, foundered, enmeshed in controversies over the curricular values underlying state standards and assessments. They were opposed in Congress by Republicans who feared possible federal encroachment and Democrats concerned about the impact of assessments on students who had not had adequate opportunities to learn the required content.

Although these controversies made top-down national standards a non-starter, the idea of voluntary, shared standards gained momentum with an initiative begun by a group of states. In 2001, Achieve, in collaboration with the Education Trust and the Thomas B. Fordham Foundation, established the American Diploma Project (ADP) to try to ensure that high school diplomas signified that students are prepared academically for entry into higher education and the workforce. These standards were based on surveys of higher

education faculty and businesses indicating what knowledge and skills students needed to pass college-level courses and workplace training programs. Sixteen states agreed to align their high school graduation requirements with the benchmarks derived from the surveys. Subsequently, 15 states agreed to develop a common assessment in Algebra II, and later five states agreed to develop and administer an Algebra I exam. As one organizational representative involved in the effort noted, “the ADP was the existence proof that you could get common standards” (personal interview).

In 2006, two former governors, James B. Hunt Jr. of North Carolina and Bob Wise of West Virginia, decided that what had seemed impossible 10 years earlier was now a realistic aspiration. In their view, policy makers, educators, and the public had become accustomed to the idea of content and performance standards. Opinion polls indicated support for national standards, and state policy makers were beginning to see potential cost advantages to common standards, particularly given the requirements of NCLB. Hunt and Wise became the policy entrepreneurs most responsible for persuading key decision makers and constituent organizations to support the idea of national (subsequently renamed “common”) standards (Hunt 2009; Rothman 2011; personal interviews). They identified their audience as the national policy community—governors, chief state school officers, other state officials, members of Congress, civil rights groups, foundations that might support the effort, and the media (personal interviews). Their definition of the problem was strategically crafted to point to common K–12 standards across multiple states as the solution. Together with the organizations they lead, Hunt and Wise drew together the findings of several lines of research to shape a clear image of the policy challenge:¹²

- The achievement of US students is low compared with the nation’s global economic competitors.
- The United States has an unacceptable achievement gap among students depending on their race and ethnicity, social class, and place of residence.
- US students leave high school inadequately prepared to succeed in college and employment.
- Countries with high-achieving students have focused, rigorous, and coherent national standards.
- US standards vary considerably across states and are “a mile wide and an inch deep.”
- Academic content standards common to multiple states are a critical ingredient in solving these problems.

Together with the proviso that the Common Core is an initiative led by states for states, these arguments have continued to form the essence of the Common Core rationale.

Although the competitiveness of the United States in the global economy

was central to the narrative, research showing dramatic variation among states came to play an increasingly prominent role in the problem definition phase.¹³ A National Research Council (NRC) report sharpened the issues and brought two important lines of research to the attention of the policy community.¹⁴ One was research by Andrew Porter and his colleagues—a fine-grained analysis of 31 state standards in three subjects, comparing state standards to each other and measuring their alignment with the National Council of Teachers of Mathematics mathematics standards and the science standards developed by the NRC. Focusing on topic coverage and the level of cognitive demand, the researchers found little evidence to support the assumption that a de facto national curriculum existed as a result of states' use of national documents such as the NCTM standards or the widespread adoption of similar textbooks. In fact, overlap in topic coverage across grade levels within the same state was greater than the alignment across states at the same grade level (National Research Council 2008). The implications that policy entrepreneurs drew from Porter's work were considerable state-to-state variability and substantial redundancy in current state standards.

This evidence of variability in content standards gained additional force as further research analyzed state-by-state variation in assessments and performance standards (National Center for Education Statistics 2007). Mapping state proficiency standards in mathematics and reading for grades 4 and 8 onto the appropriate National Assessment of Educational Progress (NAEP) scale, researchers found that state differences in the percentage of students scoring at the proficient level on state assessments did not represent real differences in achievement as measured by NAEP but reflected where a state set its proficiency levels. Most state cut points, moreover, fell below the equivalent of the NAEP proficient standard, and some even fell below the NAEP basic standard. A telling example, presented at the NRC workshop, was that a North Carolina student performing at the same level on the NAEP reading assessment as one in South Carolina would be deemed proficient in North Carolina but performing at a basic level in South Carolina, with the possibility of having to attend a remedial class because South Carolina's proficiency standard (cut score) was higher (National Research Council 2008, 23). For advocacy groups supporting national standards, this discrepancy between student performance on state assessments and NAEP was powerful information in efforts to persuade state officials because it contributed to a picture of states with significantly different, and typically low, expectations of students (personal interviews).

The case for common standards was crystallized in a report published by the National Governors Association (NGA), the Council of Chief State School Officers (CCSSO), and Achieve (NGA et al. 2008). Authored by an International Benchmarking Advisory Group, chaired by then-Governor Janet Na-

politano (AZ), then-Governor Sonny Perdue (GA), and Craig R. Barrett, the chairman of the Intel Corporation board, the report drew heavily on research using data from the Programme for International Student Assessment (PISA) and Trends in International Mathematics and Science Study (TIMSS).¹⁵ The report focused on US students' low achievement, compared with international competitors, and documented the achievement gap separating US students from different socioeconomic backgrounds, noting that the distribution of US students' scores puts the country among the most unequal in the world. It warned that "the United States is falling behind other countries in the resource that matters most in the new global economy: human capital" (NGA et al. 2008, 5). The report recommended that states upgrade their standards "by adopting a common core of internationally benchmarked standards in math and language arts for grades K–12 to ensure that students are equipped with the necessary knowledge and skills to be globally competitive" (6). The advisory group grounded its rationale for international benchmarking in William Schmidt's research on TIMSS, showing that standards in high-performing countries are characterized by focus (fewer topics covered at greater depth), rigor, and coherence (an orderly progression of topics following the logic of the discipline and minimizing repetition; 24).

Evidence use during this phase of the CCSS process was quite consistent with what would be predicted by the policy research literature. The challenge for the entrepreneurs promoting national standards was to dismantle one of the most deeply entrenched and strongest policy regimes in US education: the tradition of each state and its local districts deciding separately what students should be taught. In doing so, they had to define a set of problems to which their alternative policy idea was the solution. Research-based evidence was used to demonstrate that state-specific standards policies had resulted in considerable variability in curricular goals across states. Evidence of variability in content and performance standards across states, and differing levels of student achievement on standardized tests, was strong and generally accepted. There was, however, less agreement about the relationship between the two and the cause.

The resultant need for interpretation allowed proponents to represent the situation strategically, attributing low achievement to states' variable and low-quality standards. Other researchers—especially those who study policy implementation—drew a different conclusion, emphasizing the lack of system capacity, especially supports for teachers and students (Cohen and Moffitt 2009). They argued that the shortcomings documented in research findings are less the result of standards themselves than how policies have been implemented. Where the standards-based accountability ideal envisioned that curriculum and instructional materials, teacher training, and assessment would be closely coordinated, in practice the assessment portion had come to dom-

inate, and many schools lacked the capacity to bring all students to proficiency (Goertz 2007; Ravitch 2010).

These alternative inferences lead to different policy solutions, but all researchers could say with any certainty is that, at best, common standards might be a necessary, but not sufficient, condition for improved educational outcomes.¹⁶ The enabling conditions necessary for standards policies to work as envisioned are the various supports in which policy makers had underinvested in the past. Common Core advocates understood what researchers were telling them about enabling conditions. However, during this stage of the policy process, they chose to downplay them because they would complicate the agenda at a time when a policy window was opening but might not be open for long.¹⁷ The Common Core's policy entrepreneurs also used research evidence strategically in how they framed its rationale, emphasizing global competitiveness because it carried great appeal among governors concerned about the economic health of their states—even though educators were less persuaded that this was a compelling reason for major curricular change. During this stage of the process, then, one particular set of inferences, among the differing ones that could be drawn from research and indicator data, were selected and framed in such a way as to persuade key policy audiences that common standards held the potential to rectify pressing educational and economic problems.

In one respect, the initial phase of the CCSS process differed from what happens typically when policy entrepreneurs define problems and promote solutions. The CCSS process did not rely heavily on the use of symbols and emotional appeals. Strongly held values, such as ensuring equality of educational opportunity and keeping the United States globally competitive, were prominent in the discourse on national standards, and at one point, Governor Hunt likened the challenge to World War II (National Research Council 2008, 73). However, perhaps because most of the appeals at this stage were to political elites and organizational leaders and not to the general public, data rather than symbols and stories were more prominent.

Developing and Validating the CCSS

Because of past opposition to standards issued by the national subject-matter organizations and the federal government's role in sponsoring those efforts, it was agreed that states would have to take the lead this time. Consequently, CCSSO and NGA assumed leadership during this stage of the process. It began when 48 states signed a memorandum agreeing to participate in a process of developing a common core of state standards in ELA and mathematics and to support the development of common assessments to measure progress toward the standards. The process proceeded in two phases: a set of

college-and-career-ready (CCR) standards were first developed to define what students should know at the end of high school in order to be prepared to enter college or a workforce training program, followed by the design of K–12 standards that essentially map back from the CCR standards to grade-by-grade ones that allow students to build toward mastery of the CCR standards by the end of high school.

The second stage of the CCSS process can best be described as technical tasks undertaken in a political context, or, as one leader of the development process described it, “I would argue that the standards development was primarily a political process informed by evidence” (personal interview). Four factors shaped evidence use during this stage. The first was CCSSO and NGA’s assertion that one guiding principle of the development process “is being driven by evidence and research. In the past, standards were largely based on personal judgment. By allowing personal judgment to determine what concepts are in or out of standards, the process often becomes a negotiation, rather than a reflection on what the evidence and research tells [*sic*] us about the connection between K–12 experiences and success in higher education and promising careers” (Wilhoit 2009).

Leaders of the CCSS initiative acknowledged that their commitment to ground the effort in research and evidence was a strategy to avoid past ideological debates stemming from the “curriculum wars” of the 1990s (personal interviews). So, in essence, reliance on research operated as a political strategy to depoliticize the standards development process. Consistent with the commitment to research-based evidence, a variety of sources were used. These included research syntheses published by organizations such as the NRC, expert panels convened by federal government agencies, and national subject-matter associations; scholarly journal articles, chapters, and conference presentations; reports by Achieve, ACT, and the College Board based on faculty surveys and analyses of the relationship between student performance on admission tests and their grades in lower division course work; and reviews of international test data and the standards of high-performing countries. The standards writers also reviewed existing state standards and the NAEP frameworks. Although they were not formal research studies, these reviews constituted a form of applied policy analysis in that the policy document under development was systematically compared with existing policies.

The limited supply of relevant research was a second factor shaping evidence use. When asked why the standards documents note that they are based on “research and evidence,” a leader of the development effort replied, “we wanted to be able to cite non-peer-reviewed research because there’s not enough research available, and often the findings are inconclusive” (personal interview). The type of evidence used in the absence of appropriate research was expert judgment.

A notable example emerged in the development of the mathematics stan-

dards. Research on learning trajectories in mathematics is quite robust at the K–2 level but not at higher grade levels.¹⁸ Learning trajectories in the early grades are better developed for several reasons, including the ability to draw on a rich research base by developmental psychologists about children’s early learning and the relative simplicity of concepts and skills at the early childhood level as compared with the complexity of topics and how they might be related and sequenced in more advanced mathematics (Clements 2011, 20). Consequently, the standards writers asked several researchers who study math education for their best judgments about what trajectories might look like in higher grades based on how students learn. They then used those inferences in deciding where to place topic and skill standards. Similarly, mathematicians were also asked what progressions of standards made sense in terms of the logic of mathematics as a discipline and the foundations for higher-level mathematics study.¹⁹

Those directly responsible for drafting the standards viewed their task as drawing on available research evidence to the extent that they could and ensuring that major concepts were grounded in research: “It’s not as if every word of every standard is based on [research] evidence, like the difference between the [ELA] central idea standard in grade seven versus six—but the guardrails, the unities, the thread, the core principles that guide them and develop through them are based on the evidence” (personal interview). In their attempts to move the process along, leaders of the standards development assumed that “if we waited for the perfect research to inform the development of the standards, we would never have the standards today. So the goal was to use the best available research with the understanding that as we move deeper and deeper into implementation, researchers, academia will produce further research, and that further research will inform future iterations of the standards” (personal interview).

A third factor shaping the CCSS leaders’ evidence use was the desire to ensure that the expert judgments relied upon were more than just those of researchers. Both NGA and CCSSO had to maintain support among their state-level constituents and also among key groups such as teachers unions, whose members would be critical to whether the CCSS would ultimately be implemented in classrooms. Although only one draft of the CCR standards and one of the K–12 standards were available for public comment, state education agencies were asked to comment on multiple drafts. Six states (California, Colorado, Florida, Georgia, Massachusetts, and Minnesota) that had highly regarded standards were specifically asked to recommend standards writers and to review drafts.²⁰ State department of education (SDE) personnel in states with standards that had been judged rigorous by external groups, such as the AFT and the Fordham Foundation, reviewed the draft CCSS to ensure that they met or exceeded their current standards in terms of parsimony,

coherence, and rigor (the “fewer, clearer, and higher” promised by CCSS developers). However, SDE staff in all participating states were also given an opportunity to review several drafts. In exercising their professional judgment, they were most concerned about clarity and logic of presentation based on their experience organizing and formatting standards to facilitate their use in classrooms.

Both the AFT and the NEA convened groups of teachers to review CCSS drafts. The AFT drew its group of reviewers from members involved in providing professional development to colleagues and the NEA from national-board certified members. The AFT math review team met four times and the ELA team three times. After an extensive review of drafts, they communicated their concerns in face-to-face meetings with the standards writers.²¹ Like the SDE personnel, the teachers focused on how their colleagues were likely to respond to the draft standards. Consequently, they were particularly attentive to a logical presentation of the standards and to the language not being overly academic. The NEA group used a similar process, submitting multiple reviews and meeting with the standards writers. Its concerns centered around the inclusion of ELA standards focused on students “judging the worthiness and relevance of information” and standards that distinguish “enabling skills from goals” (personal interviews).

The SDE staff, teacher organization representatives, and those from other organizations who reviewed early drafts and whom we interviewed, reported that they felt their input was taken seriously and that they could see some of it reflected in the final versions of the CCSS. However, the tight time constraints of a process that lasted less than a year meant that changes made after each successive draft were not annotated or explained: reviewers of successive drafts could see that they had been changed but could not tell precisely in what ways. Nevertheless, with a few notable exceptions, individual participants and groups whose members would be affected by the standards felt that CCSSO and NGA staff were acting in good faith and trusted the standards writers because several had assisted individual states in the past in developing their own standards.²²

The fourth factor shaping evidence use during this stage was the role of the validation committee. The NGA and CCSSO convened a 29-member committee, including 17 university faculty and 6 others working in research positions. The committee also included three teachers, two principals, and one urban superintendent. Their charge was to review the process by which evidence was used to create the CCR and K–12 standards and to determine whether the standards writers had adhered to a set of principles including “a grounding in available research and evidence.”²³ The committee could provide feedback to the standards writers, but it could not rewrite the standards. After meeting twice in person and through e-mail exchanges with the NGA and

CCSSO staff, all but four members of the committee signed a statement certifying that the CCSS are consistent with the criteria established in the committee's change. Those who did not sign off argued that the CCSS are not sufficiently rigorous and that current standards in states such as California and Massachusetts were superior.

For other members of the validation committee, professional judgment was a major source of evidence. The compressed time frame and a realization about the limits of the research base meant that they had to fall back on judgments based on inferences drawn from their general store of expert knowledge rather than from a review of specific studies or even bodies of research. One member of the validation committee described the process in this way: "It was pretty clear from the start that nobody thought there was sufficient evidence for any of the standards. . . . The review process, in short, was inclusive and involved feedback from a lot of different perspectives. This is not 'sufficient research evidence,' but it is thoughtful professional judgment, applied systematically" (personal interview). Several members of the validation committee noted that familiarity with those drafting the standards was also a factor in the decision to validate them. Like the SDE staff, validation committee members reported knowing the standards writers and having worked with them in the past. Their work was known to be rigorous, and they were trusted. Another aspect of the validation committee's professional judgment was based on their knowledge of current state standards and international standards and their belief that the CCSS are better.

The second stage of the CCSS process differed somewhat from the typical design phase of policy making. The standards writers were not preparing legislation in the traditional sense, and states had only agreed to participate in a drafting process and not to adopt the new standards or even to put consideration of them on their decision agendas. Nevertheless, to a considerable extent, standards development was akin to the policy design stage, and evidence use was similar. Political support had to be maintained, and the process had to move quickly to take advantage of an open policy window and opportunity for action. Grounding the standards as much as possible in research was an inspired political strategy that avoided not only past ideological controversies but also the negotiations and "horse trading" that had led to bloated state standards in the past. At the same time, drawing on research ensured that the relationship between standards as a policy intervention and the desired goal of improved student learning was systematically considered during the development process, even if that relationship could not be completely and validly specified. As with policy design generally, development of the CCSS relied heavily on comparisons with past policies, evaluations of seemingly effective policies in other places, and on the professional judgment of those who would be responsible for implementing the policy.

State Adoption of the Common Core State Standards

In most states, adoption of the CCSS to replace the state's existing standards required a vote by the state board of education (SBE) because of its authority over curriculum. Three factors explain the kinds of evidence that were used in anticipation of the SBE vote. First, the adoption stage was truncated. Even the Common Core's strongest supporters assumed that it would take 3 years or more for a majority of states to adopt the standards. They based that estimate on the extensive consultation and deliberation usually required when states adopt new content standards. However, the deadlines for the federal Race to the Top competition, which awarded up to 70 points (14% of the total) on applications from states that adopted common standards and assessments meant the adoption process in most states was shortened to only a few months. Consequently, the process often resembled a political campaign targeted at individuals and groups who were likely to try to influence the SBE vote. The CCSSO and NGA provided their constituents with a "messaging tool kit" that included answers to frequently asked questions, template letters to the editor, and a sample op-ed article that could be adapted depending on whether the author was a business leader, teacher, civil rights leader, or a parent. The substance of these communication strategies highlighted central parts of the CCSS narrative: the focus on students' CCR preparation, US global competitiveness, the potential for commonality across states and local communities, the voluntary nature of state participation, and the inclusive state-led development process (Common Core State Standards Initiative 2010).

The national educator, parent, education advocacy, and civil rights groups supporting the CCSS worked with their state affiliates and allies in providing information and other assistance.²⁴ The policy entrepreneurs who had promoted the Common Core and the writers who had developed the standards made numerous appearances in states to brief policy makers on the content and rationale for the CCSS. As a result, evidence during this phase was a version of the Common Core narrative that had been used in the first two stages now customized to various state audiences.

Federalism was a second factor shaping evidence use, requiring that evidence use be tailored to particular states. Policy makers and the attentive public in states with especially rigorous standards had to be convinced that the CCSS are at least as rigorous as their current ones. So systematic, side-by-side analyses were prepared for a number of states comparing the CCSS with state standards. In some instances, these were done by in-state agencies, and in others, organizations such as Achieve and WestEd prepared the detailed comparisons showing topic and skill coverage arrayed by grade.

A third factor was that except for a small group of CCSS opponents, there was little scrutiny of the research base during this stage. State policy makers

and their staffs assumed that CCSSO and NGA had used the validation committee and other mechanisms to ensure that the CCSS had been adequately vetted and were grounded in relevant research. In addition, SDE staff, state teacher organizations, and other education-related groups knew that their colleagues had provided feedback and that it was substantially reflected in the final standards. So trust in how the standards were developed, coupled with a fast-moving adoption process, meant that the research base was rarely a topic of discussion at the state level.

The adoption stage of the CCSS was quite consistent with what typically occurs during policy enactment. Instead of moving a policy proposal from a legislative committee to a floor vote, the CCSS were considered for adoption in more than 40 states, each with slightly different information needs. Few state participants were concerned about the research supporting the CCSS. Rather, policy makers and groups whose members were likely to be affected by the CCSS wanted detailed comparisons between the Common Core and the state's status quo standards, and they sought persuasive arguments that the benefits would outweigh the substantial costs of such a major change.

Conclusion

Our examination of the Common Core movement suggests that its promoters and developers have used evidence in much the way policy analysis research would predict and that use has varied over stages of the CCSS process. The factors that best explain when and how research-based evidence was used are political context and the availability of relevant research. The political context shaped how research was used strategically at different stages. In the first stage, policy entrepreneurs used research primarily in defining a set of problems for which they already had a solution. In doing so, their political task was to shape the inferences and policy implications that were drawn from multiple studies and data sources. In the second stage, the political context provided a powerful incentive for the use of research-based evidence beyond the functional one of legitimating the CCSS as grounded in research knowledge about the determinants of students' post-high school success and the cognitive and developmental pathways that characterize effective learning. The promise to base the Common Core on research and evidence allowed its promoters to avoid past ideological battles and to produce streamlined standards. Drawing on participants' professional judgment and experience was a way to build and maintain support for the Common Core, but it also allowed the enterprise to move forward in the absence of a sufficiently robust research base. By the third stage, the rationale for the CCSS was well developed and generally accepted, so the political context only required that the narrative be tailored

to specific audiences. The new evidence introduced at this stage provided systematic comparisons of the CCSS with existing state standards.

The Common Core represents a “best-case” example of research use in education policy making, because the nature of the policy is such that there was a strong likelihood research would inform its development. Although the relevant research base is inadequate in several key areas, its major findings are largely uncontested, unlike in other areas of education policy such as school choice. The promoters and developers of the CCSS recognized that their commitment to standards grounded in research and evidence would produce clear educational and political benefits. Nevertheless, even under these advantageous conditions, they found it necessary and desirable to integrate research with other types of evidence depending on the political goals to be accomplished during each stage of the policy cycle.

The major implication from this “best case” is that researchers need to develop a more nuanced theory of research use in education policy making. Even under the most favorable conditions, research will be combined with other types of evidence. Part of the reason will be political, but integrating research and other types of evidence can also serve policy and educational purposes. The research base may be inadequate or incomplete for the policy task, it may be inconsistent or contested, and wise policy may require that basic research be combined with knowledge based on professional judgment about how best to teach students in different educational and cultural contexts. Earlier studies demonstrated that assuming a straight line between research results and policy design was naive (Weiss 1977, 1982). Now the challenge is to build on that fundamental insight by specifying the conditions under which research and other types of evidence are integrated during policy deliberations. Understanding how evidence use varies over stages of the policy cycle is a critical step in developing a theory of how research and other types of evidence together inform policy.

Notes

1. We define research-based evidence as evidence grounded in one or more systematic studies characterized by identifiable questions, an explicit design for collecting and analyzing data that is attentive to issues of reliability and validity, and some type of peer review.

2. The study on which this article was based is supported by a grant from the William T. Grant Foundation as part of its Uses of Research Evidence Program. An essential goal of the initiative is to connect studies of research use to broader disciplinary frameworks. For that reason, we have grounded our examination of evidence use in the Common Core initiative in the policy analysis and political science literature.

3. The discussion in this section draws on policy research literature, but similar patterns of variation in the use of evidence across stages of the policy cycle are apparent in political science research on national institutions and processes. For example, the

Evidence Use and the Common Core State Standards Movement

legislative process in Congress is more open to new information (e.g., from interest groups, social movements, the White House, or research entrepreneurs) at the early stages, when the need for action is being established or the agenda of concrete alternative proposals is being shaped, than at later stages, during floor debate and voting (Kollman 1998; Light 1999; Mucciaroni and Quirk 2006).

4. We chose not to include the implementation phase in this analysis because we are still collecting data on the Common Core's implementation. Also, our initial examination suggests that evidence use may vary significantly during implementation as compared with earlier stages because it is a process that moves beyond policy making into school and classroom practice with different opportunities and incentives for research use.

5. Policies are often portrayed as "if, then . . ." statements that assume if a policy mandates or offers incentives for certain actions to be taken, the change will occur. These theories of action are essentially predictive causal statements, and usually decision makers require some systematic, research-based evidence supporting the assumed link between policy and effects.

6. In conducting the interviews, we benefited from the able assistance of Jeanette Yih Harvie of UC Santa Barbara and Stephanie E. Dean at the Hunt Institute.

7. At the national level, we conducted 49 interviews, including interviews with representatives of the organizations developing the draft standards (the National Governors Association, the Council of Chief State School Officers), think tanks and policy research organizations (e.g., the Brookings Institution, Center for Education Policy), congressional and executive branch staff, as well as interest groups involved in education policy (e.g., National Education Association [NEA], American Federation of Teachers [AFT], Council of Great City Schools, National Council of La Raza, Education Trust). To track the process of policy development at the state level, we undertook detailed case studies in California, Indiana, Massachusetts, and Tennessee. These four states were selected to provide regional variation and to include representation from states receiving and not receiving Race to the Top funding.

8. Included in this database is an extensive collection of research reports, policy briefs, speeches, blog posts, press releases, media accounts, and congressional testimony related to the CCSS. These artifacts, which number approximately 1,500, provide historical and documentary background on the type of evidence cited, intended audience, salient policy issues, and links among organizations. Subsequent analyses of these data will focus on mapping information networks among various groups and individuals. However, in this analysis, they were used to verify that the full range of CCSS policy actors had been identified and to confirm interviewees' recollections about the specific documents they reported consulting during the CCSS process.

9. Among sources that the Hunt Institute staff drew upon are participant observer notes from weekly conference calls the institute convened among groups engaged in implementing the CCSS. The calls began in September 2010 with 7–14 "advocacy partners" typically participating each week. They included organizations representing elected officials, teachers and administrators, higher education, parents, and nonprofit third-party providers.

10. Interviewees were assured that their responses would be confidential and not be attributed to them, so only their role positions are noted in citing interview data. For a detailed discussion of sample selection and elite interviewing in process tracing research, see Tansey (2007).

11. In the late 1980s and early 1990s, several subject matter organizations, most prominently the National Council of Teachers of Mathematics (NCTM), developed academic content standards. Although controversial because of their constructivist ped-

agogical approach, NCTM's standards were influential in shaping individual state standards. For a summary of unsuccessful efforts to promote national standards during the 1990s, see Rothman (2011, 29–52).

12. The James B. Hunt Jr. Institute for Educational Leadership and Policy (the Hunt Institute) assists governors and other state leaders in improving their education policies and programs, largely through convening state leaders, disseminating information, and providing expert consultation. The Alliance for Excellent Education is a national advocacy organization working to improve high schools so that all students graduate prepared for postsecondary education and employment, especially those at risk of leaving high school without a diploma. Although both organizations work at all levels of the education system, the Alliance focuses primarily on federal policy and the Hunt Institute on working with state officials. They each receive financial support from a variety of funders, with foundations a major source for both.

13. We recognize that for a policy initiative as broad and far reaching as the Common Core, multiple bodies of research are potentially relevant, including ones focused on curriculum content, teaching and learning, assessment, and evaluations of the effects of a variety of policies in the United States and internationally on different groups of students. However, our purpose is not to catalog the supply of applicable research that could have informed the Common Core initiative. Rather, based on interview data and confirming documentary analysis, we identify what was reported as actually used in the process. For a number of sources that were central to policy deliberations, such as international comparisons of academic content standards and student achievement, participants in the CCSS process cited syntheses of original research rather than versions published in scholarly venues.

14. In 2007, the Hunt Institute requested that the NRC organize two workshops to examine available evidence on the ways in which standards-based accountability was currently functioning, criteria to use in evaluating common standards options, and the issues such an approach might raise. Although an NRC committee planned the workshop and commissioned papers, it was not intended to reach any conclusions or make any recommendations. However, the Hunt Institute drew on the research evidence presented at the workshops as an information source in advancing the case for common standards (Hunt Institute 2008a, 2008b).

15. Former governors Hunt and Wise were members of the group, as were Chester Finn Jr., the president of the Fordham Foundation; Kati Haycock, the president of the Education Trust; and William Schmidt, a university distinguished professor at Michigan State University and the director of TIMSS. Schmidt is the academic most associated with promoting the need for national content standards, and he functioned as a policy entrepreneur in advancing that argument.

16. Subsequently, after the CCSS were issued, critics began to argue that there is no research evidence documenting a causal relationship between standards and student performance on assessments and that there is no correlation between quality ratings of state standards and NAEP scores. States with content standards rated as weak by external bodies score about the same on NAEP as those with strong standards (Loveless 2012; Whitehurst 2009). Using a different analytical approach, Schmidt and Houang (2012) dispute this conclusion as it applies to the Common Core mathematics standards. Building on techniques developed as part of TIMSS, they first compare the proximity of state standards in 2009 to the CCSS in mathematics, and after adjusting for cut points on state assessments and controlling for state demographics related to poverty, they find that states with standards more like the CCSS in mathematics had higher performance on the 2009 NAEP. These analyses represent different ways of predicting the likely effect of the CCSS on student achievement, but the precise nature of that

Evidence Use and the Common Core State Standards Movement

relationship will not be known until the CCSS standards are actually implemented in classrooms across the country.

17. Although the systemic requirements and enabling conditions of standards policies were not given much prominence during the initial phase, policy entrepreneurs did acknowledge their role. For example, Governor Hunt (2009) noted in congressional testimony, “standards need to be supported by an integrated system, including curriculum, assessment, instruction, teacher preparation, and professional development. Unless our efforts reach the on-the-ground activity of teaching and learning, they will have been in vain. Standards-based reform was meant to be systemic reform.”

18. Learning trajectories or progressions are defined as “empirically supported hypotheses about the levels or waypoints of thinking, knowledge, and skill in using knowledge, that students are likely to go through as they learn mathematics and, one hopes, reach or exceed the common goals set for their learning. Trajectories involve hypotheses both about the order and nature of the steps in the growth of students’ mathematical understanding, and about the nature of the instructional experiences that might support them in moving step by step toward the goals of school mathematics” (Daro et al. 2011, 12). Researchers acknowledge the probabilistic nature of learning progressions and that existing ones require additional empirical examination (Sztajin et al. 2012).

19. An example from the ELA standards of the limited availability of existing research relates to text complexity. An ACT report (2006) and other research highlight the significance of students being able to master texts of increasing complexity as a predictor of their academic success in college and the workplace. Consequently, text complexity became one of the pillars of the ELA standards. While acknowledging its importance and appropriateness as a standard, literacy researchers note that “the underlying theory and research on text complexity that would support creation of state and district curricula and programs is in short supply” (Pearson and Hiebert 2012). One critical area where more robust research is needed to inform classroom practice is the measurement of text complexity. Although there are numerous formulas for measuring the readability of various types of texts, text complexity refers to more than the difficulty of words and includes structure, genre, and demands on prior knowledge. Because current formulas do not measure all these dimensions, and they are not calibrated to the ELA CCSS text complexity grade bands, the standards document calls for the development of new tools as quickly as possible. It also notes, “the use of qualitative and quantitative measures to assess text complexity is balanced in the Standards’ model by the expectation that educators will employ professional judgment to match texts to particular students and tasks” (NGA and CCSSO 2010a, 7).

20. Three people were responsible for drafting the mathematics standards: Phil Daro, a former director of the New Standards project; William McCallum, a professor of mathematics at the University of Arizona; and Jason Zimba, a professor of mathematics and physics at Bennington College. Those responsible for drafting the ELA standards were David Coleman, the founder of Student Achievement Partners, and Sue Pimentel, the cofounder of Standards Work. Each set of standards writers was assisted by a large work team (51 in mathematics and 50 in ELA) that included educators, researchers, and others with expertise in curriculum and assessment design, cognitive development, and English-language acquisition, who were called upon to provide input and review drafts on an as-needed basis. In addition, there was a feedback group for each set of standards consisting of members with expertise similar to that of the work group; they also reviewed drafts (21 for mathematics and 12 for ELA).

21. According to a participant in the standards-writing process, “the AFT teachers spent two days reviewing the standards. When we went to meet with the review team at the AFT, the math teachers had actually cut-up the standards and had deliberated

about whether the learning progressions made sense. So we started with the research, and they looked at the research and then reviewed the standards based on their experience. Some revisions we made were actually based on classroom teachers' experience" (personal interview).

22. In March 2010, NGA and CCSSO released a draft of the CCSS for public comment. Approximately 10,000 individuals, about half of whom were teachers, responded to an online survey. The overwhelming majority of respondents supported the concept of common standards, and their comments focused on areas that required clearer language, more examples, and greater detail. However, according to the summary report issued by CCSSO and NGA, a significant number of respondents perceived the CCSS to be federal standards and expressed opposition to them on principle. Another group, representing hundreds of respondents, pressed for health standards to be issued with the CCSS. One group questioned the CCSS on research grounds, arguing that they were developmentally inappropriate and at odds with the research on how children learn. They expressed concern that the standards, in starting with kindergarten students, placed too heavy an emphasis on academic knowledge and skills in the early grades and did not match the early learning standards for preK–3 that many states had adopted (CCSSO and NGA 2010).

23. The other standards development principles that served as criteria for the committee's validation of the CCSS were "evidence of the knowledge and skills that students need to be college and career ready; a proper level of clarity and specificity; [and] evidence that the standards are comparable with other leading countries' expectations" (NGA and CCSSO 2010b, 1).

24. One reason these groups were able to provide assistance is that 18 national organizations received over \$50 million between 2009 and 2010 from the Bill and Melinda Gates Foundation to provide information and help in implementing the CCSS. In addition, 10 third-party providers and 20 state and local education agencies received approximately \$39 million for the same purpose (information compiled from the Gates Foundation and funded organizations' websites).

References

- ACT. 2006. *Reading between the Lines: What the ACT Reveals about College Readiness in Reading*. Iowa City: ACT.
- Baumgartner, Frank R., and Bryan D. Jones. 1993. *Agendas and Instability in American Politics*. Chicago: University of Chicago Press.
- Beland, Daniel. 2005. "Ideas and Social Policy: An Institutional Perspective." *Social Policy and Administration* 39 (1): 1–18.
- Bennett, Andrew, and Colin Elman. 2006. "Qualitative Research: Recent Developments in Case Study Methods." *Annual Review of Political Science* 9:455–76.
- CCSSO and NGA (Council of Chief State School Officers and the National Governors Association). 2010. "Reactions to the March 2010 Draft Common Core State Standards: Highlights and Themes from the Public Feedback," <http://www.corestandards.org/assets/k-12-feedback-summary.pdf>.
- Clements, Douglas. 2011. "What Do We Know about Content and Curriculum?" In *Research in Mathematics Education: Where Do We Go from Here?* Conference Proceedings. Washington, DC: Institute for Research on Mathematics and Science Education, Michigan State University.

Evidence Use and the Common Core State Standards Movement

- Cohen, David K., and Susan L. Moffitt. 2009. *The Ordeal of Equality: Did Federal Regulation Fix the Schools?* Cambridge, MA: Harvard University Press.
- Collier, David, Henry E. Brady, and Jason Seawright. 2005. "Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology." In *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, ed. Henry E. Brady and David Collier. Lanham, MD: Rowman & Littlefield.
- Common Core State Standards Initiative. 2010. "Messaging Tool Kit," <http://programs.ccsso.org/link/CCSSI%20Toolkit%20Sept%202010.pdf>.
- Daro, Phil, Frederic A. Mosher, and Tom Corcoran. 2011. *Learning Trajectories in Mathematics: A Foundation for Standards, Curriculum, Assessment, and Instruction*. Research report R-68. New York: Consortium for Policy Research in Education.
- George, Alexander L. 1979. "The Causal Nexus between Cognitive Beliefs and Decision-Making Behavior: The 'Operational Code' Belief System." In *Psychological Models in International Politics*, ed. Lawrence S. Fulkowski. Boulder, CO: Westview.
- George, Alexander L., and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.
- Goertz, Margaret E. 2007. "Standards-Based Reform: Lessons from the Past, Directions for the Future." Paper presented at "Clio at the Table: A Conference on the Uses of History to Inform and Improve Education Policy," Brown University, Providence, RI, June.
- Hunt Institute. 2008a. "Blueprint for Education Leadership 1." Policy brief, June. James B. Hunt Jr. Institute for Educational Leadership, Durham, NC.
- Hunt Institute. 2008b. "Blueprint for Education Leadership 2." Policy brief, October. James B. Hunt Jr. Institute for Educational Leadership, Durham, NC.
- Hunt, James B., Jr. 2009. "Strengthening America's Competitiveness through Common Academic Standards." Prepared testimony before the Committee on Education and Labor of the United States House of Representatives, April 29.
- Kelman, Steven. 1987. *Making Public Policy*. New York: Basic.
- Kingdon, John W. 1995. *Agendas, Alternatives, and Public Policies*. 2nd ed. New York: HarperCollins.
- Kollman, Ken. 1998. *Outside Lobbying: Public Opinion and Interest Group Strategies*. Princeton, NJ: Princeton University Press.
- Light, Paul. 1999. *The President's Agenda: Domestic Policy Choice from Kennedy to Clinton*. Baltimore: Johns Hopkins University Press.
- Loveless, Tom. 2012. *How Well Are American Students Learning?* Brown Center on American Education. Washington, DC: Brookings Institution.
- Mahoney, James. 2000. "Strategies of Causal Inference in Small-N Analysis." *Sociological Methods and Research* 28:387-424.
- Majone, Giandomenico. 1989. *Evidence, Argument, and Persuasion in the Policy Process*. New Haven, CT: Yale University Press.
- Mintrom, Michael. 2000. *Policy Entrepreneurs and School Choice*. Washington, DC: Georgetown University Press.
- Mucciaroni Gary, and Paul J. Quirk. 2006. *Deliberative Choices: Debating Public Policy in Congress*. Chicago: University of Chicago Press.
- National Center for Education Statistics. 2007. *Mapping 2005 State Proficiency Standards onto the NAEP Scales*. Washington, DC: National Center for Education Statistics.
- NGA and CCSSO (National Governors Association and Council of Chief State School Officers). 2010a. "Appendix A: Research Supporting Key Elements of the Standards." In *Common Core State Standards for English Language Arts and Literacy in History/Social Studies, Science, and Technical Subjects*. Washington, DC: National Governors Association and Council of Chief State School Officers.

- NGA and CCSSO (National Governors Association and Council of Chief State School Officers). 2010b. *Reaching Higher: The Common Core State Standards Validation Committee*. Washington, DC: National Governors Association and Council of Chief State School Officers.
- NGA, CCSSO, and Achieve (National Governors Association, Council of Chief State School Officers, and Achieve). 2008. *Benchmarking for Success: Ensuring U.S. Students Receive a World-Class Education*. Washington, DC: National Governors Association, Council of Chief State School Officers, and Achieve.
- National Research Council. 2008. "Common Standards for K–12 Education? Considering the Evidence; Summary of a Workshop Series." Alexandra Beatty, rapporteur. Committee on State Standards in Education, Division of Behavioral and Social Sciences and Education, Washington, DC.
- Pearson, P. David, and Elfrieda H. Hiebert. 2012. "Understanding the Common Core State Standards." In *Teaching with the Common Core State Standards for English Language Arts, PreK–2*, ed. Lesley Mandel Morrow, Timothy Shanahan, and Karen K. Wixson. New York: Guilford.
- Quirk, Paul J., and Bruce Nesmith. 2011. "Reality-Based Policymaking: Information, Advice, and Presidential Success." In *Governing at Home: The White House and Domestic Policymaking*, ed. Michael Nelson and Russell L. Riley. Lawrence: University Press of Kansas.
- Ravitch, Diane. 2010. *The Death and Life of the Great American School System: How Testing and Choice Are Undermining Education*. New York: Basic.
- Rothman, Robert. 2011. *Something in Common*. Cambridge, MA: Harvard Education Press.
- Schmidt, William H., and Richard T. Houang. 2012. "Curricular Coherence and the Common Core State Standards for Mathematics." *Educational Researcher* 41 (8): 294–308.
- Sheingate, Adam. 2003. "Political Entrepreneurship, Institutional Change, and American Development." *Studies in American Political Development* 17:185–203.
- Stone, Deborah. 2012. *Policy Paradox*. 3rd ed. New York: Norton.
- Sztajin, Paola, Jere Confrey, P. Holt Wilson, and Cynthia Edgington. 2012. "Learning Trajectory Based Instruction: Toward a Theory of Teaching." *Educational Researcher* 41 (5): 147–56.
- Tansey, Oisín. 2007. "Process Tracing and Elite Interviewing: A Case for Non-probability Sampling." *PS: Political Science and Politics* 40 (4): 765–72.
- Weiss, Carol H. 1977. "Research for Policy's Sake: The Enlightenment Function of Social Research." *Policy Analysis* 3:531–45.
- Weiss, Carol H. 1982. "Policy Research in the Context of Diffuse Decision-Making." *Journal of Higher Education* 53 (6): 619–39.
- Whitehurst, Grover. 2009. "Don't Forget the Curriculum." Brown Center Letters on Education no. 3. Brookings Institution, Washington, DC, http://www.brookings.edu/papers/2009/1014_curriculum_whitehurst.aspx.
- Willhoit, Gene. 2009. "Improving Our Competitiveness: Common Core Education Standards." Prepared testimony before the Committee on Education and Labor, US House of Representatives, December 4.