

TWO CHEERS FOR EVIDENCE: LAW, RESEARCH, AND VALUES IN EDUCATION POLICYMAKING AND BEYOND

Eloise Pasachoff*

The newest federal education law, the Every Student Succeeds Act of 2015 (ESSA), reflects a recent turn toward basing social policy on research evidence. Proponents suggest that evidence-based policymaking in education and other social policy areas can help cut through ideological debate and provide meaningful limits on the choices made by the federal executive branch, states, and localities. This Essay argues that such hopes for evidence-based policymaking are overambitious. It first analyzes the evidence provisions in ESSA and demonstrates how little they constrain policy choices. It then assesses the limits of the research base in education, showing how little agreement there is on major research findings, how difficult it is to conduct this kind of research in the first place, and how challenging it is to implement the findings that exist. The Essay concludes by arguing that a major challenge for evidence-based policymaking is the deep divide among citizens and policymakers about the underlying purpose of education and other social policies in the first instance; many important policy debates are about values and cannot be resolved by appealing to research at all. This critique of evidence-based policymaking does not suggest that research is unimportant or that facts are only in the eye of the beholder but rather that claims for what evidence-based policymaking can accomplish in education and other social policy areas should be treated skeptically.

INTRODUCTION	1934
I. THE LIMITS OF THE LAW	1938
A. ESSA's Instructions for the Use of Evidence.....	1938

* Professor of Law, Georgetown University Law Center. Thanks to the *Columbia Law Review* editors for inviting me to participate in this Symposium honoring the legacy of Constance Baker Motley, and to the Symposium's other contributors and attendees for enriching conversations. Thanks also to Lily Faulhaber, Tom Glaisyer, Nora Gordon, Cara Jackson, Josef Konvitz, Chase Sackett, and Beryl Radin for constructive comments; to participants in the Reflections on Executive Power and Administrative Law Conference at the University of Wisconsin–Madison Law School for useful insight; and to Marietta Catsambas, Felycia Itza, and Claire Saba for helpful research assistance. This Essay is dedicated to the memory of my great-uncle Milton R. Konvitz, who was one of three lawyers working with Thurgood Marshall at the NAACP Legal Defense Fund when Constance Baker Motley joined the LDF staff in 1945. See Constance Baker Motley, *Equal Justice Under Law* 59 (1998) (describing Konvitz's role at LDF); see also David J. Danelski, *Rights, Liberties, and Ideals: The Contributions of Milton R. Konvitz* 15 (1983) (same). In 1964, Motley (then a state senator) and Konvitz (then on the faculty at Cornell) were joint winners of the NYU Washington Square College Alumni Association's achievement award for their respective work in the civil rights movement. See *College and School News*, 71 *Crisis* 340, 344 (1964). It is a privilege for me to pay tribute to them both at the same time in this Symposium Essay.

1. Not Directed at the Federal Government	1939
2. Limited Requirements for States, School Districts, and Schools	1940
3. Generous Definitions of Evidence	1942
4. Other, Broadly Permissive References to Evidence.....	1944
B. Opportunities to Find, Promote, and Use Evidence as Desired— Or to Ignore It	1946
1. State Choices.....	1946
2. Federal Choices	1948
II. THE LIMITS OF THE RESEARCH.....	1952
A. The Small and Contested Research Base	1953
B. The Complexities of Education Research.....	1958
C. The Research-to-Practice Dilemma	1961
III. THE LIMITS OF TECHNOCRACY.....	1966
CONCLUSION	1971

INTRODUCTION

The latest reauthorization of the 1965 Elementary and Secondary Education Act, the primary federal education law, is full of calls for evidence-based policymaking. Among other things, the Every Student Succeeds Act of 2015 (ESSA),¹ which replaced the once widely heralded but eventually widely maligned No Child Left Behind Act of 2001 (NCLB),² requires states to conduct school reform activities based on research evidence.³ ESSA incentivizes state education agencies and school districts to rely on evidence in a wide variety of other parts of their practice as well.⁴

This focus on evidence is part of a growing federal interest in evidence-based policymaking in education and other social policy programs over the last two decades.⁵ This trend gained a significant boost during the

1. Pub. L. No. 114-95, 129 Stat. 1802.

2. Pub. L. No. 107-110, 115 Stat. 1425; Pamela Barnhouse Walters & Annette Lareau, Introduction, *in* *Education Research on Trial: Policy Reform and the Call for Scientific Rigor* 1, 6 (Pamela Barnhouse Walters, Annette Lareau & Sheri H. Ranis eds., 2009) [hereinafter *Education Research on Trial*] (discussing the transition of NCLB from “[i]nitially widely celebrated” to “hotly-contested”).

3. See *infra* notes 48–61 and accompanying text.

4. See *infra* notes 73–97 and accompanying text.

5. See, e.g., Jim Nussle & Peter Orszag, Let’s Play Moneyball, *in* *Moneyball for Government* 2, 6–7 (Jim Nussle & Peter Orszag eds., 2015) (describing developments in evidence-based policymaking dating back to the George W. Bush Administration); see also *infra* notes 135–138 and accompanying text (describing legislation passed in 2002 encouraging evidence in education practices).

eight years of the Obama Administration,⁶ and it has been the subject of enthusiastic praise.⁷ ESSA's language on evidence is even being used as a model for other education laws that are currently up for reauthorization.⁸

One might think that requirements for evidence-based policymaking and practice would constrain the federal executive branch, now under President Trump, in overseeing cooperative federalism programs like ESSA. Perhaps a requirement for evidence will limit the White House and the Department of Education from making ideologically driven choices⁹—a potential respite from climate change denialism,¹⁰ presidential rejection of national security analysis,¹¹ and the circulation of conspiracy theories as fact¹² that otherwise seem to be dominating the new Administration.

6. See, e.g., Ron Haskins & Greg Margolis, *Show Me the Evidence: Obama's Fight for Rigor and Results in Social Policy 2–12* (2015) (describing the Obama Administration's evidence-based initiatives as novel for their breadth and interconnectedness).

7. See, e.g., *id.* at 21 (arguing that the Obama evidence agenda “has the potential to yield measurable improvement in the nation's social programs”); Nussle & Orszag, *supra* note 5, at 3–4 (arguing that “data, evidence, and evaluation [can] revolutionize America's government”); Robert Slavin, OMB to Government: Show Us the Evidence, *Education Week: Sputnik* (May 31, 2012), http://blogs.edweek.org/edweek/sputnik/2012/05/omb_to_government_show_us_the_evidence.html [<http://perma.cc/XA64-B7RP>] (suggesting that the Office of Management and Budget's memo to federal agencies asking them to promote evidence in their decisionmaking could “change history”).

8. See, e.g., *Strengthening Career and Technical Education for the 21st Century Act*, H.R. 2353, 115th Cong. § 7 (2017) (defining “evidence-based” as having the meaning given that term in ESSA); see also *infra* notes 63–69 and accompanying text (discussing this term in ESSA).

9. See, e.g., Frederick M. Hess & Bethany Little, “Moneyball” for Education: Using Data, Evidence, and Evaluation to Improve Federal Education Policy 2 (2015), <http://results4america.org/wp-content/uploads/2015/03/2015-3-18-Moneyball-for-Education-Report-1.pdf> [<http://perma.cc/94QM-EPRK>] (suggesting the evidence-based approach in education “could . . . present a bipartisan pathway forward at a time when much of education policy seems to be increasingly stuck in fruitless debate”); see also Ron Haskins, *Can Evidence Trump Ideology?*, *Brookings: Social Mobility Memos* (Dec. 5, 2014), <http://www.brookings.edu/blog/social-mobility-memos/2014/12/05/can-evidence-trump-ideology> [<http://perma.cc/DJJ2-GQUQ>] (arguing that evidence-based policymaking “offers a way to avoid . . . ideological roadblocks on at least some important social issues”).

10. See, e.g., Coral Davenport, E.P.A. Chief Doubts Consensus View of Climate Change, *N.Y. Times* (Mar. 9, 2017), <http://www.nytimes.com/2017/03/09/us/politics/epa-scott-pruitt-global-warming.html> (on file with the *Columbia Law Review*); Coral Davenport, E.P.A. Head Stacks Agency with Climate Change Skeptics, *N.Y. Times* (Mar. 7, 2017), <http://www.nytimes.com/2017/03/07/us/politics/scott-pruitt-environmental-protection-agency.html> (on file with the *Columbia Law Review*).

11. See, e.g., Greg Miller & Adam Entous, *Trump Turning Away Intelligence Briefers Since Election Win*, *Wash. Post* (Nov. 23, 2016), http://www.washingtonpost.com/world/national-security/trump-turning-away-intelligence-briefers-since-election-win/2016/11/23/5cc643c4b1ae-11e6-be1c-8cec35b1ad25_story.html?utm_term=.d1f24cda78ee (on file with the *Columbia Law Review*).

12. See, e.g., Michael Finnegan & Mark Z. Barabak, *How the Phony Conspiracy Theory over Wiretapping Caught Fire*, *L.A. Times* (Mar. 22, 2017), <http://www.latimes.com/nation/la-na-pol-trump-wiretap-conspiracy-20170322-story.html> [<http://perma.cc/5ATY-2CVV>].

Perhaps, too, the law's focus on evidence will provide meaningful substantive checks on the devolution to state and local authority that otherwise characterizes the move from No Child Left Behind to the Every Student Succeeds Act.¹³ If states and school districts must widely implement evidence-based interventions instead of relying on hunches, legacy programs, or ideology,¹⁴ perhaps education in America will be significantly improved.¹⁵ After all, the argument goes, "[m]uch good can be done with the evidence scientific studies produce," not only "for the most vulnerable students" but also "[m]ore broadly, [for] the public-education system" as a whole.¹⁶

As the rest of this Essay argues, however, one would be wrong to predict that ESSA's evidence requirements will impose any meaningful constraints on the federal executive, states, or school districts. It is also not likely that these requirements will themselves result in significant improvements to education in America or resolve raging debates about the best way to engage in school reform.

Why not? First, as Part I illustrates, although ESSA's references to evidence-based decisionmaking are numerous, its actual requirements for evidence-based decisionmaking are anything but stringent.¹⁷ The Department of Education, states, and school districts retain many opportunities to push their preferred policy choices.¹⁸

Moreover, as Part II lays out, requirements to identify "what works" overestimate the current knowledge base and potential for future research,

13. Derek W. Black, *Abandoning the Federal Role in Education: The Every Student Succeeds Act*, 105 *Calif. L. Rev.* (forthcoming 2017) [hereinafter Black, *Abandoning the Federal Role*] (manuscript at 124–26), <http://papers.ssrn.com/abstract=2848415> (on file with the *Columbia Law Review*) (lamenting devolution while identifying evidence provisions as a "silver lining").

14. See Lance D. Fusarelli, *Flying (Partially) Blind: School Leaders' Use of Research in Decisionmaking*, in *When Research Matters: How Scholarship Influences Education Policy* 177, 185 (Frederick M. Hess ed., 2008) [hereinafter *When Research Matters*] ("Decisionmaking and program adoption in education is shaped by and often determined by ease of use, good marketing, lack of threat to current practice, 'philosophical commitments, political necessities, and the attractiveness or popularity of ideas' rather than research-based evidence of program effectiveness." (quoting Tom Corcoran, *The Use of Research Evidence in Instructional Improvement 2* (2003), http://www.cpre.org/sites/default/files/policybrief/880_rb40.pdf [<http://perma.cc/6WUM-CJ4L>])); Paul Manna & Michael J. Petrilli, *Double Standard? "Scientifically Based Research" and the No Child Left Behind Act*, in *When Research Matters*, supra, at 63, 64 (noting that "elected officials will rely much on gut instincts, ideology, riveting anecdotes, opinion polls, or the need to repay favors to colleagues . . . when formulating their positions").

15. See, e.g., Robert Balfanz, *How Boosting Education Research Could Revolutionize US Schooling*, *Stan. Soc. Innovation Rev.* (Mar. 14, 2017), http://ssir.org/articles/entry/how_boosting_education_research_could_revolutionize_us_schooling [<http://perma.cc/9AB5-KRU9>] (arguing that ESSA's focus on evidence "presents a strong path for strengthening public schooling").

16. *Id.*

17. See *infra* section I.A.

18. See *infra* section I.B.

and they underestimate the challenges of implementation in the messy, human endeavor of teaching and learning that goes on every day in thousands of jurisdictions all over America. While the analysis in this Part focuses on education research, the critiques apply more broadly to social science research in general.

Finally, as Part III explains, the elusive search for “what works” can obscure critical and often contested values-based decisions about what we should be trying to do in the first place and about necessary tradeoffs along the way. This Part, too, focuses on education policy, but the argument extends broadly to other kinds of social policy.

This Essay should not be read as a critique of the evidence requirements in ESSA but rather as a caution against the fetishization of evidence in social science policymaking more generally. Too often, policymakers envision what evaluation scholars call an “instrumental role” for evidence, in which research will provide an uncontested answer to a clear policy problem.¹⁹ But research findings almost never provide a straightforward answer about what to do,²⁰ and decisions about education and other social policies are enmeshed in a complex web of political realities, institutional capacities, and societal values.²¹ Too often, then, decision-makers on the ground end up using evidence in what evaluation scholars call a “symbolic” way, “creat[ing] legitimacy for solutions that are already favored or even enacted,” thereby largely undermining the point of requiring evidence-based decisions in the first place.²²

Instead, evidence plays the most valuable role as what evaluation scholars call a “conceptual” tool.²³ In the words of Carol Weiss, a pioneering scholar in the evaluation of social policy:

Social research can be “used” in reconceptualizing the character of policy issues or even redefining the policy agenda. Thus, social research may sensitize decisionmakers to new issues and turn what were nonproblems into policy problems In turn, it may convert existing social problems into nonproblems It may drastically revise the way that a society thinks about issues . . . , the facets of an issue that are viewed as susceptible to alteration, and the alternative measures that it considers.²⁴

19. Cynthia E. Coburn, Meredith I. Honig & Mary Kay Stein, What’s the Evidence on Districts’ Use of Evidence?, *in* *The Role of Research in Educational Improvement* 67, 75 (John D. Bransford, Deborah J. Stipek, Nancy J. Vye, Louis M. Gomez & Diana Lam eds., 2009).

20. See *infra* notes 141–186 and accompanying text.

21. See *infra* notes 222–249 and accompanying text.

22. See Coburn et al., *supra* note 19, at 77.

23. See *id.* at 76–77.

24. Charles E. Lindblom & David K. Cohen, Usable Knowledge: Social Science and Social Problem Solving 77 (1979) (quoting Carol H. Weiss, Introduction, *in* *Using Social Research in Public Policy Making* 1, 15–16 (Carol H. Weiss ed., 1972)); see also Kathryn E. Newcomer, Carol H. Weiss, Evaluation Research: Methods for Studying Programs and Policies, *in* *The Oxford Handbook of Classics in Public Policy and Administration* 326, 326–41

Evidence can more usefully change the way we think about a problem of social policy than it can tell us what to do about that problem.²⁵ In this way, it is an important part of policy analysis. But it should be viewed as only one component of the democratic deliberation and political contestation in which education, like other social policies, is rooted.

The focus on evidence-based policymaking in education has some valuable aspects, then, but it should not be seen as a panacea or as a neutral, technocratic solution. Evidence alone is insufficient to resolve complex social problems with contested underlying values and goals, and we ought not expect that it will.

I. THE LIMITS OF THE LAW

Notwithstanding the frequency of their appearance, the evidence requirements in ESSA are limited.²⁶ The Department of Education, states, and local school districts have many opportunities to promote whatever policies they would like to promote for whatever reason.²⁷

A. *ESSA's Instructions for the Use of Evidence*

ESSA refers to evidence numerous times throughout its thousand pages. Depending on how one counts such a reference, there may be as many as fifty-eight such instances.²⁸ However, the actual requirements for evidence-based decisionmaking under ESSA are few. As this section explains, the references to evidence are (1) not directed at the federal government; (2) include only limited requirements for states, school districts, and schools; (3) provide a generous definition of what counts as “evidence”; and (4) are in most cases merely suggestions.

(Steven J. Balla, Martin Lodge & Edward C. Page eds., 2015) (analyzing Weiss’s influential work in the field of evaluation theory and practice).

25. See, e.g., Lindblom & Cohen, *supra* note 24, at 48 (noting that research typically “raise[s] new issues, stimulate[s] new debate, and multipl[ies] the complexities of the social problem at hand”); David K. Cohen, Stephen W. Raudenbush & Deborah Loewenberg Ball, Resources, Instruction, and Research, *in* *Evidence Matters: Randomized Trials in Education Research* 80, 117 (Frederick Mosteller & Robert Boruch eds., 2002) [hereinafter *Evidence Matters*] (noting that research can “inform[] thought and debate . . . [and] might tend to close out unfruitful arguments as well as highlight new problems”); Frederick M. Hess, Conclusion: Education Research and Public Policy, *in* *When Research Matters*, *supra* note 14, at 239, 256 [hereinafter Hess, Conclusion] (“[R]esearch cannot provide all the answers officials seek—but is frequently most valuable when it helps shed[] new light on problems and favored solutions.”).

26. See *infra* section I.A.

27. See *infra* section I.B.

28. Memorandum from the Penn Hill Grp. on Evidence-Based and Its Use in ESSA to the Council of Chief State Sch. Officers (CCSSO) (2016), <http://www.ccsso.org/Documents/2016/ESSA/ESSAEvidenceBasedSummaryAndAnalysis.pdf> [<http://perma.cc/WHT5-CHJJ>]; see also Results for America, Scan of the Evidence Provisions in the Every Student Succeeds Act (ESSA) (2016), <http://results4america.org/wp-content/uploads/2017/02/RFA-ESSA-Scan.pdf> [<http://perma.cc/KUG9-3HN5>].

1. *Not Directed at the Federal Government.* — It is important first to distinguish among the entities that might be subject to ESSA's evidence instructions: the federal executive branch, state education authorities, and local school districts and schools.

Unlike in many other regulatory programs, with broad delegations to federal agencies to regulate a policy area “in the public interest” or in a manner that is “fair and equitable,” congressional delegations to the Department of Education largely authorize the agency to award grants to states and districts under terms that are intricately detailed by statute.²⁹ Because of this structure, there are no directions to the Department of Education to use the “best available evidence” to set education policy.³⁰ Instead, the provisions discussing evidence are largely directed at states and school districts.³¹ ESSA's focus on evidence will therefore do little to constrain the Department of Education.

To be sure, plenty of other parts of ESSA limit what the Secretary of Education can do. For example, the Secretary may not reject a state's funding application for any substantive reason; the Secretary is limited to asking for compliance with the technical requirements laid out by statute.³² The Secretary may not require states or school districts seeking a waiver under any part of the Act to commit to taking any particular steps in order to receive the waiver.³³ And the Secretary is prevented from promulgating expansive rules or issuing detailed nonregulatory guidance on

29. See Derek W. Black, *Federalizing Education by Waiver?*, 68 *Vand. L. Rev.* 607, 676–78 (2015) (describing cases upholding delegations to agencies to regulate “in the public interest” in commodities price-fixing and in “fair and equitable” ways in broadcasting regulations, and contrasting such broad delegations with the narrower delegation to the Department of Education); Eloise Pasachoff, *The President's Budget as a Source of Agency Policy Control*, 125 *Yale L.J.* 2182, 2205 & n.101 (2016) [hereinafter Pasachoff, *The President's Budget*] (describing broad delegations to agencies to regulate “to protect the public health” with respect to air quality and “to promote honesty and fair dealing in the interest of consumers” with respect to food safety, and contrasting such broad delegations with the narrower delegation to the Department of Education).

30. Cf. 16 U.S.C. § 1533(b)(1)(A) (2012) (requiring the Secretary of the Department of the Interior to make decisions under the Endangered Species Act using “the best scientific and commercial data available”); 33 U.S.C. § 1311(b)(2)(A) (2012) (directing the Administrator of the Environmental Protection Agency to consider the “best available technology economically achievable” in regulating “effluent limitations” under the Clean Water Act).

31. Chiefs for Change, *ESSA and Evidence: Why It Matters 1* (2016), <http://chiefsforchange.org/wp-content/uploads/2016/07/ESSA-and-Evidence-Why-It-Matters.pdf> [<http://perma.cc/JJ2U-XS9F>] (“ESSA incentivizes states to use evidence-based programs and interventions in districts and schools.”); Martin R. West, *From Evidence-Based Programs to an Evidence-Based System: Opportunities Under the Every Student Succeeds Act*, Brookings Inst. (Feb. 5, 2016), <http://www.brookings.edu/research/from-evidence-based-programs-to-an-evidence-based-system-opportunities-under-the-every-student-succeeds-act> [<http://perma.cc/4VA2-VPZG>] (explaining that ESSA's focus on evidence may “sustain a new model for decision-making within state education agencies and school districts”).

32. 20 U.S.C. § 7871(c) (Supp. 2015).

33. *Id.* § 7861(b)(4), (d)(3).

particular components of state compliance with the Act.³⁴ This aspect of ESSA grows out of congressional charges, especially from Republicans, that President Obama's Department of Education aggrandized both the federal role in general and the executive branch's authority in particular.³⁵

But these provisions bar the Secretary from certain kinds of policy-making entirely, rather than imposing evidence requirements around the scope of policymaking. And, as discussed below, other opportunities remain for the Trump Administration (or any administration) to get its policy preferences through.³⁶

2. *Limited Requirements for States, School Districts, and Schools.* — ESSA contains two types of grants for the Department of Education to oversee: formula grants and competitive grants.³⁷ Formula grants are those to which all eligible entities, whether states or school districts, are entitled, assuming they submit an application that explains how they will do what the relevant provisions tell them to do.³⁸ Competitive grants, on the other hand, are those under which money is not guaranteed even after submitting an application.³⁹

The core part of ESSA, and its predecessors since 1965, is a formula grant called Title I, Part A (or sometimes just Title I). Title I contains the key provisions on standards, testing, and accountability with which states, districts, and schools must comply,⁴⁰ since all states receive Title I funds.⁴¹ As this subsection explains, ESSA's only requirements for evidence appear in Title I. ESSA contains other, smaller formula grants, as well as a wide variety of competitive grants, but as section I.A.4 explains below, the focus on evidence in these grants is permissive rather than mandatory.

34. *Id.* §§ 6311(e), 7915.

35. See, e.g., Patrick McGuinn, *From No Child Left Behind to the Every Student Succeeds Act: Federalism and the Education Legacy of the Obama Administration*, 46 *Publius: J. Federalism* 392, 401–02, 408–09 (2016) (describing the mobilization against “federal overreach” that followed the Obama Administration’s support for the Common Core State Standards and aligned standardized tests).

36. See *infra* section I.B.2.

37. See, e.g., Eloise Pasachoff, *Agency Enforcement of Spending Clause Statutes: A Defense of the Funding Cut-Off*, 124 *Yale L.J.* 248, 268 (2014) [hereinafter Pasachoff, *Agency Enforcement*] (categorizing grant types using the Department of Education as an example).

38. *Id.*

39. *Id.*

40. See, e.g., Black, *Abandoning the Federal Role*, *supra* note 13 (manuscript at 114–26) (describing how ESSA’s regime for standards, testing, and accountability in Title I applies to states, districts, and schools and is the central part of the Act); Lorraine M. McDonnell, *No Child Left Behind and the Federal Role in Education: Evolution or Revolution?*, 80 *Peabody J. Educ.* 19, 22–33 & n.2 (2005) (describing the evolution of Title I as the key feature of the original Elementary and Secondary Education Act of 1965 and subsequent reauthorizations).

41. See, e.g., Rebecca R. Skinner & Leah Rosenstiel, *Cong. Research Serv.*, R44486, *FY2016 State Grants Under Title I-A of the Elementary and Secondary Education Act (ESEA) 4–6 tbl.1* (2017), <http://fas.org/sgp/crs/misc/R44486.pdf> [<http://perma.cc/5MN9-D9YV>] (listing each state’s share of Title I grants).

The requirements for evidence-based practices appear in only two places in Title I: in the development of school districts' plans to engage parents and families,⁴² and in the development of school "support and improvement plan[s]" for a small group of schools that are not adequately meeting the state's accountability requirements.⁴³ Neither of these requirements is broadly systemic.

The first of these activities—designing parent engagement plans—is not a central part of the Act; it is merely one of twenty-odd things a local school district must include in its application to receive a subgrant under ESSA from the state.⁴⁴ There is no real consequence for failure to do any of the things the school districts say they will do in asking for their subgrant.⁴⁵ The requirement for the parent engagement plan to explain how the school district will conduct a needs assessment and then use the findings "to design evidence-based strategies for more effective parental involvement"⁴⁶ thus cannot be seen as transformative or constraining in any way.

The second of these activities—designing school improvement plans—is a much bigger deal, as these plans are core to the accountability regime that lies at the center of the Act.⁴⁷ Here, too, however, the calls for evidence-based interventions are not that significant.

There are two situations in which school improvement plans are required. The first situation is narrower, based on low achievement of one or more "subgroups" in a school.⁴⁸ "Subgroup of students" is a term of art in the statute, as it was in No Child Left Behind,⁴⁹ meaning "economically disadvantaged students," "students from major racial and ethnic groups," "children with disabilities," and "English learners."⁵⁰ For all schools in which one or more of these subgroups "are consistently underperforming"⁵¹ under the state's accountability requirements—test scores, graduation rates for high schools, progress in achieving English proficiency, and any "other indicator of school quality or student success" that the state wishes to adopt⁵²—the state educational agency must inform the local school district of these schools' existence and must ensure that the

42. See 20 U.S.C. §§ 6312(b)(7), 6318(a)(2)(E) (Supp. 2015).

43. See *id.* § 6311(d)(1)(B).

44. See *id.* § 6312(a), (b)(7).

45. See *id.* §§ 6312(a)(5), 6318(h) (providing for local review but not state consequences); see also Pasachoff, Agency Enforcement, *supra* note 37, at 284 (describing infrequency of funding cut-offs for a grantee's noncompliance with terms of a grant).

46. 20 U.S.C. § 6318(a)(2)(D)–(E).

47. See, e.g., Black, Abandoning the Federal Role, *supra* note 13 (manuscript at 115–17, 124–26) (contrasting NCLB and ESSA accountability regimes as the central parts of each Act).

48. 20 U.S.C. § 6311(d)(2).

49. 20 U.S.C. § 6311(b)(2)(C)(v)(II) (2002).

50. 20 U.S.C. § 6311(c)(2) (Supp. 2015).

51. *Id.* § 6311(d)(2)(A)(ii).

52. *Id.* § 6311(c)(4)(B)–(C).

school district in turn notifies the schools.⁵³ Each school must then work with community “stakeholders” to develop a “targeted support and improvement plan” that “includes evidence-based interventions” to remedy the problem for the relevant subgroup(s).⁵⁴ This targeted plan must then be approved and subsequently monitored by the school district.⁵⁵

The second situation that requires a school improvement plan is broader, based on the performance of the school as a whole.⁵⁶ For at least those schools that perform the worst in the state on tests or on high school graduation rates, as well as for those schools in which a subgroup’s performance on its own would rank at the bottom of the state’s performance requirements,⁵⁷ the state educational agency must again identify these schools to their school district.⁵⁸ Now, instead of simply notifying the schools that they must implement an improvement plan, the school district itself has to work with community “stakeholders” to develop a school-level “comprehensive support and improvement plan” that “includes evidence-based interventions” to fix the school’s achievement problems.⁵⁹ This comprehensive plan must be approved not only by the individual school and school district but also by the state educational agency,⁶⁰ which must also then monitor its implementation.⁶¹

These provisions constitute the sum total of the evidence requirements in the Act. Note the limited categories of schools and students to which the more significant requirements—those for school improvement plans—apply. The requirements apply only to decisions about how to respond to the persistently low performance of certain subsets of children and extremely small percentages of schools. While this is an important category of children and schools, these requirements are simply not going to significantly transform or constrain the education decisions of states, districts, and schools as a whole.⁶²

3. *Generous Definitions of Evidence.* — To be sure, these requirements could, in principle, play a dramatic role for the neediest children and schools if the term “evidence-based” itself imposed meaningful constraints on decisionmaking. That is, if “evidence-based” closed off a significant universe of options, then it could indeed be constraining. But it does not.

53. Id. § 6311(d)(2)(A).

54. Id. § 6311(d)(2)(B).

55. Id.

56. Id. § 6311(d)(1).

57. Id. § 6311(c)(4)(D)(i).

58. Id. § 6311(d)(1)(A).

59. Id. § 6311(d)(1)(B).

60. Id. § 6311(d)(1)(B)(v).

61. Id. § 6311(d)(1)(B)(vi).

62. Cf. Black, *Abandoning the Federal Role*, supra note 13 (manuscript at 125–26) (criticizing ESSA’s accountability regime for applying to “almost no one,” and describing requirements for evidence-based interventions as positive but nonetheless “relatively minor” in light of the vast flexibility otherwise granted to states).

The statute devotes several paragraphs to a definition of the term “evidence-based.”⁶³ ESSA presents four possible ways, with different degrees of strenuousness, by which an entity may show that “an activity, strategy, or intervention” is “evidence-based.”⁶⁴

The three most demanding ways require a grantee to “demonstrate[] a statistically significant effect on improving student outcomes or other relevant outcomes,”⁶⁵ with different tiers of evidence required to make that showing: There must be either (1) “strong evidence from at least 1 well-designed and well-implemented experimental study”; (2) “moderate evidence from at least 1 well-designed and well-implemented quasi-experimental study”; or (3) “promising evidence from at least 1 well-designed and well-implemented correlational study with statistical controls for selection bias.”⁶⁶ These are the tiers of evidence that must apply to the school improvement plans and related state activities for the lowest-achieving schools and student subgroups.⁶⁷

There is a fourth, less demanding tier of evidence. To be evidence based under this tier, an “activity, strategy, or intervention” need not be statistically significant, but must merely “demonstrate[] a rationale based on high-quality research findings or positive evaluation that such activity, strategy, or intervention is likely to improve student outcomes or other relevant outcomes.”⁶⁸ Such an “activity, strategy, or intervention,” if selected by an educational authority, must also “include[] ongoing efforts to examine the effects of such activity, strategy, or intervention.”⁶⁹ This fourth tier is applicable to the requirement for evidence in developing parent and family engagement plans; such plans may be based on evidence in one of the top three tiers, but need not be, as long as they are based on evidence that would satisfy this fourth tier.⁷⁰

As Part II below will make clear, it is not difficult to find evidence to support a vast range of educational interventions under these definitions, even under the top three tiers that are applicable to school improvement plans. Tier three (“promising evidence”) and tier four (a “rationale” with a promise to assess it as implemented) are especially easy to meet. To be sure, the commitment that grantees make under tier four to engage in ongoing assessment of their activities could encourage them to make better decisions—but the reality is that only those grantees who already

63. 20 U.S.C. § 7801(21)(A)–(B).

64. *Id.* § 7801(21)(A).

65. *Id.* § 7801(21)(A)(i).

66. *Id.* § 7801(21)(A)(i)(I)–(III).

67. *Id.* § 7801(21)(B).

68. *Id.* § 7801(21)(A)(ii)(I).

69. *Id.* § 7801(21)(A)(ii)(II).

70. Compare *id.* § 7801(21)(B) (defining “evidence-based” activities as only those that satisfy one of the top three tiers of evidence), with *id.* § 6318(a)(2)(E) (imposing no such limitation).

want to engage in the process of continuous improvement are likely to take this opportunity seriously.⁷¹

Moreover, to the extent ESSA's evidence requirements include constraints, the constraints are not on determining what state, district, and school educational goals should be, but only on how to achieve those goals. Because ESSA as a whole is shot through with providing substantive discretion to states, districts, and schools,⁷² the statute's evidence definitions do little to limit policy choices.

4. *Other, Broadly Permissive References to Evidence.* — Beyond the parent engagement plans and the school improvement plans in Title I, the other references to evidence in ESSA are all permissive rather than mandatory. This permissiveness will not fundamentally transform practices or cabin discretion.

For example, consider the references to evidence in three smaller formula grants: one to support at-risk youth;⁷³ one to support teachers and principals;⁷⁴ and one for "student support and academic enrichment."⁷⁵ States need not apply for these grants at all, but if they do, and if they want to engage in certain activities identified in the statute, they must simply agree to find evidence-based ways to do so.⁷⁶

These provisions may do some work to promote evidence-based thinking, especially when the items on the list are popular interventions

71. See West, *supra* note 31 (noting "[t]hose six words [“ongoing efforts to examine the effects’ of the activity on important student outcomes”], if taken seriously and implemented with care, hold the potential to create and provide resources to sustain a new model for decision-making within state agencies and school districts,” but “[t]he opportunity to use federal funds for evaluation purposes will only make a difference if state officials choose to exploit it.”).

72. Black, *Abandoning the Federal Role*, *supra* note 13 (manuscript at 127–29).

73. 20 U.S.C. §§ 6431–6439 (Title I, Part D, Subpart 1).

74. *Id.* §§ 6611–6614 (Title II, Part A).

75. *Id.* §§ 7111–7122 (Title IV, Part A, Subpart 1).

76. *Id.* § 6434(c)(20)(B) (providing that services and interventions for youth who have been in contact with both the child welfare and juvenile justice systems should be evidence based); *id.* § 6611(c)(4)(B)(v)(I), (vii)(III) (permitting states to use funds under the teacher and principal grant to allow “effective teachers to lead evidence-based . . . professional development” for their peers and to develop “induction and mentoring programs” for new teachers and principals that are evidence based); *id.* § 6613(b)(3)(D), (E) (providing that school districts may use funds under the teacher and principal grant to “reduc[e] class size to a level that is evidence-based” and to provide “high-quality, personalized professional development that is evidence-based”); *id.* § 7114(b)(3)(B)(ii)(I), (B)(iii), (C)(iii) (permitting states to use funds under the student support and academic enrichment grant to help school districts implement “mental health awareness training programs,” integrate “health and safety practices into school or athletic programs,” and deliver “specialized or rigorous academic courses and curricula through the use of technology”); *id.* §§ 7118(5)(A), (B)(ii)(II)(aa), (F)(ii) (providing that school districts may use their funds on evidence-based programs in “drug and violence prevention activities,” “school-based mental health services,” and in “reduc[ing] exclusionary discipline practices”).

like class size reductions.⁷⁷ But the reach of the work should not be overstated. The statute includes the caveat that evidence need be used only “to the extent the State determines that such evidence is reasonably available.”⁷⁸ The statute also identifies many other potential activities for which these funds may be used without requiring that states tie those activities to evidence at all.⁷⁹ And given the minimal requirements in the statutory definition of “evidence-based” itself,⁸⁰ all of these references to evidence could be satisfied with not much more than an articulated rationale based on some research found somewhere. These provisions may refer to evidence, then, but they are hardly constraining.

The same is true of references to evidence in ESSA’s competitive grants. While certain competitive grants encourage applicants to embed evidence-based activities in their programs, the encouragement is not particularly strenuous. For example, some grants provide “priority points” for applicants who can demonstrate that their proposed activities are supported by evidence from the top three tiers, but they do not require applicants to make such a demonstration.⁸¹ Some further limit a portion (but not all) of any federal funds awarded to those activities that are generally “evidence-based” (including the bottom tier of evidence).⁸² Some also require ongoing evaluation of the extent to which the grant programs

77. Id. § 6613(b)(3)(D); see also Dan D. Goldhaber & Dominic J. Brewer, What Gets Studied and Why: Examining the Incentives that Drive Education Research, *in* When Research Matters, supra note 14, at 197, 197 (discussing the popularity of this policy).

78. 20 U.S.C. § 6434(c)(20)(B); id. § 6611(c)(4)(B)(v)(I), (vii)(III), (xxi); id. § 6613(b)(3)(D), (P); id. § 7114(b)(3)(B)(ii)(I), (iii), (C)(iii).

79. Id. § 6434(c) (listing twenty aspects of a plan a state must describe if it wants certain funds for “neglected, delinquent, and at-risk children and youth,” only one of which mentions the use of evidence); id. § 6611(c)(4)(B) (listing twenty-one allowable uses of funds for state activities under the teacher and principal grant, only three of which mention the use of evidence); id. § 6613(b)(3) (listing sixteen allowable uses of funds for school districts receiving subgrants under this grant, only four of which mention the use of evidence); id. § 7114(b) (listing seventeen allowable uses of funds for state activities under the student support and academic enrichment grant, only three of which mention the use of evidence); id. § 7118 (listing twenty-eight allowable uses of funds under this grant, only three of which mention the use of evidence).

80. See supra notes 63–70 and accompanying text.

81. 20 U.S.C. § 6642(e) (grants for “[c]omprehensive literacy State development”); id. § 6672(e) (grants for “[s]upporting effective educator development”); id. § 6673(e) (grants for “[s]chool leader recruitment and support”); id. § 7243(c) (grants for “statewide family engagement centers”); id. § 7274(b) (grants for “[p]romise neighborhoods”); id. § 7275(b) (grants for “full-service community schools”); id. § 7294(f) (grants for “[s]upporting high-ability learners and learning”); see also id. § 7231e(2)(A) (listing evidence-based projects as one of several ways applicants for grants to support magnet schools may establish priority).

82. Id. § 6672(a)(2), (5) (grants for “[s]upporting effective educator development”); id. § 6673(a)(6) (grants for “[s]chool leader recruitment and support”); id. §§ 7242, 7243(b)(6)(G) (grants for “statewide family engagement centers”).

are successful.⁸³ One is designed to encourage broad efforts to “create, develop, implement, replicate, or take to scale entrepreneurial, evidence-based, field-initiated innovations to improve student achievement and attainment for high-need students.”⁸⁴

To embed incentives for evidence across numerous multimillion dollar grants is not nothing.⁸⁵ But nor is it everything. There remain many competitive grants in ESSA that do not require any particular evidence-based decisionmaking to support an application.⁸⁶ The eligible entities for some of the grants encouraging evidence-based thinking are not even state educational authorities, school districts, or schools,⁸⁷ so the effect of these evidence requirements on the nation’s core educational institutions is further diluted. And, of course, the nature of competitive grants is that not every eligible entity will apply, and certainly not every entity that applies will win one.

The permissive references to evidence throughout ESSA thus may do more to express a mood than to provide an actual set of transformative constraints.

B. *Opportunities to Find, Promote, and Use Evidence as Desired—Or to Ignore It*

To illustrate the lack of constraints imposed by ESSA’s references to evidence, it is worth considering the ways in which the states and the federal government can continue to push their own policy preferences within the bounds of the law.

1. *State Choices.* — Because states are the primary recipients of funding under ESSA—most local money is allocated through subgrants⁸⁸—

83. Id. § 7274(h), (i) (grants for “[p]romise neighborhoods”); id. § 7275(f), (g) (grants for “[f]ull-service community schools”); id. § 7294(h) (grants for “[s]upporting high-ability learners and learning”).

84. Id. § 7261(a)(1)(A) (grants for “education innovation and research”).

85. See, e.g., id. § 6603(b) (authorizing close to \$500 million each year for a variety of competitive grants); id. § 7246 (authorizing \$10 million each year for another competitive grant); id. § 7251(a) (authorizing between \$200 and \$221 million each year for another set of competitive grants).

86. See, e.g., id. § 6662 (grants for “Presidential and Congressional Academies for American History and Civics”); id. § 7221b(g) (grants to support “high-quality charter schools”); id. § 7281 (grants for “[n]ational activities for school safety”); id. § 7292 (grants for “assistance for arts education”); id. § 7293 (grants for “[r]eady to learn programming”).

87. See, e.g., id. § 6672(f) (including institutions of higher education and nonprofits as eligible entities for “[s]upporting effective educator development” grants); id. § 7261(b) (including nonprofits acting on their own as well as a governmental agency acting in partnership with a nonprofit, a business, “an educational service agency,” or “an institution of higher education” as eligible entities for “education innovation and research” grants); id. § 7272(1) (including institutions of higher education as well as nonprofits working in partnership with certain other entities as eligible entities for “[p]romise neighborhood[]” grants, and including a consortium of community-based organizations or nonprofit organizations as eligible entities for “full-service community school[]” grants).

88. See, e.g., id. § 6312(a)(1) (authorizing subgrants to local school districts).

state authority to steer school districts and schools toward state priorities for education is significant.

As should be clear from the discussion of ESSA's evidence provisions above, there are major parts of the Act that simply require no evidence whatsoever. States need not justify their standards with respect to any evidence of what works.⁸⁹ The same is true of their tests,⁹⁰ their teacher certification plans,⁹¹ their plans for teaching English learners⁹²—the list goes on.⁹³ States can simply make assertions in their applications for formula grants about how they plan to work on these issues without examining any research evidence first.⁹⁴

That some aspects of the Act call for evidence while others do not is notable. It might suggest that some aspects of education are too complicated to demand evidence for; that in some categories of action, federalism trumps evidence; or simply that political dealmaking reached the median voter here. Whatever the reason, it is clear that in vast swaths of their educational undertakings, states need not examine research but may simply make policy choices.

When states do have to use research evidence, ESSA also provides numerous opportunities for states to require districts and schools to adopt whatever evidence-based interventions the state prioritizes.⁹⁵ This type of activity illustrates what evaluation scholars call a “sanctioning role” for evidence, when state or federal agencies deem particular interventions “research-based” and thus approved for use with state or federal funds.⁹⁶ Here, too, then, states are empowered to select the evidence they find most persuasive to support the interventions they want to support anyway.

Another way states can specify the evidence-based interventions they prefer is designing criteria for ESSA's school improvement subgrants to local educational agencies.⁹⁷ Controlling the purse strings provides authority

89. See *id.* § 6311(b)(1)(A) (detailing requirements for “challenging” standards without requiring any evidentiary basis).

90. See *id.* § 6311(b)(2)(A)–(B) (detailing requirements for “high-quality” assessments without requiring any evidentiary basis).

91. See *id.* § 6311(g)(2)(J) (detailing assurances a state must offer on teacher certification without requiring any evidentiary basis).

92. See *id.* § 6311(b)(1)(F), (b)(2)(G) (detailing requirements for standards and assessments for English learners without requiring any evidentiary basis).

93. See *id.* § 6311 (detailing requirements for state plans in general but including evidence requirements only at § 6311(d)(1)–(2)); see also *supra* notes 51–61 and accompanying text (describing these evidence requirements).

94. See, e.g., 20 U.S.C. § 6311(g)(1)–(2) (detailing descriptions states must offer and assurances they must provide without requiring any consultation of evidence).

95. See *id.* § 6311(d)(3)(B)(ii).

96. See Coburn et al., *supra* note 19, at 75, 78.

97. See 20 U.S.C. § 6303(a)–(b).

to make policy choices without much constraint from the requirement for evidence.

2. *Federal Choices.* — There are numerous ways in which the Trump Administration can implement ESSA's evidence provisions while still pushing its own agenda, whether through regulation, nonregulatory guidance, grant competitions, or the budget process.

Consider first the role of regulation. Shortly before leaving office, the Obama Administration promulgated a final rule to implement the Act's accountability requirements.⁹⁸ While Republicans in Congress used the Congressional Review Act to rescind the accountability rules early in the Trump Administration,⁹⁹ the rules are nonetheless illustrative of how an administration can use (or skirt) the Act's evidence provisions to further its preferred policies.

For example, when describing the steps that school districts must take to develop a comprehensive support and improvement plan for their lowest-performing schools, the rules provided a detailed list of types of interventions that the Obama Administration believed would satisfy the requirement for being evidence based¹⁰⁰—an illustration of the “sanctioning role” for evidence at the federal level.¹⁰¹ This list contained interventions that the Administration obviously believed were good choices—among them, “strategies designed to increase diversity by attracting and retaining students from varying socioeconomic, racial, and ethnic backgrounds” and “in the case of a public charter school . . . revoking or non-renewing the school's charter”¹⁰²—but from which the Trump Administration would likely wish to distance itself.¹⁰³

98. Elementary and Secondary Education Act of 1965, as Amended by the Every Student Succeeds Act—Accountability and State Plans, 81 Fed. Reg. 86,076 (Nov. 29, 2016) (disapproved by Act of Mar. 27, 2017, Pub. L. No. 115-13, 131 Stat. 77).

99. See Act of Mar. 27, 2017, Pub. L. No. 115-13, 131 Stat. 77; Emma Brown, Senate Scraps Obama Regulations on School Accountability, Wash. Post (Mar. 9, 2017), http://www.washingtonpost.com/local/education/senate-scraps-obama-regulations-on-school-accountability/2017/03/09/e9279932-04e5-11e7-b1e9-a05d3c21f7cf_story.html (on file with the *Columbia Law Review*) (discussing the party-line vote in Congress).

100. Elementary and Secondary Education Act of 1965, as Amended by the Every Student Succeeds Act—Accountability and State Plans, 81 Fed. Reg. at 86,079, 86,230.

101. See *supra* note 96 and accompanying text.

102. Elementary and Secondary Education Act of 1965, as Amended by the Every Student Succeeds Act—Accountability and State Plans, 81 Fed. Reg. at 86,230.

103. See, e.g., Emma Brown, Trump's Education Department Nixes Obama-Era Grant Program for School Diversity, Wash. Post (Mar. 29, 2017), <http://www.washingtonpost.com/news/education/wp/2017/03/29/trumps-education-department-nixes-obama-era-grant-program-for-school-diversity> (on file with the *Columbia Law Review*); Stephen Henderson, Betsy DeVos and the Twilight of American Public Education, Det. Free Press (Dec. 3, 2016), <http://www.freep.com/story/opinion/columnists/1948bama1948n-henderson/2016/12/03/betsy-devos-education-donald-trump/94728574/> [<http://perma.cc/BX7X-K2HE>] (describing how charter schools in Michigan, where DeVos has long championed them, are almost never closed for poor performance).

If the Trump Administration decides to reissue a new set of accountability rules within the limits of the Congressional Review Act,¹⁰⁴ presumably it would feel free to select other examples that fit its policy preferences instead. Nothing would require states, districts, and schools to select these choices, but it is reasonable to imagine that at least some will find it easier to use what the Administration is promoting.¹⁰⁵

Another way an administration can use the regulatory process to put its stamp on evidence is to clarify the definition of that term. For example, in 2013, as part of the Obama Administration's effort to promote evidence-based policymaking more generally,¹⁰⁶ the Department of Education revised its General Administrative Regulations to include for the first time a detailed definition of evidence that it could use in the competitive grant application process.¹⁰⁷ The Trump Administration recently issued a final rule making technical changes to these definitions to conform to the language in ESSA,¹⁰⁸ but one can imagine a subsequent revision that would be substantive—perhaps removing the requirement that to be considered “strong” or “moderate,” evidence from a study must include a sample that overlaps with the population or setting relevant to the grantee,¹⁰⁹ or perhaps deleting the definitions altogether and leaving entirely to administrators' discretion what would satisfy the statute's requirements.¹¹⁰

104. See 5 U.S.C. § 801(b)(2) (2012) (providing that a rule rescinded under that Act “may not be reissued in substantially the same form”).

105. Others, of course, may prefer to avoid any strategy that the Trump Administration promotes (just as in the Obama era, the reverse may have been true for a differently aligned set of political actors), but at least some will just select from the offered list. See, e.g., Coburn et al., *supra* note 19, at 78 (describing a study in which several districts chose programs from a list without “particularly support[ing] [those programs] solely as a way to maintain their federal funding”).

106. See *supra* note 6 and accompanying text.

107. Compare Direct Grant Programs and Definitions that Apply to Department Regulations, 78 Fed. Reg. 49,338, 49,355–56 (Aug. 13, 2013) (revising, *inter alia*, 34 C.F.R. § 77.1, the section on definitions, and including definitions of terms relevant to evidence), with 34 C.F.R. § 77.1 (2012) (including no definitions of terms relevant to evidence).

108. See Definitions and Selection Criteria that Apply to Direct Grant Programs, 82 Fed. Reg. 35,445, 35,445 (July 31, 2017) (to be codified at 34 C.F.R. pts. 75, 77) (explaining that “these regulations make only technical changes”).

109. For existing versions of this requirement, see *id.* at 35,449–50 (including a version of this requirement for the definitions of “moderate evidence” and “strong evidence”); Direct Grant Programs and Definitions that Apply to Department Regulations, 78 Fed. Reg. at 49,355–56 (including a version of this requirement for the definitions of “moderate evidence of effectiveness” and “strong evidence of effectiveness”).

110. The Department of Education has indicated that the General Administrative Regulations are under review for potential “repeal, replacement, or modification” under President Trump's Executive Order directing all federal agencies to identify regulations for one of those fates. See Exec. Order No. 13,777, 82 Fed. Reg. 12,285, 12,286 (Mar. 1, 2017) (ordering agencies to identify “regulations for repeal, replacement, or modification”); U.S. Dep't of Educ., Regulatory Reform Task Force Progress Report 2, 4 (2016), <http://www2.ed.gov/documents/press-releases/regulatory-reform-task-force-progress-report.pdf>

The absence of regulatory language on evidence offers another avenue to further an administration's policy goals. For example, in 2016, the Obama accountability rules placed an additional limit on how to establish that interventions in school improvement plans under ESSA would be evidence based: It would not be acceptable to pick an intervention that is justified only by a lower tier of evidence if an intervention justified by a higher tier of evidence would also be appropriate.¹¹¹ When Congress rejected the 2016 accountability rules,¹¹² this more stringent overlay on evidence for ESSA's school improvement plans became void. This absence further opens the door for the Trump Administration to push its own policy choices more generally, even if supported by weaker evidence.

Nonregulatory guidance provides one opportunity to do so. Several months before the Obama Administration issued the accountability regulations in November 2016, it issued nonregulatory guidance on using evidence.¹¹³ That guidance contained detailed recommendations for how states, districts, and schools should identify evidence at each of ESSA's four tiers.¹¹⁴ Most importantly, the guidance explains that, while the statutory definition of evidence-based requires only "at least one study,"¹¹⁵ stakeholders in fact "should consider the entire body of relevant evidence" and should not choose interventions that only one study supports if there are other equally strong studies that reach the opposite conclusion about the intervention.¹¹⁶

As of October 2017, that guidance document is still on the Trump Department of Education's website,¹¹⁷ but rescinding that guidance document would be one way for the Trump Administration to put its mark on the interventions it wants states, districts, and schools to select. The recommendation to consider "the entire body of relevant evidence,"¹¹⁸ for

[<http://perma.cc/8WYP-YL7N>] (noting that the Department of Education's central policy office "will facilitate discussions about Department-wide regulations," such as the General Administrative Regulations, to conform with this Executive Order).

111. Elementary and Secondary Education Act of 1965, as Amended by the Every Student Succeeds Act—Accountability and State Plans, 81 Fed. Reg. 86,076, 86,231 (Nov. 29, 2016) (disapproved by Act of Mar. 27, 2017, Pub. L. No. 115-13, 131 Stat. 77) (describing the revision to 34 C.F.R. § 200.21(d)(3)(iii)).

112. See *supra* notes 98–99 and accompanying text.

113. U.S. Dep't of Educ., Non-Regulatory Guidance: Using Evidence to Strengthen Education Investments (2016) [hereinafter U.S. Dep't of Educ., Using Evidence], <http://www2.ed.gov/policy/elsec/leg/essa/guidanceusesinvestment.pdf> [<http://perma.cc/A2T5-8GL5>].

114. *Id.* at 8–9.

115. *Id.* at 8 (internal quotation marks omitted) (paraphrasing 20 U.S.C. § 7801(21)(A)(i)(I)–(III) (Supp. 2015)).

116. *Id.*

117. See U.S. Dep't Educ., Using Evidence, *supra* note 113 (providing a permalink that captured the document as of October 20, 2017).

118. *Id.* at 8.

example, could be detrimental to the Administration's favored policy of school vouchers,¹¹⁹ in light of evidence on both sides of that issue.¹²⁰

The Trump Administration could also issue other more general non-regulatory guidance, sanctioning its preferred best practices,¹²¹ just as the Obama Administration did in a variety of contexts.¹²² The Trump Administration's version would no doubt draw on a different set of think tanks than the Obama Administration did and promote a different set of interventions.¹²³

In addition to promulgating new regulations and issuing nonregulatory guidance documents, the Department of Education can also push its policy preferences through the way it structures grant competitions.¹²⁴ In designing the program criteria and allocating points to different areas, the Department can weigh its desired policies more heavily than others.¹²⁵ The Department can also make choices about who should be peer reviewers for the competitions, a procedural decision that can end up having substantive effects.¹²⁶ Selecting and rejecting grant recipients provide ample opportunity to entrench an administration's policy and political priorities.¹²⁷

119. See, e.g., U.S. Dep't of Educ., Fiscal Year 2018 Budget Summary and Background Information 1–3 (2017), <http://www2.ed.gov/about/overview/budget/budget18/summary/18summary.pdf> [<http://perma.cc/L63T-TRYD>] (describing the goal of increasing school choice, including through private-school scholarships).

120. See *infra* note 155.

121. See *supra* note 96 and accompanying text.

122. See, e.g., Press Release, U.S. Dep't of Educ., Obama Administration Releases Resources for Schools, Colleges to Ensure Appropriate Use of School Resource Officers and Campus Police (Sept. 8, 2016), <http://www.ed.gov/news/press-releases/obama-administration-releases-resources-schools-colleges-ensure-appropriate-use-school-resource-officers-and-campus-police> [<http://perma.cc/6QSE-2KXH>]; Press Release, U.S. Dep't of Educ., U.S. Departments of Education and Justice Release School Discipline Guidance Package to Enhance School Climate and Improve School Discipline Policies/Practices (Jan. 8, 2014), <http://www.ed.gov/news/press-releases/us-departments-education-and-justice-release-school-discipline-guidance-package> [<http://perma.cc/9BSK-A8GQ>]; see also 20 U.S.C. § 7261(f) (Supp. 2015) (directing the Department of Education to “disseminate best practices” in education research).

123. Cf. Bruce Baker & Kevin G. Welner, Evidence and Rigor: Scrutinizing the Rhetorical Embrace of Evidence-Based Decision Making, 41 *Educ. Researcher* 98, 98 (2012) (critiquing the Obama Administration's Department of Education for relying on “speculative think-tank reports” instead of “high-quality research”); Goldhaber & Brewer, *supra* note 77, at 205–07 (discussing research conducted and disseminated by think tanks, which often have “a definitive ideological bent”).

124. See Pasachoff, *The President's Budget*, *supra* note 29, at 2270 (“[G]rant competition priorities can appear neutral while in fact privileging certain sets of applicants, whether those whose work is favored on substantive policy grounds or those who are politically important.”).

125. See *id.*

126. See, e.g., Hess & Little, *supra* note 9, at 7 (discussing the “seemingly political nature” of grant competition reviews).

127. See Pasachoff, *The President's Budget*, *supra* note 29, at 2256 (discussing political science literature illustrating the use of grants to support the President's interests). The Trump Administration appears to be using grant awards to reject at least some evidence-

Finally, the budget process provides another way for the federal government to promote its policy choices, notwithstanding ESSA's language on evidence. Each year, the President's budget will propose funding different programs and activities at different amounts, and different Presidents prioritize different things.¹²⁸ Congress, too, can decide whether and to what extent to fund the President's priorities.¹²⁹ A program that embeds evidence-based thinking may simply not be prioritized. At the same time, Congress can make clear, whether through formal appropriations riders¹³⁰ or informally in legislative history during the appropriations process, how it wants the Department to allocate funds.¹³¹ Nothing in ESSA's embrace of evidence would restrict the budget process from undercutting that embrace.

For all of these reasons, then, ESSA's focus on evidence, both on the face of the law and in its operation, does little to meaningfully limit the policy options available to the Department of Education and the states.

II. THE LIMITS OF THE RESEARCH

Even if ESSA's language were tightened, however, challenges would remain, for no provision requiring evidence-based decisionmaking is likely to dramatically improve education in America. The hope that evidence will drive such a transformation stems from "assum[ptions] that evidence is clear, unambiguous, and available; that decision makers use evidence in an instrumental fashion, weighing the merits of alternate courses of action and choosing the solutions that best fit the problem; and therefore that evidence leads directly to decisions."¹³² As this Part demonstrates, however, these assumptions turn out to be wrong, in light of limits in the state of the research, difficulties in conducting this kind of social science research in the first place, and challenges in implementing what research findings exist.

based programs in the social policy space. See Robert Gordon & Ron Haskins, *Trump Team Doesn't Understand Evidence-Based Policies Regarding Social Problems*, Hill (July 26, 2017), <http://thehill.com/blogs/pundits-blog/the-administration/343908-trump-team-doesnt-understand-evidence-based-policies> [<http://perma.cc/Q2EK-49U2>] (describing termination notices given by the Department of Health and Human Services to eighty-one evidence-based programs showing some success in reducing teen pregnancy).

128. See, e.g., Allen Schick, *The Federal Budget: Politics, Policy, Process* 101 (3d ed. 2007); Pasachoff, *The President's Budget*, *supra* note 29, at 2211–12.

129. Schick, *supra* note 128, at 109.

130. Jason A. MacDonald, *Limitation Riders and Congressional Influence over Bureaucratic Policy Decisions*, 104 *Am. Pol. Sci. Rev.* 766, 766–67 (2010) (describing limitation riders as "an effective tool for congressional influence over . . . bureaucratic policy making").

131. See, e.g., Schick, *supra* note 128, at 136 (discussing legislative history).

132. Coburn et al., *supra* note 19, at 69.

A. *The Small and Contested Research Base*

As two leading education funders and former federal officials wrote almost a decade into the No Child Left Behind era, “[G]iven the amount of paper and breath expended on the various pedagogical ‘wars,’ one might think that by now there would be more definitive knowledge and agreement on what to do.”¹³³

That having been said, there is no doubt that the quality and quantity of research in education have exploded over the course of the last two decades. In large part, this development was prompted by incentives in federal law. Reauthorizations of the Elementary and Secondary Education Act (ESEA) of 1965 and additional stand-alone legislation dating back to 1988 started to encourage a focus on finding and implementing effective programs.¹³⁴ This focus reached a high point in 2002, when Congress reauthorized the ESEA as NCLB, which required numerous interventions relying on “scientifically based research,”¹³⁵ and then subsequently passed the Education Sciences Reform Act, which expanded and institutionalized the federal infrastructure for funding and encouraging education research.¹³⁶

The first director of the Institute for Education Sciences, an entity created by the Education Sciences Reform Act, has described the state of knowledge at that time as extremely sparse.¹³⁷ When state education agencies

133. Frederic A. Mosher & Marshall S. Smith, *The Role of Research in Education Reform from the Perspective of Federal Policymakers and Foundation Grantmakers*, in *The Role of Research in Educational Improvement*, supra note 19, at 19, 32.

134. See, e.g., Benjamin Michael Superfine, *New Directions in School Funding and Governance: Moving from Politics to Evidence*, 98 Ky. L.J. 653, 678–80 (2010) (describing this history); see also Michael J. Feuer, Lisa Towne & Richard J. Shavelson, *Scientific Culture and Educational Research*, *Educ. Researcher*, Nov. 2002, at 4, 4–6 (comparing this history to other fields).

135. The term “scientifically based research” appeared 111 times in NCLB. See Kamina Aliya Pinder, *Using Federal Law to Prescribe Pedagogy: Lessons Learned from the Scientifically-Based Research Requirements of No Child Left Behind*, 6 Geo. J.L. & Pub. Pol’y 47, 59 (2008). NCLB’s definition of “scientific research” was controversial, as it seemed to limit acceptable forms of research to randomized control trials. See, e.g., David C. Berliner, *Educational Research: The Hardest Science of All*, *Educ. Researcher*, Nov. 2002, at 18, 18 (critiquing NCLB’s definition). ESSA’s expansion of the phrase “evidence-based” to include four tiers of research instead of just the top tier is in part a response to these critiques. See West, supra note 31 (explaining the difference between ESSA’s evidence requirements and NCLB’s “scientifically based research”).

136. See Superfine, supra note 134, at 692 (calling the Education Sciences Reform Act “a reasoned response to the inconsistent quality of educational research and the lack of a solid evidentiary base for making effective educational funding decisions,” while also noting its “fail[ure] to account for the complex and heavily contextualized nature of education”). For a history of the federal role in education research since 1867, see generally Andrew Rudalevige, *Structure and Science in Federal Education Research*, in *When Research Matters*, supra note 14, at 17.

137. See Christopher S. Elmendorf & Darien Shanske, *Solving “Problems No One Has Solved”: Courts, Causal Inference, and the Right to Education*, 2018 U. Ill. L. Rev. (forthcoming)

asked him what kinds of programs they should implement in order to meet NCLB's new requirements, he confesses that he would tell them that he didn't know of any decent research that would help them decide what to do.¹³⁸

"[H]ow far we have come since 2002," the immediate past director of the Institute recently observed.¹³⁹ "We actually know and use things from education research" developed over the last fifteen years, she explained, and educational systems have changed directions in response to research findings.¹⁴⁰

While the research base has expanded, however, it is still limited, and its expansion has not led to clear results or obvious solutions to the nation's pressing educational problems.¹⁴¹ For example, the What Works Clearinghouse—an initiative under the auspices of the Institute for Education Sciences that identifies helpful research studies¹⁴²—does not contain straightforward answers for education decisionmakers. A meta-analysis of the 10,000 studies reviewed by the Clearinghouse found that "only 29 different interventions showed significant effects—and the average effect was small."¹⁴³ Another analysis of the twenty-seven randomized control

(manuscript at 12–13), <http://papers.ssrn.com/abstract=2886754> (on file with the *Columbia Law Review*).

138. *Id.*

139. Ruth Curran Neild, *Federally-Supported Education Research Doesn't Need a Do-Over*, Brookings Inst. (Apr. 7, 2016), <http://www.brookings.edu/research/federally-supported-education-research-doesnt-need-a-do-over/> [<http://perma.cc/X7QF-HZ5M>].

140. *Id.*; see also Ben Levin, *Making Research Matter More*, *Educ. Pol'y Analysis Archives*, Oct. 17, 2004, at 1, 3 (describing how research influenced the adoption of early education and school reform strategies focused on building educator capacity); West, *supra* note 31 (describing the adoption of early-reading-instruction programs based on evidence). Despite critiques that NCLB's focus on experiments was too narrow a conception of what scientifically based research should be, see *supra* note 135, the 2002 laws led to an expansion in experiments in education. See, e.g., Alan Ginsburg & Marshall S. Smith, *Am. Enter. Inst., Do Randomized Controlled Trials Meet the "Gold Standard"? A Study of the Usefulness of RCTs in the What Works Clearinghouse* 6, 29 n.19 (2016), <http://www.aei.org/wp-content/uploads/2016/03/Do-randomized-controlled-trials-meet-the-gold-standard.pdf> [<http://perma.cc/3RP9-TXE2>] (describing the expansion in Randomized Control Trials (RCTs) between 2004 and 2014).

141. See, e.g., Elmendorf & Shanske, *supra* note 137 (manuscript at 4) ("[T]here is no social scientific (or political) consensus about what *changes* to the education system would most likely bring about substantial improvements in the adult outcomes of high-poverty, high-need student populations."); Levin, *supra* note 140, at 4 ("Much of education is concerned with producing significant and lasting change in how people think or behave, yet on the whole we do not yet know very much about how to do this, either in schools or in other settings."); Superfine, *supra* note 134, at 692 ("[W]hile our knowledge about educational reform is growing, it is still quite limited.").

142. What Works Clearinghouse, *Inst. of Educ. Scis.*, <http://ies.ed.gov/ncee/wwc/> [<http://perma.cc/23JC-4KMD>] (last visited Aug. 2, 2017).

143. Sarah D. Sparks, *How to Find Evidence-Based Fixes for Schools that Fall Behind*, *Educ. Wk.* (Sept. 27, 2016), <http://www.edweek.org/ew/articles/2016/09/28/how-to-find-evidence-based-fixes-for-schools.html> [<http://perma.cc/7ANA-A2WG>] [hereinafter Sparks, *Evidence-Based Fixes*] (last updated Sept. 28, 2016).

trials about mathematics curricula in the Clearinghouse concluded that “none of the [studies] provides sufficiently useful information for consumers wishing to make informed judgments about which mathematics curriculum to purchase” and suggested that the problems the reviewers identified likely extended both to studies of other types of curricula and to non-curricular interventions.¹⁴⁴ Some, in fact, have called it the Nothing Works Clearinghouse.¹⁴⁵

Even for programs with positive research findings, it has been hard to take them to scale in a statewide, systemic way.¹⁴⁶ As one education researcher has observed, discussing the disjunctive nature of individual studies, researchers “have been focusing on their parts of the elephant, and I’m not sure there would be a whole elephant if you brought them all together.”¹⁴⁷

Even positive research findings remain deeply contested.¹⁴⁸ For example, there is no agreement that money matters to educational outcomes,¹⁴⁹ that post-Katrina New Orleans schools are a success story,¹⁵⁰

144. Ginsburg & Smith, *supra* note 140, at 23–25. For other critiques of the What Works Clearinghouse, see, e.g., Sheri H. Ranis, *Blending Quality and Utility: Lessons Learned from the Education Research Debates*, in *Education Research on Trial*, *supra* note 2, at 125, 131–33; Alan H. Schoenfeld, *Instructional Research and the Improvement of Practice*, in *The Role of Research in Educational Improvement*, *supra* note 19, at 163, 184–85.

145. Rudalevige, *supra* note 136, at 36. Many researchers do not view this fact with despair but rather see it as an expected part of conducting high-quality research. See, e.g., Judith M. Gueron, *The Politics of Random Assignment: Implementing Studies and Affecting Policy*, in *Evidence Matters*, *supra* note 25, at 15, 40 (citing one of sociologist Peter Rossi’s laws, which provides “the better the study, the smaller the likely net impact”); Richard M. Ingersoll, *Researcher Meets the Policy Realm: A Personal Account*, in *When Research Matters*, *supra* note 14, at 113, 132 (citing another of Rossi’s laws, which says “the expected value for any measured effect of a social program is zero”). But the absence of ready answers complicates education decisionmakers’ ability to use interventions that are supported by evidence.

146. See, e.g., Elmendorf & Shanske, *supra* note 137 (manuscript at 13) (explaining that, despite promising interventions in small-scale pilots, “no state has achieved big, sustained improvements at scale”).

147. Sparks, *Evidence-Based Fixes*, *supra* note 143; see also John D. Bransford, Nancy J. Vye, Deborah J. Stipek, Louis M. Gomez & Diana Lam, *Equity, Excellence, Elephants, and Evidence*, in *The Role of Research in Educational Improvement*, *supra* note 19, at 1, 1, 3 [hereinafter Bransford et al., *Equity*] (citing John Godfrey Saxe, *The Blind Men and the Elephant*, in *Poetry of America* 151–52 (William James Linton ed., 1878)) (discussing the difficulty of metaphorically seeing the whole elephant in an educational system as decentralized as that of the United States).

148. See, e.g., Fusarelli, *supra* note 14, at 181 (discussing the ambiguity of social science and education research).

149. See Superfine, *supra* note 134, at 670 (describing the scholarly debate on this point).

150. See, e.g., Anya Kamenetz, *New Orleans Schools, 10 Years After Katrina: Beacon or Warning?*, NPR (Aug. 15, 2015), <http://www.npr.org/sections/ed/2015/08/15/431967706/new-orleans-schools-10-years-after-katrina-beacon-or-warning> [<http://perma.cc/3LZ8-F8Y3>] (“Critics of the reforms . . . look at the same data as supporters of the new system and draw wildly different conclusions.”).

that preschool programs,¹⁵¹ smaller class sizes,¹⁵² reading interventions,¹⁵³ standards- and testing-based accountability regimes,¹⁵⁴ or vouchers¹⁵⁵ improve student achievement. In a classic essay, *The Awful Reputation of Education Research*, education historian Carl Kaestle quoted Christopher Cross, a leading federal official for education research, as naming this conundrum “Cross’s corollary”: “[F]or every study in education research, there are an equal or greater number of opposing studies.”¹⁵⁶

Some suggest that in at least some of these cases, “real scholarly differences are at issue.”¹⁵⁷ For example, possibly well-intentioned methodological differences may result in significantly overblown conclusions.¹⁵⁸ Disagreements about what metrics are appropriate to use¹⁵⁹ and what constitutes high-quality research¹⁶⁰ explain another set of scholarly differences.

Beyond these relatively neutral explanations for differences in published findings is the possibility of ideological bias.¹⁶¹ The concern with

151. See Superfine, *supra* note 134, at 671 (describing the scholarly debate regarding preschool programs).

152. See Thomas D. Cook, Why Have Educational Evaluators Chosen Not to Do Randomized Experiments?, *Annals Am. Acad. Pol. & Soc. Sci.*, Sept. 2003, at 114, 138–39 (noting the varying results in analyses of class-size experiments).

153. See generally James S. Kim, Research and the Reading Wars, *in* *When Research Matters*, *supra* note 14, at 89 (tracing the historical debate over phonics in reading instruction).

154. Superfine, *supra* note 134, at 682–84 (discussing the mixed results of these regimes).

155. See Cook, *supra* note 152, at 133, 138 (summarizing conflicting studies on voucher programs); Jeffrey R. Henig, The Evolving Relationship Between Researchers and Public Policy, *in* *When Research Matters*, *supra* note 14, at 41, 43–44 [hereinafter Henig, *Evolving Relationship*] (describing this scholarly dispute).

156. Carl Kaestle, *The Awful Reputation of Education Research*, *Educ. Researcher*, Jan.–Feb. 1993, at 23, 29 (internal quotation marks omitted).

157. Cook, *supra* note 152, at 138.

158. See, e.g., Manna & Petrilli, *supra* note 14, at 88 (“Within any study—even those meeting the highest design standards and that pass peer-review—researchers make judgment calls that some other credible researcher would see as flawed.”); Adrian Simpson, The Misdirection of Public Policy: Comparing and Combining Standardised Effect Sizes, 32 *J. Educ. Pol’y* 460, 463 (2017) (critiquing standardized effect sizes as a “policy tool for directing whole educational areas” and criticizing scholarly framing of “areas where it is easier to make what may be educationally unimportant differences stand out through methodological choices”).

159. See, e.g., Ingersoll, *supra* note 145, at 118–19 (observing that even researchers studying the same phenomenon may disagree about which measures are appropriate to evaluate); see also Jason Russell, Opinion, Why Do School Choice Critics Elevate Test Scores Over Choice?, *Wash. Examiner* (May 4, 2017), <http://www.washingtonexaminer.com/why-do-school-choice-critics-elevate-test-scores-over-choice/article/2622124> [http://perma.cc/P345-BFLC] (arguing it is wrong to conclude that vouchers do not work because of a study showing decreased math scores from voucher use if the appropriate metric is parent satisfaction).

160. See, e.g., Fusarelli, *supra* note 14, at 186 (describing “serious and fundamental disagreement about what constitutes valid, reliable research” in the field of education).

161. See, e.g., Hess, Conclusion, *supra* note 25, at 245 (“[W]hen research gets caught up in larger political debates and is wielded by interested parties, it can become more difficult

ideological bias is amplified when researchers agree on the importance of an educational input but draw sharply divergent conclusions along ideological lines about what policy interventions would make best use of that input.¹⁶²

As the editors of a volume on “[k]ey issues and challenges” for evidence-based policy have explained, “To the extent that research findings are widely used as weapons in strongly emotive debates, it may be only a short step to accusations that most research on these matters is biased and lacks objectivity.”¹⁶³ This state of play in the research limits the confidence that education decisionmakers can have that research findings represent a neutral determination of “what works.”

Critiques on political grounds of the Institute for Education Sciences and its predecessor, the Office of Educational Research and Improvement, further complicate matters. Various charges over the decades have been that the federal research entity then in existence has “overly politicized . . . its reviews of programs,”¹⁶⁴ too narrowly circumscribed the types of research it deems acceptable,¹⁶⁵ or even suppressed research findings for political reasons.¹⁶⁶ The Department of Education office in charge of

for scholars to argue about technical considerations, such as sample size or measurement error, as researchers rather than as partisans.”); Carol H. Weiss, *The Politicization of Evaluation Research*, 26 *J. Soc. Issues* 57, 59–60 (1970) [hereinafter Weiss, *Politicization*] (noting that methodological disagreements may “derive less from methodology than from ideology”).

162. See, e.g., Ingersoll, *supra* note 145, at 122–23 (explaining that political liberals and conservatives agree that teachers matter to student outcomes but fiercely disagree about the extent to which regulations or markets will better supply high-quality teachers); Weiss, *Politicization*, *supra* note 161, at 61–62 (noting that there is often “discontinuity between the study and recommendations of a course of action” and that, “in many cases, the data do not provide even a jumping-off point” for what the recommendations should be).

163. Brian W. Head, *Reconsidering Evidence-Based Policy: Key Issues and Challenges*, 29 *Pol’y & Soc’y* 77, 77, 81 (2010). The metaphor of research being used as a partisan weapon is common. See, e.g., Pamela Barnhouse Walters & Annette Lareau, *Education Research that Matters: Influence, Scientific Rigor, and Policymaking*, *in* *Education Research on Trial*, *supra* note 2, at 197, 214 (noting that research is used more as an “arsenal” by those with policy goals already in place); Kenneth K. Wong, *Considering the Politics in the Research Policymaking Nexus*, *in* *When Research Matters*, *supra* note 14, at 219, 235 (“As Brookings Institution economist Henry Aaron points out, ‘[P]eople wield their social science research studies like short swords and shields in the ideological wars.’” (quoting Andrew Rich, *Think Tanks, Public Policy, and the Politics of Expertise* 215 (2004))).

164. Superfine, *supra* note 134, at 688; see also Maris A. Vinovskis, *A History of Efforts to Improve the Quality of Federal Education Research: From Gardner’s Task Force to the Institute of Education Sciences*, *in* *Education Research on Trial*, *supra* note 2, at 51, 56, 59 [hereinafter Vinovskis, *History of Efforts*] (describing varying complaints during different eras from both sides of the aisle about the political slant of the agency’s research staff).

165. Pamela Barnhouse Walters, *The Politics of Science: Battles for Scientific Authority in the Field of Education Research*, *in* *Education Research on Trial*, *supra* note 2, at 17, 40 (discussing the American Educational Research Association’s public statement “express[ing] dismay” at the agency’s limited view of appropriate scientific research in education); Wong, *supra* note 163, at 225 (“Partisan shift tends to destabilize appropriations for research because the policy priorities are likely to change.”).

166. Kaestle, *supra* note 156, at 30.

an NCLB-era evidence-based reading program faced similar allegations, culminating in a critical report by the Department's Inspector General Office.¹⁶⁷ Similar critiques exist of state-level government entities.¹⁶⁸

During the era that No Child Left Behind and the Education Sciences Reform Act were being debated, the National Research Council published a major report on scientific research in education. After considering the appropriate extent of “[p]olitical insulation” of any federal educational research agency, the authors concluded that “[i]t would be simply incompatible with the American tradition of democratic governance to exclude political and social influences from decisions about research priorities.”¹⁶⁹ There is a real logic to this argument.¹⁷⁰ But it also comes with a cost. When the government promotes evidence-based decisionmaking but does not seem neutral in so doing, it can be hard to see evidence itself as anything other than political.

B. *The Complexities of Education Research*

Another set of issues further limits the likelihood that evidence-based policymaking will ultimately transform education in America: the complexities inherent in conducting this kind of social science research.

In a twist on the familiar contrast between the hard sciences (“[p]hysics, chemistry, geology, and so on”) and the soft sciences (“the social sciences in general and education in particular”), one education scholar instead contrasts the “[h]ard-to-do science[s]” (like social science in general and education in particular) and the “[e]asy-to-do science[s]” (like physics, chemistry, and the rest).¹⁷¹ “We do our science under conditions that physical scientists find intolerable,” wrote this education scholar, referring to messy, human, context-bound interactions.¹⁷² Or, as another

167. Pinder, *supra* note 135, at 62, 75; see also Robert B. Schwartz & Susan M. Kardos, *Research-Based Evidence and State Policy*, in *The Role of Research in Educational Improvement*, *supra* note 19, at 47, 53 (“[S]keptics [of evidence-based policymaking] view the movement, especially as it applies to the implementation of the federal government’s reading policy, as one more example of ideology masquerading as science.”).

168. See, e.g., Fusarelli, *supra* note 14, at 183 (describing critiques of former Texas Commissioners of Education who owned companies promoting “‘research-based’ programs” that also appeared on an approved list to qualify for state funding).

169. Feuer et al., *supra* note 134, at 10.

170. Cf. Pasachoff, *The President’s Budget*, *supra* note 29, at 2269–70 (noting that funding decisions are “value-laden decisions” that have a proper place in government priority setting).

171. Berliner, *supra* note 135, at 18.

172. *Id.*; see also Levin, *supra* note 140, at 4 (“[K]nowledge about human behavior is in principle different from knowledge of the inanimate world . . .”). Of course, even knowledge about the inanimate world has research challenges. See, e.g., Wendy E. Wagner, *The Science Charade in Toxic Risk Regulation*, 95 *Colum. L. Rev.* 1613, 1619–22 (1995) (discussing “trans-science” questions that can be hidden, intentionally or unintentionally, during the research process). These same challenges exist in education research. See, e.g.,

tongue-in-cheek reflection on these challenges suggests, all social science knowledge can be generally summarized as follows: “(1) Some do, some don’t. (2) The differences aren’t very great. (3) It’s more complicated than that.”¹⁷³

One challenge to education research is that it involves “complex social systems such as classrooms and schools”¹⁷⁴ with “humans . . . embedded in complex and changing networks of social interaction.”¹⁷⁵ Studies of interventions that seem to work in one context can be hard to replicate in a wildly different context, given the number of often indeterminable and likely unmeasurable variables potentially at issue.¹⁷⁶ It is also hard to make sure that the control group in an education experiment does not, in fact, receive the experimental treatment, since members across school communities often “have social and professional relations” that expose the control group to the experimental conditions.¹⁷⁷

In part, the challenge relates to our highly decentralized educational system. Almost 14,000 school districts have different needs, populations, curricula, instructional practices, and governance structures.¹⁷⁸ But the challenge is not entirely due to our federalist design, as human and social complexities exist even within any given locale; in education, “all else is too rarely equal to make ready claims of causality.”¹⁷⁹

Simpson, *supra* note 158, at 463 (noting researchers’ ability to “legitimately directly manipulate effect size when they are looking to increase their chance of detecting a difference”).

173. Edward R. Tufte, *Beautiful Evidence* 138 (2006).

174. Ginsburg & Smith, *supra* note 140, at 6.

175. Berliner, *supra* note 135, at 19; see also Feuer et al., *supra* note 134, at 7 (discussing features of education that complicate its study, including the volition and diversity of people and the variability of curriculum, instruction, and governance); Superfine, *supra* note 134, at 690–92 (discussing the “complex and heavily contextualized nature of education”).

176. Berliner, *supra* note 135, at 19 (describing “the ordinary events of life,” from “a sick child, a messy divorce, [or] a passionate love affair” to “a new principal, a new child in the classroom, [or] rain that keeps the children from a recess outside the school building,” that “limit[] the generalizability of educational research findings”); see also Ginsburg & Smith, *supra* note 140, at 1 (noting that even randomized controlled trials with internal validity do not necessarily have external validity); Lawrence W. Sherman, *Misleading Evidence and Evidence-Led Policy: Making Social Science More Experimental*, *Annals Am. Acad. Pol. & Soc. Sci.*, Sept. 2003, at 6, 9 (explaining that “threats to *internal validity*” concern “the conclusions drawn within the sample,” while threats to “*external validity*” concern “how far the conclusions may be generalized to other populations”).

177. Annette Lareau, *Narrow Questions, Narrow Answers: The Limited Value of Randomized Control Trials for Education Research*, in *Education Research on Trial*, *supra* note 2, at 145, 152–53.

178. Nat’l Ctr. for Educ. Statistics, U.S. Dep’t of Educ., *Digest of Education Statistics: Table 214.10. Number of Public School District and Public and Private Elementary and Secondary Schools: Selected Years, 1869–70 Through 2014–2015* (2016), http://nces.ed.gov/programs/digest/d16/tables/dt16_214.10.asp [<http://perma.cc/N3B5-JTZK>]; see also Bransford et al., *Equity*, *supra* note 147, at 3–4 (discussing the many institutional players and variations in American education); Fusarelli, *supra* note 14, at 190 (discussing skepticism that research conducted in one school in one state will translate to a very different context).

179. Rudalevige, *supra* note 136, at 18.

Another challenge in education research is that teaching involves multiple sets of interactions moving back and forth in multiple ways over multiple time periods. For example:

Any teaching behavior interacts with a number of student characteristics, including IQ, socioeconomic status, motivation to learn, and a host of other factors. Simultaneously, student behavior is interacting with teacher characteristics, such as the teacher's training in the subject taught, conceptions of learning, beliefs about assessment, and even the teacher's personal happiness with life. But it doesn't end there because other variables interact with those just mentioned—the curriculum materials, the socioeconomic status of the community, peer effects in the school, youth employment in the area, and so forth. Moreover, we are not even sure in which directions the influences work, and many surely are reciprocal.¹⁸⁰

It is therefore difficult to isolate connections that might matter to a study's conclusions.

A further challenge with conducting education research emerges from the difficulty of getting education decisionmakers to keep an intervention in place long enough to assess its value (even given the difficulties with assessment just discussed). Some interventions require long-term study,¹⁸¹ but decisionmakers often have short-term political needs that lead them to change a program before it has been fully implemented or assessed¹⁸² or to “insist that measurable results be available in a short time-frame.”¹⁸³

These challenges do not doom education (or broader social science) research, of course, but they do make it harder to identify “what works” in any straightforward sense.¹⁸⁴

180. Berliner, *supra* note 135, at 19; see also Cohen et al., *supra* note 25, at 86–87 (discussing the difficulties of “observation and measurement of complex social and intellectual processes” involved in teaching).

181. See, e.g., Gueron, *supra* note 145, at 39 (“The life cycle of a major experiment or evaluation is often five or more years.”); Vinovskis, *History of Efforts*, *supra* note 164, at 54 (describing an ideal “systematic five-stage strategy for education research and development that would require 10–12 years to complete”).

182. See, e.g., Head, *supra* note 163, at 84 (“[G]overnments have a propensity to change a program before outcomes have been assessed, so that any evaluation would thus be measuring moving targets with variable criteria of success.”); Lareau, *supra* note 177, at 152 (identifying “principals’ flagging interest in participating in a study” as part of what brings “daunting, and arguably insurmountable, challenges” to “randomized controlled trials that focus on longitudinal change in schools”).

183. Head, *supra* note 163, at 84.

184. For selected responses to this research challenge, see, e.g., Cook, *supra* note 152, at 131–32 (arguing that “the complexity and heterogeneity of schools leads to the need for larger school sample sizes and the need to anticipate and measure specific sources of variation to reduce their unwanted influence through statistical control”); Lareau, *supra* note 177, at 146 (arguing that a broader set of research methods than the randomized control trial—“qualitative methods, including participant observation and in-depth interviews”—is

Recent history also challenges the idea that educational research can, with enough dogged effort, identify solutions to social problems. In considering the utility of evidence-based policies in education and other social science fields, several scholars call for humility in the face of now-discredited but once widely held beliefs (about men and women, about race-based differences, about how children learn) that were based on research evidence.¹⁸⁵ Recognizing that “[r]esearch has been used to support positions that were later shown to be wrong or, even worse, are now considered morally repugnant, such as the supposed inferiority of some groups of people,” suggests caution in assuming that the state of knowledge in the field provides scientifically valid answers simply because it is based on research evidence.¹⁸⁶

C. *The Research-to-Practice Dilemma*

Where high-quality, nonpartisan research findings exist, another challenge remains in translating these findings from research to practice.¹⁸⁷ Many implementation difficulties exist.¹⁸⁸

The rise of education research has been accompanied by a rise in practitioners’ awareness and use of this research.¹⁸⁹ At the same time,

necessary in this context); Weiss, *Politicization*, supra note 161, at 62 (suggesting that instead of “all-or-nothing, go/no-go conclusions,” researchers should attend to “the effectiveness of variant conditions within programs . . . and begin to explain which elements and sub-elements are associated with more or less success”); Carol H. Weiss, *What to Do Until the Random Assigner Comes*, in *Evidence Matters*, supra note 25, at 198, 220–22 (discussing the value of “theory-based evaluation” and “ruling out” to evaluate “the sprawling changeable world of community programs”).

185. See Ellen Condliffe Lagemann, *An Elusive Science: The Troubling History of Education Research* 246 (2000) (“With the benefit of hindsight, one can see the limitations of beliefs that once seemed indisputably true.”); Berliner, supra note 135, at 20 (describing “the short half-life of [education research] findings” in light of “changes in the social environment that invalidate the research or render it irrelevant”); cf. Cecelia Klingele, *The Promises and Perils of Evidence-Based Corrections*, 91 *Notre Dame L. Rev.* 537, 575 (2016) (urging a similar humility in the context of correctional policy).

186. Levin, supra note 140, at 3.

187. See Lagemann, supra note 185, at 239–41 (explaining that the lack of a centralized educational research community or organization has made it difficult to coordinate a research agenda that might aid policymakers).

188. Cf. Jeffrey L. Pressman & Aaron Wildavsky, *Implementation: How Great Expectations in Washington Are Dashed in Oakland; Or, Why It’s Amazing that Federal Programs Work at All, This Being a Saga of the Economic Development Administration as Told by Two Sympathetic Observers Who Seek to Build Morals on a Foundation of Ruined Hope* 93–94 (3d ed. 1984) (describing the difficulty of any program implementation given “the number of steps involved, [and] the number of participants whose preferences have to be taken into account, the number of separate decisions that are part of what we think of as a single one”).

189. Neild, supra note 139 (reporting that, among staff in the nation’s thirty-two largest school districts, use of education research is high and skepticism of research is low). But see Fusarelli, supra note 14, at 179 (“[S]chool leaders are more likely to cite general research traditions or concepts such as brain research or emotional intelligence rather than specific studies.”).

“concerns remain about the capacities and dispositions of our governmental institutions to effectively interpret scientific evidence to the extent that it is present.”¹⁹⁰ These concerns may exacerbate geographical disparities, as smaller, rural, or less wealthy districts and states may not have as much access to research findings as do larger urban or wealthier suburban ones (let alone the ability to interpret them).¹⁹¹

Entities like the research–practice partnerships funded by the Institute for Education Sciences¹⁹² and other research intermediaries¹⁹³ can help connect education decisionmakers to studies. But this process is not as straightforward as it sounds, for reasons both technical (as it is unclear how well such partnerships work)¹⁹⁴ and political (as many intermediaries are ideologically driven).¹⁹⁵

Academics are not immune from these problems. As cuts to university funding have driven researchers to find other sources of research support, foundations and business philanthropy have stepped in; but such

190. Superfine, *supra* note 134, at 692; see also Carrie L. Conaway, *The Problem with Briefs*, in *Brief*, 8 *Educ. Fin. & Pol’y* 287, 293 (2013) (suggesting that the most common complaint about academic writing is that it is difficult for non-academics to understand); Fusarelli, *supra* note 14, at 186 (discussing the difficulty of keeping up with huge quantities of education research published each year even for “fulltime education researchers . . . let alone school leaders busy managing the increasingly complex daily operations of schools”).

191. See Coburn et al., *supra* note 19, at 70 (discussing the wide variety in districts’ capacity to access and consume research); Conaway, *supra* note 190, at 296 (noting that “the vast majority” of state education agencies do not have a research director); Sparks, *Evidence-Based Fixes*, *supra* note 143 (noting the importance of school context in education policy and asserting that ESSA’s flexibility “runs the risk of putting smaller or more rural districts at a disadvantage”).

192. See Neild, *supra* note 139 (describing recent growth in such partnerships); see also What Works Clearinghouse, *Practice Guides*, Inst. of Educ. Scis., <http://ies.ed.gov/ncee/wwc/PracticeGuides> [<http://perma.cc/7GHX-G8XB>] (last visited Aug. 2, 2017) (linking to summaries of education research in particular areas of school practice).

193. Conaway, *supra* note 190, at 297 (highlighting the importance of “independent research intermediaries”); see also Ctr. for Research & Reform in Educ., *Evidence for ESSA*, <http://www.evidencefoessa.org/> [<http://perma.cc/CX52-ME8E>] (last visited Aug. 2, 2017) (providing descriptions of math and reading programs “that meet ESSA evidence standards”).

194. See, e.g., Meredith I. Honig, Nitya Venkateswaran & Patricia McNeil, *Research Use as Learning: The Case of Fundamental Change in School District Central Offices*, 54 *Am. Educ. Res. J.* (forthcoming 2017) (manuscript at 29), <http://journals.sagepub.com/doi/abs/10.3102/0002831217712466> (on file with the *Columbia Law Review*) (reporting the “limited effect” of “research-practice intermediaries”).

195. Gary Anderson, Pedro de la Cruz & Andrea López, *New Governance and New Knowledge Brokers: Think Tanks and Universities as Boundary Organizations*, 92 *Peabody J. Educ.* 4, 5 (2017) (describing the skillful way that think tanks have successfully promoted ideological views through the dissemination of research); Hess, *Conclusion*, *supra* note 25, at 247 (describing “membership groups” with “a natural interest in promoting research findings which align with the interests of their members and their existing policy agendas” and “mission-driven or ideological organizations” as two large subsets of research intermediaries).

sponsorship again raises questions about the neutrality of the research, because these institutions are sometimes affiliated with clear advocacy positions.¹⁹⁶ At the same time, when academics engage on their own with education decisionmakers in the hope of making their research relevant,¹⁹⁷ the potential emerges for policy “capture” that negatively affects the “scope of research projects.”¹⁹⁸ From another direction, academics unaffiliated with any institution other than their home institution have sometimes been accused of conducting ideologically driven research anyway.¹⁹⁹ Academics have also found their work captured without their knowing assistance.²⁰⁰

Other potential mediating institutions fare no better in the quest for neutrality. For example, the media sometimes report what turns out to be partisan think-tank work uncritically, strive for balance in presenting two sides of a story when facts lie only on one side, or hype researchers’ divergent views to drum up readership.²⁰¹ For-profit companies have their own incentives and may produce the bare minimum of evidence required to sell their products, providing research that is largely spin.²⁰² And, as already indicated, it is simply not plausible to believe that the

196. Anderson et al., *supra* note 195, at 6, 11–12 (noting that when university researchers rely on sponsored funding, it calls into question whether their research is truly independent); Hess, Conclusion, *supra* note 25, at 248 (noting the financial and reputational incentives for researchers to “depict [their] work in ways that the [supporting] organizations will find congenial and to remain quiescent if they stretch the findings or recommendations in the course of their efforts”).

197. See, e.g., Conaway, *supra* note 190, at 290 (arguing for this kind of collaboration); Hess, Conclusion, *supra* note 25, at 252 (discussing the value of building relationships between education researchers and the subjects of their studies).

198. Head, *supra* note 163, at 86 (discussing the danger of such “capture” when researchers “engage closely with policy bureaucrats”); Henig, *Evolving Relationship*, *supra* note 155, at 62 (questioning “how far down the path of relevance researchers can travel without something of value being put at stake”); Hess, Conclusion, *supra* note 25, at 252–53 (discussing the challenges within researcher–practitioner relationships).

199. See Anderson et al., *supra* note 195, at 7 (tracing the development of conservative think tanks to the claim that “liberal academics had some direct and indirect influence on social policy”).

200. See, e.g., Ingersoll, *supra* note 145, at 128–29 (discussing this experience).

201. See, e.g., Jeffrey R. Henig, *Spin Cycle: How Research Is Used in Policy Debates: The Case of Charter Schools* 177–216 (2008) (describing such challenges for both old and new media); Carol H. Weiss & Eleanor Singer, *Reporting of Social Science in the National Media* 129–40 (1988) (identifying key complaints about social science reporting, including “oversimplification,” “undue closure and certainty,” and “inadequate scrutiny of the quality of social science studies”).

202. See, e.g., Fusarelli, *supra* note 14, at 182–83 (noting skepticism about the relationship between governments and for-profit “‘research-based’ programs”); Alis Oancea & Richard Pring, *The Importance of Being Thorough: On Systematic Accumulations of ‘What Works’ in Education Research*, in *Evidence-Based Education Policy* 11, 16 & n.9 (David Bridges, Paul Smeyers & Richard Smith eds., 2009) (describing the possibility of vendors gaming “‘research-based’” requirements by pointing to a self-appointed intermediary organization designed to provide speedy “certification” that “‘products are backed by valid research’” (quoting the intermediary’s now-defunct website)).

federal government is an unbiased institution that simply collects and presents research findings.²⁰³ It was not plausible even during the wonky, evidence-championing Obama Administration,²⁰⁴ and it is surely not plausible during the “alternative-fact”-embracing Trump Administration.²⁰⁵

A further challenge for connecting research to practice is the tension between what social science research can offer (“tentative and contextual” assessments of “the chance that something might happen given certain other conditions”) and what decisionmakers often want (“certainty” to inform decisions that have to be made now and the hope that policy problems can simply be solved).²⁰⁶ This tension can result in a number of difficulties. On the one hand, decisionmakers might simply decide to ignore research as not offering the answers they need.²⁰⁷ On the other hand, researchers or a research mediator might “oversell[] the present state of knowledge”²⁰⁸ because they “want their message to be clear and predominant (even when the research base is fuzzy).”²⁰⁹

Either of these paths presents challenges for the future influence of research findings on policy choices, because policy choices tend to be sticky.²¹⁰ Unless decisionmakers have a political need to make a change,²¹¹ it can be politically and bureaucratically difficult to move in a different

203. See *supra* notes 164–169 and accompanying text.

204. Baker & Welner, *supra* note 123, at 100 (critiquing Obama’s Department of Education for “rel[ying] overwhelmingly on work that is not peer reviewed, most of which is neither credible nor rigorous”).

205. Mahita Gajanan, Kellyanne Conway Defends White House’s Falsehoods as ‘Alternative Facts,’ *Time* (Jan. 22, 2017), <http://time.com/4642689/kellyanne-conway-sean-spicer-donald-trump-alternative-facts/> [<http://perma.cc/7QKW-BW25>].

206. Levin, *supra* note 140, at 4; see also Feuer et al., *supra* note 134, at 6, 9 (noting that, despite a “commonly heard lament . . . posed as a biting rhetorical question: When will education produce the equivalent of a Salk vaccine?”, “there is [no] simple panacea for ills of schools just waiting to be discovered”).

207. See, e.g., Fusarelli, *supra* note 14, at 181 (“[T]he ‘it depends’ response of researchers to many issues tends to freeze out researchers from having a significant impact on decisionmaking.”); Henig, *Evolving Relationship*, *supra* note 155, at 47 (suggesting that one practitioner response to unclear research is to say “a pox on all your houses”).

208. Klingele, *supra* note 185, at 576; see also D.C. Phillips, *A Quixotic Quest? Philosophical Issues in Assessing the Quality of Education Research*, in *Education Research on Trial*, *supra* note 2, at 163, 172 (noting Harold Larrabee’s 1964 lament that although “all statements . . . that something is reliably known should, strictly speaking, be made only with extensive qualifications . . . [,] [t]o save time, breath, and inked paper, we are likely to go right on with our broad, sweeping, abstract generalizations about what we claim to know” (alteration omitted)).

209. Karen Seashore Louis, *Politics, Advocacy, and Research: What Have We Learned and What Remains?*, 92 *Peabody J. Educ.* 141, 142 (2017).

210. See Goldhaber & Brewer, *supra* note 77, at 200 (providing an example of a policy still in place after more than a decade despite a negative evaluation); Wong, *supra* note 163, at 225 (discussing the “inertia of the status quo”).

211. See *supra* notes 182–183 and accompanying text.

direction.²¹² New findings may emerge that indicate that decisionmakers should change tack. But if decisionmakers have given up on research as unhelpful, or have relied on research to provide the final “answer” of what they should do, further developments in the research may not result in policy changes.²¹³

The difficulties with research-to-practice are hard enough for education decisionmakers, but additional difficulties emerge with education’s “street-level bureaucrats”—teachers.²¹⁴ Teachers may unintentionally be part of unfaithful implementation of a policy because of complexities that emerge during the school day and year or other institutional barriers.²¹⁵ Or they may intentionally push back subtly and unobservably out of frustration with a policy that conflicts with what they know from their own experience to be true about teaching and learning.²¹⁶ Or they may incorporate what they believe to be an evidence-based practice in their teaching that later studies have superseded, or that was never a proven strategy at all, and may be reluctant to give it up.²¹⁷

There is a school of research that respects the importance of professional knowledge rather than simply the findings of experiments.²¹⁸ There is also a school of research that says even experiments should incorporate

212. Levin, *supra* note 140, at 4 (noting that “[p]ractitioners may be deeply enmeshed in practices and beliefs that are highly resistant to change”).

213. See Sherman, *supra* note 176, at 17 (“Many consumers may treat [systematic reviews of studies] as final conclusions . . .”).

214. See Michael Lipsky, *Street-Level Bureaucracy: Dilemmas of the Individual in Public Services* 3 (2010) (“[T]he individual decisions of these workers become, or add up to, agency policy.”).

215. Ginsburg & Smith, *supra* note 140, at 11 (discussing “implementation-fidelity” problems with the mathematics curricula studies they reviewed); see also Lareau, *supra* note 177, at 153–57 (discussing the complexity of policy implementation because of social and institutional factors); Mosher & Smith, *supra* note 133, at 38 (stating that if policy reforms “require big changes, and/or the knowledge, resources, and technology to support the required changes are in short supply or absent, then the odds [of success] go way down”).

216. See, e.g., Larry Cuban, *Inside the Black Box of Classroom Practice: Change Without Reform in American Education* 161–63 (2013) (describing forms of and reasons for “[a]ctive or passive teacher resistance” to reform initiatives); Jack Schneider, *From the Ivory Tower to the Schoolhouse: How Scholarship Becomes Common Knowledge in Education* 186–87 (2014) (explaining that if research does not “possess a sense of philosophical compatibility” with teachers’ beliefs, “even if the idea is ostensibly ‘proven,’ it stands little chance of survival in classrooms”).

217. See Fusarelli, *supra* note 14, at 185–93 (discussing educators’ decisionmaking processes).

218. Andy Hargreaves & Corrie Stone-Johnson, *Evidence-Informed Change and the Practice of Teaching*, in *The Role of Research in Educational Improvement*, *supra* note 19, at 89, 89–90 (noting that one of the stronger approaches to evidence-informed improvement lies in professional learning communities, which use both research-based and practically grounded evidence); Head, *supra* note 163, at 83 (describing the value of professional and technical knowledge on the ground).

the realities of teacher discretion into their design.²¹⁹ It is key to study how schools implement programs, not just to evaluate what comes out of the black box of classroom practice, to determine whether a particular intervention “worked” or not.²²⁰ The bottom line, though, is that the volition of teachers is an important part of why implementing evidence-based policies in schools is not as easy as just a direction from the top.

The research-to-practice problem thus joins the research-difficulty problem and the state-of-knowledge problem in underscoring the improbability that calls to rely on evidence will cure what ails education in America.

III. THE LIMITS OF TECHNOCRACY

A final reason why the calls for evidence-based decisionmaking in ESSA are not going to transform education in America is that citizens are deeply divided about the underlying purpose of education (just as the country is divided about so many other social policies). As one set of scholars has written, “It is noteworthy that there is clarity about what defines a successful airplane but a lack of consensus on what defines a successful school and how we measure successes.”²²¹ These are questions about values that evidence cannot answer.²²²

219. Cohen et al., *supra* note 25, at 107–15 (describing research design to account for teachers as “active agents of instruction”).

220. See, e.g., Dean L. Fixsen, Sandra F. Naoom, Karen A. Blase, Robert M. Friedman & Frances Wallace, *Implementation Research: A Synthesis of the Literature 74–75* (2005), <http://nirn.fpg.unc.edu/sites/nirn.fpg.unc.edu/files/resources/NIRN-MonographFull-01-2005.pdf> [<http://perma.cc/8ZGB-GLYQ>]; Conaway, *supra* note 190, at 294 (calling for more studies on implementation). For discussion of the related field of “improvement science,” see generally Anthony S. Bryk, Louis M. Gomez, Alicia Grunow & Paul G. LeMahieu, *Learning to Improve: How America’s Schools Can Get Better at Getting Better* (2015); Sarah D. Sparks, ‘Improvement Science’ Seen as Emerging Tool in K–12 Sphere, *Educ. Wk.* (Oct. 1, 2013), <http://www.edweek.org/ew/articles/2013/10/02/06improvement.h33.html> [<http://perma.cc/RET9-6GCG>].

221. Louis M. Gomez, Janet A. Weiss, Deborah Stipek & John D. Bransford, *Toward a Deeper Understanding of the Educational Elephant*, in *The Role of Research in Educational Improvement*, *supra* note 19, at 209, 211–12 (describing “the lack of agreement about the goals of educational practice and policy,” even about something as seemingly straightforward as “ambitious instruction,” because of “other important goals of schools—such as maintaining safety, supporting mental health, promoting moral and social development, preparing students for the world of work”); see also Jennifer L. Hochschild & Nathan Scovronick, *The American Dream and the Public Schools 12–17* (2003) (describing conflicts over educational policies as rooted in the tension among three core American values: “[t]he [s]uccess of [i]ndividuals,” “[t]he [c]ollective [g]ood,” and “[t]he [w]elfare of [g]roups”); Levin, *supra* note 140, at 2 (“People may agree on educational goals only at the most general level, with many conflicts not only about goals but about the best means of carrying them out.”).

222. See, e.g., Trisha Greenhalgh & Jill Russell, *Evidence-Based Policymaking: A Critique*, 52 *Persp. Biology & Med.* 304, 310 (2009) (“[A]n answer to the question ‘What should we do’ will never be plucked cleanly from massed files of scientific evidence These are questions about society’s values, not about science’s undiscovered secrets.”).

Many recognize the important limits of research findings in light of the role that values need to play in policymaking, both in education and well beyond.²²³ But there is some danger that governmental decision-makers are not always aware of these limits.²²⁴ They may assume that, in deciding to do just what the evidence says that they should, they are making “good government” decisions about how best to spend taxpayer dollars rather than engaging in any value-laden decision.²²⁵ And they may assume that the available evidence itself provides neutral answers, rather than being contingent on the kinds of questions that were asked,²²⁶ “trans-science” decisions made by researchers along the way,²²⁷ and the uncertainties inherent in social science research in general and education research in particular.²²⁸

223. See, e.g., Cohen et al., *supra* note 25, at 117 (“[R]esearch would not prescribe decisions about resources, for those require interactions among a range of persons and groups whose qualifications to decide are civic rather than scientific, and whose values often differ.”); Wong, *supra* note 163, at 223 (arguing that researchers should not see themselves as “expert problem solvers” but as “participants in democratic deliberation” (internal quotation marks omitted) (quoting Mary Jo Bane, *Presidential Address—Expertise, Advocacy and Deliberation: Lessons from Welfare Reform*, 20 *J. Pol’y Analysis & Mgmt.* 191, 195 (2001))); Robert Gordon & Ron Haskins, *The Trump Administration’s Misleading Embrace of ‘Evidence,’ Politico* (Mar. 31, 2017), <http://www.politico.com/agenda/story/2017/03/the-trump-administrations-misleading-embrace-of-evidence-000385> [<http://perma.cc/5X8C-58SZ>] (explaining that “evidence can only go so far” because “[t]he art of governing means setting priorities for what is worth trying to fix”).

224. See, e.g., Lesley Saunders, *Grounding the Democratic Imagination: Developing the Relationship Between Research and Policy in Education* 10 (2004) (“One risk associated with [the ostensibly ideology-free nature of evidence-based education] . . . is that value-positions disappear from sight as if by sleight of hand.”); Greenhalgh & Russell, *supra* note 222, at 315 (“[T]echnical fixes remain the holy grail of many government departments.”); Beryl A. Radin, *Neutral Information, Evidence, Politics, and Public Administration*, 76 *Pub. Admin. Rev.* 188, 189, 191 (2016) (critiquing both Democrats and Republicans “who believe that it is possible to find clear, neutral, and lasting answers” to policy questions because “[w]hile the concept of evidence-based decisions may have great appeal, information is rarely neutral and, instead, cannot be disentangled from . . . value, structural, and political attributes”).

225. Coburn et al., *supra* note 19, at 79 (“[A]dvocates of research-based programs and evidence-based decisionmaking often position their use as an antidote to overly politicized and ideological decisionmaking on the part of school and district leaders.”); Greenhalgh & Russell, *supra* note 222, at 310 (critiquing evidence-based policymaking for turning “political problems . . . into technical ones, with the concomitant danger that political programmes are disguised as science”); Simpson, *supra* note 158, at 451 (critiquing “‘metricophilia’: the expectation that quantitative data—virtually on their own—will give us the answers on which to base policy in education” (quoting Richard Smith, *Beneath the Skin: Statistics, Trust, and Status*, 61 *Educ. Theory* 633, 633 (2011))).

226. See, e.g., Carol H. Weiss, *Evaluation: Methods for Studying Programs and Policies* 314 (2d ed. 1998) (“Values are built into the study through the choice of questions . . .”).

227. See *supra* note 172 (discussing the challenge of trans-scientific questions); see also Giandomenico Majone, *Evidence, Argument, and Persuasion in the Policy Process* 65 (1989) (“A different conceptualization of the problem, other tools and models, or a few different judgments made at crucial points of the argument could lead to quite different conclusions.”).

228. See *supra* notes 141–186 and accompanying text.

The question of what works is no doubt an important one, but it should not be used to obscure the question that necessarily precedes it: What should we be trying to accomplish? Once we bring that question to the fore, it becomes clear that debates about evidence can sometimes be a cover for an underlying substantive disagreement about goals. Framing decisions as evidence based presents a danger that values-based decisions will be masked as neutral.²²⁹

In making the case for conducting experiments in education and relying on evidence to make education decisions, a leading scholar asks these rhetorical questions:

What if policy elites incorrectly concluded that Catholic schools are superior to public ones, and did something about this in the policies they created? What if they erroneously concluded that vouchers stimulate academic achievement, and did something about this in terms of funding priorities? What if they falsely concluded that school desegregation does not affect minority achievement when it does, and acted accordingly? Incorrect causal conclusions have costs in terms of dollars, achievements, and dreams.²³⁰

But what if the choices made by “policy elites” are really about values—for example, to take the above policies, a belief that religion should play a greater role in public and private life, that “government schools” stifle liberty,²³¹ and that maintaining racial hierarchies is justified—rather than evidence? When President Trump’s Office of Management and Budget (OMB) Director explained the Administration’s decision to eliminate funding for afterschool programs and associated nutritional supports by invoking a lack of evidence that these programs work to boost student achievement,²³² did he really mean that there is no evidence that they work? That he read the studies showing that they do and decided that the studies were flawed? Or was this decision part of a moral universe defining what role the government should play, as opposed to the

229. See Maris A. Vinovskis, *Missing in Practice? Development and Evaluation at the U.S. Department of Education*, in *Evidence Matters*, supra note 25, at 120, 124–25 (describing the way that research can operate “by shifting attention from moral commitment to analytical problems that rarely have clear-cut or simple solutions” (internal quotation marks omitted) (quoting Bruce K. MacLaury, Foreword to Henry J. Aaron, *Politics and the Professors: The Great Society in Perspective* (1978))); Gordon & Haskins, supra note 223 (describing “misleading” efforts by the Trump Administration’s Office of Management and Budget director to “use[] the language of evidence” while “selectively citing research,” and arguing that instead the Administration “should make a forthright argument about priorities”).

230. Cook, supra note 152, at 117.

231. Julie Bosman, *Public Schools? To Kansas Conservatives, They’re ‘Government Schools,’* N.Y. Times (July 9, 2016), <http://www.nytimes.com/2016/07/10/us/schools-kansas-conservatives.html> (on file with the *Columbia Law Review*).

232. Julia Zorthian, *White House: There’s No Evidence After-School Programs Help Kids’ Performance*, Fortune (Mar. 16, 2017), <http://fortune.com/2017/03/16/donald-trump-after-school-programs-performance-mick-mulvaney-budget/> [<http://perma.cc/W4V2-M7GZ>].

family, the market, or the church, with the references to evidence merely acting as a cover?²³³

It is dangerous when policy choices that are really rooted in values are framed as evidence based because debates about those policy choices end up taking place on a plane that is disconnected from reality. The Trump Administration is not going to be convinced to fund afterschool programs with a data dump of studies showing that the programs are successful. Instead, if the Administration changes its position on whether to fund a given program, it will be due to voters' moral outrage.²³⁴ And that outrage is not going to be prompted by studies finding that the programs are successful—if anything, such studies would be rejected as political.²³⁵ Appeals to values seem most likely to be successful to change the proposed policy.²³⁶

It is also dangerous to frame policies that are really rooted in values as instead based on evidence because policies can take on a life of their own, sometimes ending up disconnected from, or even in opposition to, the values that originally prompted the policy.²³⁷ Consider the rise of single-sex education over the last two decades. Early proponents included feminists who suggested that single-sex education could increase women's achievement.²³⁸ Yet they were soon joined by advocates peddling studies purporting to show differences in the way girls' and boys' brains learn; this development soon prompted schools to create single-sex classrooms to teach boys "heroic behavior" and allow them to run around while teaching girls "good character" and encouraging them to sit while discussing

233. See Gordon & Haskins, *supra* note 223 (discussing the research on afterschool programs in light of the OMB director's statement that they do not work, and critiquing the Trump Administration's decision to cut these programs as a reflection of unacknowledged values rather than actual reliance on evidence).

234. Cf. Michael Tonry, *Making Peace, Not a Desert*, 10 *Criminology & Pub. Pol'y* 637, 639 (2011) ("[Contemporary crime-control] policies will be repudiated, or support for them weakened, only if enough of their proponents can be persuaded that they are unjust and cannot be morally justified.").

235. See *supra* notes 161–170 and accompanying text for discussion of responses to conflicting studies connected to ideological disagreement.

236. See, e.g., Letter from David N. Cicilline & Lou Barletta, Members of Cong., to Mick Mulvaney, Dir., Office of Mgmt. & Budget (Mar. 15, 2017), <http://barletta.house.gov/sites/barletta.house.gov/files/documents/21st%20Century%20CLC%20OMB%20Letter.pdf> [<http://perma.cc/8UM7-FWBP>] (asking President Trump to reinstate funding for afterschool programs based on appeals to values such as preventing turmoil for working families, promoting safety for kids, and allowing working parents to maintain jobs).

237. Cf. Klingele, *supra* note 185, at 540–41 ("Although most proponents of evidence-based correctional practices frame them as rehabilitative tools designed to reduce the use of incarceration and make correctional interventions more modest and humane, these tools are capable of doing the very opposite.").

238. See Rosemary Salomone, *Rights and Wrongs in the Debate over Single-Sex Schooling*, 93 *B.U. L. Rev.* 971, 976–77, 1027 (2013) (noting publications in the 1980s and 1990s that "reaffirmed the idea that American schools, overwhelmingly coed, were 'shortchanging' girls").

their feelings.²³⁹ Such policies entrench the very sex stereotypes that many of the original proponents of single-sex education intended to demolish.

Similarly, when evidence is used ostensibly to justify a policy that actually stems from values, opponents of those values can use counterevidence to undercut the policy. For example, in *Brown v. Board of Education*, the Supreme Court famously (and controversially) relied in part on a study showing that black children preferred pink dolls to brown dolls to justify its ruling that segregation violated the Equal Protection Clause.²⁴⁰ Segregationists later used that same study to argue in favor of segregation, since the study showed that black children in segregated schools actually “exhibited less negative reactions to the brown doll and showed less desire to play with the pink doll.”²⁴¹ So did the evidence show that it was better to maintain segregated schools? When values are in play, neither side is likely to convince the other side with a neutral appeal to evidence.²⁴²

Moreover, there is a debate among education scholars and practitioners about what school reform should properly focus on: fixing the system as it exists, or reimagining the system entirely.²⁴³ Evidence is more likely to play a helpful role in the former than the latter.²⁴⁴ And tinkering can be valuable.²⁴⁵ But focusing too much on evidence to help us tinker can distract us from asking fundamental questions about the point of the whole endeavor.²⁴⁶ At a moment when the direction of (and perhaps

239. See *id.* at 980–83.

240. 347 U.S. 483, 494 & n.11 (1954) (citing social science research to justify the proposition that segregation “with the sanction of law . . . has a tendency to [retard] the educational and mental development of negro children and to deprive them of some of the benefits they would receive in a racial[ly] integrated school system” (internal quotation marks omitted) (alterations in original)).

241. Joshua Dunn & Martin West, *Calculated Justice: Education Research and the Courts*, in *When Research Matters*, *supra* note 14, at 155, 162–63.

242. See, e.g., Lindblom & Cohen, *supra* note 24, at 52 (connecting the perceived “authoritativeness” of social science research for its users with the extent to which the research “squares with their ideology or conforms to their general world view or epistemological position”).

243. See, e.g., David Tyack & Larry Cuban, *Tinkering Toward Utopia: A Century of Public School Reform 1* (1995) (describing the tension in Americans’ attitude toward school reform as one between “utopian thinking” and “tinkering,” each with both positive and negative connotations).

244. See Head, *supra* note 163, at 84 (“Evidence-based arguments about ‘fine-tuning’, based on careful research about effectiveness, might be more likely to gain traction in those areas that are away from the political heat.”).

245. See, e.g., Tyack & Cuban, *supra* note 243, at 10 (discussing “a positive kind of tinkering, adapting knowledgeably to local needs and circumstances, preserving what is valuable and correcting what is not”); cf. Paul Butler, *The System Is Working the Way It Is Supposed to: The Limits of Criminal Justice Reform*, 104 *Geo. L.J.* 1419, 1466 (2016) (explaining that in some circumstances, “ratchets”—those things that “somewhat work sometimes”—can be useful).

246. See, e.g., Tyack & Cuban, *supra* note 243, at 10–11 (arguing for “a vision of a just democracy” instead of the “radically restricted” debates over “educational and social goals” that have been recently prevalent); cf. Butler, *supra* note 245, at 1466–68 (arguing that

existence of) public education is at a crossroads,²⁴⁷ this latter conversation is critical. If the debate simmering beneath technocratic discussions about evidence is really between Black Lives Matter²⁴⁸ and God's Kingdom,²⁴⁹ those technocratic discussions are not likely to accomplish much.

CONCLUSION

ESSA's calls for evidence-based policymaking have an intuitive appeal. Of course we should spend our limited public dollars on what works, and we should not spend our limited public dollars on what does not. The complexities outlined above do not indicate that we should disregard evidence or give up in despair on the research endeavor, believing that data exist only in the eye of the beholder or that we are each entitled to our own facts.²⁵⁰

We should, however, be more realistic about our expectations for evidence-based policymaking in education and other social policies. We must understand that the answer to the question of "what works" will always be more complicated than a sound bite; research brooks no easy answers, and implementation is messy. We must also understand that asking "what works" is itself a value-laden question. What works for what? For whom? To what end?²⁵¹

addressing "ratchets" in criminal justice reform should not take the place of larger substantive changes).

247. Compare Nathan Diament, Opinion, *The Power of School Choice*, Wash. Times (Jan. 12, 2017), <http://www.washingtontimes.com/news/2017/jan/12/the-power-of-school-choice/> [http://perma.cc/Q63P-6KHD] (arguing that "Donald Trump and Betsy DeVos can improve K-12 education through choice"), with Rebecca Mead, *Betsy DeVos and the Plan to Break Public Schools*, New Yorker (Dec. 14, 2016), <http://www.newyorker.com/news/daily-comment/betsy-devos-and-the-plan-to-break-public-schools> [http://perma.cc/9MSS-QM9F] (arguing that the new Administration's "ideological embrace of choice" ignores the idea of the "public school as a public good" and unsettles the "fundamental premises that underlie our institutions of public education").

248. See, e.g., Emily Deruy, *How Black Lives Matter Activists Plan to Fix Schools*, Atlantic (Aug. 5, 2016), <http://www.theatlantic.com/education/archive/2016/08/the-ambitious-education-plan-of-the-black-lives-matter-movement/494711/> [http://perma.cc/AC99-KKC9].

249. See, e.g., Benjamin Wermund, *Trump's Education Pick Says Reform Can 'Advance God's Kingdom'*, Politico (Dec. 2, 2016), <http://www.politico.com/story/2016/12/betsy-devos-education-trump-religion-232150> [http://perma.cc/VA6R-4QML].

250. Cf. Majone, *supra* note 227, at 10–11 ("Evidence is not synonymous with data or information Facts can be evaluated in terms of more or less objective canons, but evidence must be evaluated in accordance with a number of factors peculiar to a given situation").

251. See Greenhalgh & Russell, *supra* note 222, at 315 ("[R]esearch evidence can and should inform policy judgments—but this evidence does not in and of itself provide the answer to the ethical question of 'what to do' (and in particular, 'how to allocate resources')."); Hess, *Conclusion*, *supra* note 25, at 253–54 ("Research has a vital role to play in democratic policy debate . . . not to dictate outcomes or to presume that public officials should be the handmaidens of researchers, but to ensure that public decisionmaking is informed by all the facts, insights, and analyses that the tools of science can provide.").

The authors of the important 2002 National Research Council report on scientific research in education referenced above argued that the “compelling culture of democratic accountability . . . demands evidence that public monies are wisely spent.”²⁵² This is true. But the compelling culture of democratic accountability also demands vigorous debate about the underlying goals to which public monies are put. We ought to highlight this debate whenever calls for evidence-based policymaking imply that technocratic, value-free solutions are only a research study away.

There is nothing wrong with the invocations of evidence in ESSA. It is good to encourage decisionmakers to canvas their needs and examine what research might help them meet those needs. But evidence requirements do not provide a meaningful way to constrain decisionmaking in education, nor are they likely to provide the answers that will fix the system once and for all. We should not let a focus on evidence distract us from the democratic debate at the core of education. In conversations about the right path forward for education policy—and other policies that form the fabric of our democracy—we ignore discussions of values at our peril.

252. See Feuer et al., *supra* note 134, at 5; see also *supra* text accompanying note 169.