

critics, then, feared that acknowledging the real complexity of the relationship between education and equality might undermine the larger goal of improving educational quality for students in need.

doi: 10.1017/heq.2017.35

“Seeing Like a State” in the Postwar Era: The Coleman Report, Longitudinal Datasets, and the Measurement of Human Capital

Ethan L. Hutt

Nearly a half-century later, Christopher Jencks’s 1969 quip that “like a veritable Bible, the ‘Coleman Report,’ is cited today on almost every side of every major educational controversy” continues to ring true.¹ Whether the issue is the efficiency of schools, the imperative of integration, or the capacity of public education to solve social problems, the debate—public and scholarly—occurs in the shadow of the Coleman Report.

Though it is difficult to deny the Coleman Report’s singular influence on conversations about American schooling, historians of education have an important role in properly situating it not just in the research on inequality or school effectiveness but also within larger historical narratives. There are two such narratives, in particular, that I hope to highlight here. The first concerns the historical development and operation of the “American education state”—that is, the variety of people, institutions, and governance structures that have both composed and constructed the American public education

Ethan Hutt (ehutt@umd.edu) is an Assistant Professor at the University of Maryland-College Park. This essay is part of ongoing work on the history of data collection and research on American schools. He would like to thank the audiences at the annual meetings of the History of Education Society and the Organization of American Historians as well as Nancy Beadie and Joy Williamson-Lott for their thoughtful feedback.

¹James S. Coleman et al., *Equality of Educational Opportunity* (1966; repr., New York: Arno Press, 1979); and Christopher Jencks, “A Reappraisal of the Most Controversial Educational Document of Our Time: The ‘Coleman Report,’” *New York Times Magazine*, Aug. 10, 1969, 12.

system.² A perennial operational challenge at the federal level has been developing ways to comprehend the sprawling system.³ Though the collection and dissemination of statistics had been the responsibility of the Department of Education since 1867, as Douglas Reed notes in the opening of his book, nearly a hundred years later the federal government still lacked a basic capacity to gather information about the operation of local schools.⁴

A partial solution to this information problem points to the second major story: the development of the infrastructure the federal government built to inform itself and the public about the operation of America's schools. Historians have increasingly documented the ways in which quantification serves as a technique of governance and a tool of statecraft, as well as the ways in which the data systems designed to produce these quantifications shape and become entangled with the underlying phenomenon.⁵ Narratives of the history of education research have tended to focus on the shifting role and strategies of

²Douglas S. Reed, *Building the Federal Schoolhouse: Localism and the American Education State* (Oxford: Oxford University Press, 2014). Though Reed uses the phrase "American education state" primarily to refer to the constellation of governmental and nongovernmental actors who "collectively operate as the education state," to the extent that it also is intended to denote various "forms of authority" that allow the state to operate, I think the development of educational statistics is an important part of this story (xii). For an example from an earlier period see, Nancy Beadie, "The Federal Role in Education and the Rise of Social Science Research: Historical and Comparative Perspectives," *Review of Research in Education* 40, no. 1 (March 1, 2016), 1–37.

³Raymond E. Callahan, *Education and the Cult of Efficiency: A Study of the Social Forces That Have Shaped the Administration* (Chicago, University of Chicago Press, 1964); Tracy L. Steffes, *School, Society, and State: A New Education to Govern Modern America, 1890–1940* (Chicago: University of Chicago Press, 2012); and David Tyack, *The One Best System: A History of American Urban Education* (Cambridge: Harvard University Press, 1974).

⁴Reed, *Building the Federal Schoolhouse*, xi.

⁵On the way data collection and analysis has shaped the ideas under examination see, for example, Paul N. Edwards, *A Vast Machine: Computer Models, Climate Data, and the Politics of Global Warming* (Cambridge, MA: MIT Press, 2010); Sarah E. Igo, *The Averaged American: Surveys, Citizens, and the Making of a Mass Public* (Cambridge, MA: Harvard University Press, 2007); Alice O'Connor, *Poverty Knowledge: Social Science, Social Policy, and the Poor in Twentieth-Century U.S. History* (Princeton, NJ: Princeton University Press, 2001); Theodore M. Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (Princeton, NJ: Princeton University Press, 1995); and Thomas A. Stapleford, *The Cost of Living in America: A Political History of Economic Statistics, 1880–2000* (New York: Cambridge University Press, 2009). For examples in the history of education, see Ethan L. Hutt and Mitchell Stevens, "From Soldiers to Students: The Tests of General Educational Development (GED) as Diplomatic Measurement," *Social Science History* 41 (forthcoming); Steffes, *School, Society, and State*, 15–46.

the federal government in supporting educational research or on the general failure of federally funded research to produce a basic science of education.⁶ Largely missing from this story is the inclination and capacity (often secured through contracts) of those in the federal government to produce its own information about schools. The Coleman Report is part of this story—the provision in the Civil Rights Act requiring the survey reflects, after all, this desire for information—but the story does not begin with Coleman or the Civil Rights Act of 1964.

In this essay, I try to provide a view into this story by examining the development of a new kind of federally funded national education data project: the longitudinal dataset. Enabled by advances in sampling design, computer data processing, and the expanded university and think-tank research infrastructure of the Cold War, the national longitudinal dataset was unique among prior federal data-collection efforts, both in its intention to provide a nationally representative sample of American schools and students and in its aim to capture the relationship between student traits and abilities, school characteristics, and life outcomes. The first of these efforts, entitled Project Talent (1960–1975), spanned the commission, release, and reaction to the Coleman Report, and therefore provides a useful context for tracing broader shifts in the thinking about the role of schools in shaping life trajectories.

This context helps illustrate the way in which “manpower” development and the application of quantitative techniques such as systems analysis continued to inform federal data-collection efforts and interpretation, even as the rhetoric of education policy became increasingly studded with discussions of race, educational equity, and equal opportunity during the 1960s. Just as importantly, it highlights the ways in which large, nationally representative surveys like Project Talent and the Coleman Report invited policymakers and scholars to think in increasingly national, decontextualized ways about the operation and effects of American schools in general. That this data was, to an unprecedented degree, sufficiently large and accessible enough to allow for analysis and reanalysis also provided the opportunity for scholars and policymakers to draw conflicting conclusions about the character of American schools—the contrasting, but equally stylized,

⁶On the history of education research see Ellen Condliffe Lagemann, *An Elusive Science: The Troubling History of Education Research* (Chicago: University of Chicago Press, 2002); and Maris A. Vinovskis, “The Changing Role of the Federal Government in Educational Research and Statistics,” *History of Education Quarterly* 36, no. 2 (Summer 1996), 111–28.

statistical portraits framing the need for different research and policies going forward.

Scholars often trace the interest in researching educational outcomes or examining the relationship between inputs and outcomes to the Coleman Report and debates over educational opportunity. However, prior to the passage of the Civil Rights Act of 1964, this relationship had become a matter of interest for a growing number of analysts who sought to understand how America could optimize its investment in schools in order to develop the intellectual talents necessary to win the Cold War. In studying this relationship, analysts sought to utilize the quantitative analytic techniques developed to guide military weapons development, strategy, and investment during World War II to solve the problem of school organization. Thus, in 1959, at the behest of the Ford Foundation, two analysts at RAND Corporation applied the company's trademark analytic tool—systems analysis—to the study of a school system.

RAND was one of a growing number of independent, though largely military-funded, research organizations that sought to develop quantitative techniques capable of analyzing the increasingly complex and interrelated systems that make up modern society.⁷ Whether it was the design of urban spaces, the electrical grid, health care systems, or schools, researchers believed that applying these techniques would improve the design and operation of these systems in a way that would optimize their outputs.⁸ The impetus for their development, and one factor driving their proliferation, was the increasingly strong belief among many social scientists that traditional analytic tools were insufficient to guide decision-making in a society that comprises increasingly complex systems and that is increasingly awash in data on their operation. Tools that could structure and simplify this complexity in a way that made rational choices possible were thus at a premium.⁹

⁷On the history of RAND, see David Raymond Jardini, "Out of the Blue Yonder: The RAND Corporation's Diversification into Social Welfare Research, 1946–1968" (PhD diss., Carnegie Mellon University, 1996).

⁸On the adaptation and application of these techniques in civilian contexts see, for example, Agatha C. Hughes and Thomas P. Hughes, eds., *Systems, Experts, and Computers: The Systems Approach in Management and Engineering, World War II and After* (Cambridge, MA: MIT Press, 2000); and Jennifer S. Light, *From Warfare to Welfare: Defense Intellectuals and Urban Problems in Cold War America* (Baltimore: Johns Hopkins University Press, 2003). In the context of education, see the ongoing work of Elizabeth P. Berman, "The Organizational Legacy of PPBS," (working paper, University at Albany, State University of New York, 2017).

⁹One the importance of rational choice models see, for example, Paul Erickson, "Mathematical Models, Rational Choice, and the Search for Cold War Culture," *Isis*

RAND's systems analysis approach combined and elaborated a variety of quantitative techniques, including cost-benefit analysis and a branch of military analysis known as Operations Research, in ways that provided for a quantitative comparison of a system's output given a variety of different system-input specifications. The resulting analysis would allow the analyst to recommend the optimal choice among a variety of competing options. This overriding concern for choice and optimization was evident in the RAND study of education. Emphasizing that from a system perspective no difference existed between an air force radar system, a business, or a school district—"in all of these systems there are various ways of combining elements or inputs in order to produce outputs"—systems analysis provided the opportunity to "try out' innovations" by manipulating various inputs within the analysis and calculating their effects on cost and output. Though their analysis was preliminary, involving records from a single school district, the analysts concluded that their study demonstrated that "it would soon be feasible to make comparisons ... that can help administrators and others *choose* improved educational systems [emphasis added]"—ones that maximized any number of potentially desired outcomes, from scholastic achievement and creativity to "social poise" and physical health.¹⁰

The primary obstacle to achieving feasibility was not so much the complexity of the analysis, which they acknowledged, as it was the paucity of available data to use in their analyses. In contrast to so many other fields, there was simply insufficient school system data to feed into the systems analysis to produce useful comparisons of alternative choices. Still, the analysts took solace in the fact that the federal government, with appropriations made through the Cooperative Research Act (1954), had launched two massive data-collection efforts, the results of which they believed "will tell us what we need to know about the relationship between school characteristics and educational output."¹¹

This research not only captures the new ambition to view American schools as a rational system composed of variables available for manipulation—whether via hypothetical analysis or policy—but also the way in which these views and the analytic methods that informed them shaped data collection on America's schools. It

101, no. 2 (June 2010), 386–92; and Hunter Heyck, *Age of System: Understanding the Development of Modern Social Science* (Baltimore: Johns Hopkins University Press, 2015)

¹⁰Joseph A. Kershaw and Ronald N. McKean, *Systems Analysis and Education* (Santa Monica, CA: RAND Corporation, 1959), iii, 8. The analysts ultimately opted for academic achievement as measured by the California Assessment Test.

¹¹Kershaw and McKean, *Systems Analysis and Education*, 57.

would be a former RAND analyst, Alexander Mood, who transformed the Civil Rights Act's call for a survey of educational opportunity into a massive quantitative survey analysis.¹² The desire for standardized data on America's schools proved easier to dream than to deliver. The experience of the early data-collection efforts referenced by the RAND analysts would reveal just how much the idiosyncratic reality of American schooling diverged from their vision, how much work it would take to bring it into view, and how much statistical airbrushing would be required to make it accessible to statistical analysis.

One of these federally funded efforts was called Project Talent and involved the most ambitious education research project ever attempted to date.¹³ The project was led by John C. Flanagan, a professor at the University of Pittsburgh and founder of the American Institutes for Research (AIR) think tank. Flanagan, a Harvard trained psychologist, had spent World War II in the Army Air Forces Aviation Psychology Program designing test batteries more capable of predicting which recruits would succeed as pilots and which were better suited for alternative roles like bombardier or navigator.¹⁴ After the war, he hoped to continue researching in this vein and to use AIR as a vehicle for applying these techniques to governmental and private-sector problems involving the development and selection of human resources.¹⁵

In conceiving of Project Talent, Flanagan merged educators' long-standing concerns with individual development and vocational guidance with contemporary Cold War concerns for maximizing American productivity.¹⁶ These concerns included not only the narrow issue of identifying and expanding American scientific expertise but also the broader matters of maximizing labor force productivity by efficiently matching people with jobs well-suited to their abilities

¹²Gerald Grant, "Shaping Social Policy: The Politics of the Coleman Report," *Teachers College Record* 75, no. 1 (Sept. 1973), 19.

¹³John C. Flanagan, *Design for a Study of American Youth* (Boston: Houghton Mifflin, 1962). The other major data collection project was Samuel Goodman, *The Quality Measurement Project: A Research Activity Conducted by the New York State Education Department* (New York: New York State Department of Education, 1958).

¹⁴John C. Flanagan and the Army Air Forces, *The Aviation Psychology Program in the Army Air Forces* (Washington, DC: Government Printing Office, 1948).

¹⁵John C. Flanagan, "Some Notes on the Administrative History of the American Institutes for Research in the Behavioral Sciences" (Washington, DC: Records of the American Institutes for Research, n.d.).

¹⁶On the shifting focus on guidance counseling, see: David Gamson, "From Progressivism to Federalism: The Pursuit of Equal Educational Opportunity, 1915–1965," in *To Educate a Nation: Federal and National Strategies of School Reform*, ed. Carl Kaestle and Alyssa Lodewick (Lawrence: University Press of Kansas, 2007).

and educational opportunities to further develop their skills.¹⁷ Flanagan believed these problems could be more effectively addressed if researchers could understand the fundamental relationships between educational systems, the development of student abilities, and their ultimate career outcomes. Knowing this relationship would allow both policymakers to better allocate school resources and school officials to provide more timely information to students about their likely career trajectories. As a sales brochure for the project's findings proclaimed: "To discover youth's aptitudes, talents, and creativity ... to meet the country's acute need for trained personnel in all fields ... Project Talent will yield accurate facts, understanding, and knowledge to turn potentialities into skilled manpower."¹⁸

The only way, in Flanagan's view, to ascertain these relationships—to know whether and how the potential was fulfilled—was to conduct a massive, longitudinal "census" of American talent and survey of American school organization.¹⁹ Only a massive survey carried out over an extended period of time would allow him to determine the relationship between student talents, school variables, and career success across the entire occupational spectrum. The final design called for a nationally representative sample of 440,000 American high school students (roughly one out of twenty) and 1,353 high schools, with follow-up surveys conducted with students at one, five, and ten years after graduation. Beyond the immense logistical challenge the study design posed, the biggest obstacle to the study's execution was that, while Flanagan proposed a study of American talent, there were no standardized definitions for school features, pathways, or curricula. If Flanagan was going to bring into view a picture of the American school and the American student, he would have to do so not only through

¹⁷For an example of the issues included under "manpower" concerns, see William Haber and the Industrial Relations Research Association, *Manpower in the United States: Problems and Policies* (New York: Harper, 1954). On the search for science talent in schools, see Sevan G. Terzian and John L. Rury, "A Highly Selected Strain of Guinea Pigs?: The Westinghouse Science Talent Search and Educational Meritocracy, 1942–1958," *Teachers College Record* 116, no. 5 (2014), 1–33.

¹⁸Houghton Mifflin Company, *The Talents of American Youth, An Important Series of Books about Project Talent, An Historic First* (sales brochure) (Washington, DC: Records of the American Institutes for Research, n.d.).

¹⁹An important Project Talent precursor was Robert L. Thorndike and Elizabeth Hagen, *Ten Thousand Careers* (New York: John Wiley & Sons, 1959). Thorndike and Hagen conducted a follow-up survey with ten thousand people who had taken test batteries during World War II as part of Flanagan's Aviation Psychology Program to determine whether the information contained in the test batteries could have been used to predict future career and career success. Flanagan served on the advisory committee for this effort; Thorndike, in turn, would serve on the advisory committee for Project Talent.

conducting a survey but through its construction as well. To help with this task, Flanagan enlisted a technical panel of thirty-one prominent researchers—including Henry Chauncey, E. Franklin Frazier, Samuel A. Stouffer, and Robert L. Thorndike—to develop from scratch a test battery that ultimately consisted of twenty-five academic and psychological subtests, a student interest and activity inventory, a measure of personal preferences, and two short open-ended essays, the entirety of which took two-and-a-half days to administer.²⁰

Despite Flanagan's hope, the resulting billion pieces of data mostly offered support for the "small relationship between the amount of student learning" and such school variables as "school size, class size, school building age, rural versus urban location, and dropout rate." Flanagan also found considerable evidence that socioeconomic status was at least as important as academic achievement in predicting college enrollment.²¹ Flanagan spun these findings as evidence of the ineffective guidance programs and failure of American high schools to develop individual talent. But they did not come close to fulfilling the promise of being able to divine the relationship between school characteristics, individual talent development, and career success. Though Flanagan and his associates were fond of likening standardized testing to the physical scientist using X-rays to study the crystalline structure of molecules, and the Project Talent data bank to the centuries of astronomical and botanical observations that led to scientific breakthroughs for Johannes Kepler and Charles Darwin, the seeming failure of his immense dataset to reveal the core structure of the school system was immensely disappointing.²²

This failure has led many historians to ignore or dismiss Project Talent as, in the words of one historian, "an exercise in overkill."²³ But I want to suggest that the contemporary response to Project Talent provides insights into a major shift in educational research embodied in both Project Talent and the Coleman Report and the subsequent direction of large scale, federally directed research surveys.

First, it showed the intent of researchers to nationalize the conversation about the conception and quality of American schooling. At a time when many scholars, including James Conant, expressed skepticism

²⁰Flanagan, *Design for a Study*, 34–66.

²¹John C. Flanagan, *Project Talent and Related Efforts to Improve Secondary Education* (Bloomington, IN: Phi Delta Kappa International, 1978), 17–19.

²²Flanagan, *Design for a Study*, 1; and William W. Cooley, "The Project Talent Data Bank," *Journal of Educational Measurement* 2, no. 2 (Dec. 1965), 134.

²³Joseph F. Kett, *Merit: The History of a Founding Ideal from the American Revolution to the Twenty-First Century* (Ithaca, NY: Cornell University Press, 2012), 157.

about the value and wisdom of generalizing about the “American school” given the history of local control, Project Talent demonstrated that both the technical tools and analytic techniques necessary to conjure a stable, if stylized, image of the American school system had arrived.²⁴ While considerable local and state variation remained—and Coleman’s research would highlight the importance of *within school* variation—conversations were increasingly driven by decontextualized generalizations about national and regional averages. Ironically, concerns state and local officials expressed about researchers’ ability to make direct comparisons between districts forestalled alternative designs that would have allowed for greater discussion of state and local variation—something that affected not only Project Talent and the Coleman Report but also the design of the National Assessment of Educational Progress (NAEP).²⁵ Beyond discussing school in national terms, Project Talent set a new standard for evaluating school effectiveness both longitudinally and in terms of life and career outcomes.²⁶

Second, the large-scale, quantitative, computer-readable data these surveys produced allowed the datasets themselves to become part of the story, as scholars analyzed and reanalyzed the data in an effort to extract new insights and discern its “real” meaning. Of course, while these massive datasets offered an unprecedented opportunity to study the relationships between students and schools, it did not do so equally. The resources and technical abilities necessary to analyze this data clearly favored scholars with statistical training at larger institutions with

²⁴James B. Conant, *The American High School Today: A First Report to Interested Citizens* (New York: McGraw-Hill, 1959), 16; and Donald T. Campbell, “Administrative Experimentation, Institutional Records, and Nonreactive Measures,” in *Improving Experimental Design and Statistical Analysis*, ed. Julian C. Stanley (Chicago: Rand McNally, 1967), 257–91.

²⁵Maris A. Vinovskis, *Overseeing the Nation’s Report Card: The Creation and Evaluation of the National Assessment Governing Board* (Washington, DC: National Assessment Governing Board, 2001), 6–7. The evolution of Title I reporting results offers an interesting contrast. Initial federal tolerance for localized and idiosyncratic evaluations gave way to the Title I Evaluation Reporting System (TIERS) (1979), which required that Title I evaluations be conducted using norm-referenced standardized tests and that program effects be reported in terms of Normal Curve Equivalents—a metric newly created for the purpose of these evaluations.

²⁶Notable examples of studies funded by the Department of Education include the *Youth in Transition Study* (1965); the American “Freshman Survey” (1966); and the National Longitudinal Study of 1972 (NLS-72). These should be considered in light of a broader interest in longitudinal studies initiated in this period, including the *Michigan Panel Study of Income Dynamics* (1968) and the Bureau of Labor Statistics’ *National Longitudinal Surveys* (1965). In the realm of popular culture, the classic British documentary *Up* series began following fourteen school children in 1964 and has continued airing installments at seven-year intervals with the most recent, *56 Up*, airing in 2012.

computing capabilities. To the extent that this data had an outsized influence on future research and policy discussions because of its size and national representation, it did so in a way that reflected the specific concerns of these scholars and the constraints of the survey creators. For instance, despite the vast amount of data Project Talent collected on students—and subsequent widespread use by scholars—one variable was omitted: race. This decision reflected the project's concern for individual development, not equal opportunity or racial justice.²⁷

Finally, despite the hope that the unprecedented size and detail of the Coleman Report and Project Talent would reveal the relationships between students, schools, educational opportunity, and career trajectories, they ultimately cast as much shadow as illumination. Whether one chose to interpret the darkness or the light—and what one saw in those spaces—offered a Rorschach test of ideological and methodological commitments. Christopher Jencks, for instance, argued repeatedly that, even beyond the Coleman Report, Project Talent provided the “best available evidence” of the inability of schools—regardless of their characteristics—to address inequity.²⁸ Others, however, like economist Alice Rivlin, who served as President Johnson's Assistant Secretary of Planning and Evaluation, explained away the Project Talent results by arguing that the dataset was large, but not large enough. What was needed was “a longitudinal data system for keeping track of individual students as they move through school”—a critique of Coleman's “snapshot” view and Project Talent's failure to collect information on course-taking and specific school resources directed at individual students. The real value of Project Talent, Rivlin argued, was that it justified the funding of “more complex and expensive longitudinal studies”—studies for which Project Talent served as the explicit blueprint.²⁹

Though the next federal longitudinal survey (NLS-72) would update its statement of purpose to include the study of “access to educational . . . opportunity,” the commitment to conceptualize education at the national level and to view schools as systems composed of different inputs, but nevertheless governed by generalizable rules that could be made visible through statistical analysis, remained the

²⁷After the publication of the Coleman Report, there was an effort to include the item in follow-up surveys. See Alan E. Bayer, “Construction of a Race Item for Survey Research,” *Public Opinion Quarterly* 36, no. 4 (Dec. 1972), 592–602.

²⁸Jencks's effort would, itself, be critiqued as flawed, inconclusive, and politically motivated. See the special issue on Jencks, “Perspectives on Inequality,” *Harvard Educational Review* 43, no. 1 (April 1973).

²⁹Alice Rivlin, *Systematic Thinking for Social Action* (Washington, DC: Brookings Institute, 1972), 65.

foundation of the enterprise.³⁰ Though Rivlin conceded that “the problem may be that the real world is not organized to generate information about [economic] production functions, no matter how cleverly the statistics are collected,” these concerns did not prevent policymakers and scholars over the last half century from endeavoring to try.³¹ These efforts, the choices they involved, and the consequences for how we have conceptualized and evaluated the American education system remain an important, and underexamined, legacy of the Coleman Report.

doi: 10.1017/beq.2017.36

³⁰Samuel Peng et al., *National Longitudinal Study of the High School Class of 1972: Review and Annotation of Study Reports* (Washington, DC: National Center for Education Statistics, 1977), vii.

³¹Rivlin, *Systematic Thinking*, 65.