

October 2017

Measuring Up: Standardized Testing and the Making of Postwar American Identities, 1940-2001

Keegan J. Shepherd

University of South Florida, keegan.shepherd@gmail.com

Follow this and additional works at: <http://scholarcommons.usf.edu/etd>

 Part of the [American Studies Commons](#), and the [United States History Commons](#)

Scholar Commons Citation

Shepherd, Keegan J., "Measuring Up: Standardized Testing and the Making of Postwar American Identities, 1940-2001" (2017).
Graduate Theses and Dissertations.
<http://scholarcommons.usf.edu/etd/7086>

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.

Measuring Up: Standardized Testing and the Making of Postwar American Identities, 1940-2001

by

Keegan J. Shepherd

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of History
College of Arts and Sciences
University of South Florida

Major Professor: David Johnson, Ph.D.
Julia Irwin, Ph.D.
Brian Connolly, Ph.D.
Mitchell Stevens, Ph.D.

Date of Approval:
August 04, 2017

Keywords: education, college, assessment, brainpower, political culture

Copyright © 2017, Keegan J. Shepherd

DEDICATION

Dissertations are rather ridiculous and exhausting things to make. If you are lucky, you find yourself surrounded by people who encourage you to keep going—not because you already have a great deal of student loan debt to your name, but because you may have a genuinely worthwhile set of ideas that deserve to exist in public. I am incredibly lucky.

I must thank my husband, Evan Purchell, for sticking with me even through I was often a complete anxious wreck, raccoon-eyed and caffeine-addled, and definitely on more than one occasion a total jerk. You make everything easier, and there is no way I could have even come close to finishing without your encouragement and, sometimes, pho and a Renee Harmon film.

I must thank my parents, Judith Shepherd and John Shepherd, for always being behind my aim to become a historian, even when it made no practical sense. (It still doesn't.)

I must thank the historians and sociologists I met at the University of South Florida who made me a better thinker and kindly put up with me for years: Barbara Berglund, Catherine Conner, Julia Irwin, Brian Connolly, Nadia Jones-Gailani, Robert Benford, and Sara Crawley. Whether or not I directly studied under you, I learned so much.

I must especially thank my advisor, David Johnson. Thank you for helping me find my voice, in my writing and elsewhere. Thank you for also getting Mitchell Stevens to agree to join my committee. Speaking of: Thank you, Mitchell, Julia, and Brian. I have to thank the friends I made in the history department at the University of South Florida—especially Brittany Merritt,

Ashley Buchanan, Sarah Jünke, and Bradley L’Herrou—for being tremendous listeners, conversationalists, and confidantes.

I really need to thank the friends I made outside my department. My friendship with Neal Fischer started as we worked together in our graduate employee union, and grew so much through Scrabble (and board game trash talk) that we commemorated it with a tattoo. My other dear comrades at Graduate Assistants United—especially Megan Flocken, Tanya de Dios, Josh Lenes, Suzanne Young, and Erin Sauer—made me a committed activist, deeper analyst, stronger communicator, and more outgoing person. Andrea Ranieri was not only the best roommate anyone could ask for but also helped me get a job: you are an incredible friend. Dana Rognlie completely changed how I looked at the world when we were co-chairs of AGEL. Mary Catherine Whitlock was a queer role model—a true badass in and out of academia—and such a good squirrelfriend that we, too, got matching tattoos. Hilary Dotson and Lindsey Turnbull are the truest of friends with the greatest hospitality. Dan Guadagnolo and Jacob Barto made me feel like I had sisters in every city. Michael Newberger, Jordan Schultz, and Mike Onate are real ride-or-dies, and the best company for professional wrestling, B-movies, and *Twin Peaks*.

Finally, I have to thank Tampa and all the amazing people I met during my five years in the city. Tampa is grungy and sweaty and sunburnt, brightly patterned and half-drunk—and I love it dearly. The folks who run Florida Avalon, the bartenders at the Independent, the baristas at Felicitous: you all make the city a fantastic place to be. Tampa, I love you all, and miss you so.

As I said, I was incredibly lucky.

ACKNOWLEDGMENTS

This research would not have been possible without the generous help of several institutions. I must thank both the Harry S Truman Library Institute in Independence, Missouri, and Gerald R. Ford Foundation in Ann Arbor, Michigan for awarding me travel research grants to their respective presidential libraries. Their archival material made several of these chapters possible. I must also thank the University of South Florida for awarding me a dissertation completion fellowship, so that I could turn by ideas into a defendable draft. Thank you.

TABLE OF CONTENTS

Abstract.....	iii
Introduction: The Scale and Scope of Standardized Testing.....	1
What, Exactly, Is a Standardized Test?	11
A Brief History of Standardized Testing to the late 1930s	15
Testing and the Postwar World	29
Chapter One: Manpower from the Masses: Standardized Testing, State Legitimacy, and Postwar Order	36
Harnessing Manpower: Resource Planning, Vocational Guidance, and Wartime Education	39
Measuring the Masses: The General Classification Test and Preparatory Guides.....	51
Postwar Readjustment: The GI Bill, Guidance Counseling, and Personality Inventories	64
Chapter Two: American Talent Show: Standardized Testing and the Cold War Scramble for Brainpower.....	78
Harnessing Brainpower: Educational Testing Service, the Korean War, and the SSCQT	80
Test Early, Test Often: Guidance Counseling, Gifted Youth, and Scholarship Contests	96
Sputnik and the Federal Embrace of Testing	112
Chapter Three: Mapping Marginal Men: Standardized Testing as a Form of Racial Social Control in the Postwar State	126
Racial Ideology, the GCT, and Desegregation.....	128
Standardized Testing as a Means for Postwar Racial Social Marginalization	149
Chapter Four: Bad at Math: Standardized Testing, Mathematics Achievement, and Gender Politics	162
Gender Ideology, Postwar Science, and Functionalist Theory.....	163
Federal Fascination with the Gender Gap	187
The Enduring Mystery of the Gender Gap	196
Chapter Five: The Birth of the Student Consumer: Standardized Testing, Consumer Advocacy, and the Educational Marketplace	202
Consumerism as a Framework for Understanding Standardized Testing	205

Ralph Nader and Public Distrust of Educational Testing Service.....	215
The Legitimation and Expansion of the Test-Prep Industry	226
The Knotty Logic of the Contemporary Educational Marketplace	244
Conclusion: Standardized Testing Regimes in an Uncertain Educational Future.....	252
References.....	257

ABSTRACT

Standardized testing is a defining feature of contemporary American society. It not only governs how people are channeled through their schooling; it amplifies existing social disparities. Nonetheless, standardized testing endures, namely because it has served as a vital tool for the post-1945 American state. The postwar state prioritized, on the one hand, the cultivation of intellects resilient enough to sustain American geopolitical supremacy through scientific discovery and technological innovation and, on the other hand, the maintenance of an obedient population that would not disrupt existing social hierarchies. Standardized testing helped the postwar state solve this mind-body dilemma. As a function for social order, standardized testing provided the means for governing bodies to make sense of their citizens—particularly to gauge the skills, knowledge, and ability youth could eventually bring into a labor force. Standardized testing thus makes it clear not just who the so-called best and brightest are, but how well a population of students has adjusted to a set of educational norms the state has deemed necessary for future social success. Standardized testing is thus a way to monitor and enforce educational compliance with projected state and social needs.

This dissertation examines how standardized testing became a vital instrument for the postwar state—and, in turn, how this state dependence on standardized testing fueled several larger postwar political cultures. This dissertation also focuses on the tension between the ways the state made sense of its citizens and the ways citizens made sense of society through standardized tests. Standardized testing acted as a massive social sorter in postwar America, but

with time, many groups of Americans began to question the wisdom of educational reliance on standardized testing, as well as challenge the foundational assumptions about what testing actually measured. African-Americans, women, and working-class Americans pursued legal, legislative, and academic methods to push back against unjust standardized testing practices. This resistance also often provoked responses from those who had great stakes—whether socioeconomic or corporate—in maintaining the use and meaning of standardized tests. Politicians, parents, consumer advocates, academics, educational reformers, feminists, civil rights activists, entrepreneurs, and white supremacists all interpreted standardized testing scores and trends for larger political ends, often using similar information as their opponents to stake far different positions about race, gender, class, and merit. Yet, as more Americans questioned the legitimacy of standardized testing, they often reinforced standardized testing data’s use as a rhetorical tool, ultimately entrenching testing data as a way to make sense of society, even as more and more Americans find testing regimes purposeless.

This dissertation shows that, by the beginning of the twenty-first century, standardized testing data had become a potent tool for contending who mattered and who did not for America’s future. Because debates about standardized testing’s social utility often hinged on the meaning of test data—whether it reflected objective truths about the natural distribution of mental aptitudes or, instead, exposed the biases psychologists built into their devices as well as the prejudicial baggage that informed laypeople’s interpretation of statistical information—standardized testing itself continues to have value as a political weapon. Whether or not one has a specific policy proposal for the future use of standardized testing, its rhetorical function as a symbol for what is wrong with America will continue to fuel numerous, often oppositional, debates.

INTRODUCTION: THE SCALE AND SCOPE OF STANDARDIZED TESTING

Standardized testing governs American life. From the earliest stages of education, American children face a steady stream of standardized tests.¹ The results from standardized tests shape how teachers, parents, guidance counselors, and administrators view students—and, in turn, how students see themselves. These test scores form the path a student takes during their schooling: whether they are placed in advanced or remedial courses; whether they pass from one grade to the next; whether they are accepted into a more selective school; and whether they are granted entry into a graduate program. Standardized test scores also determine whether schools gain or lose funding and if teachers earn promotion. Standardized testing may carry high stakes for all educational participants—but it also has widespread effects well beyond the classroom.

Realtors refer to the average SAT scores in a district when marketing property in affluent

¹ Many researchers agree that strictly using standardized tests to assess learning in children under the age of 8 is an unethical practice that produces wildly unreliable results—and that educators should, at the very least, incorporate additional methods of observation and tracking before third grade in order to identify and monitor children who may have learning deficits or exceptional educational needs. Such professional reticence, however, did not stop the G.W. Bush and Obama Administrations from trying to implement national standardized assessments for preschool- and early-elementary-aged children. The ethical haziness in both cases resided in whether such assessments would have relied on data-gathering methods unsuitable for small children—and what potential consequences early-childhood educators would have faced if children entered elementary school without having demonstrated proficiency in certain areas. See: Lorrie Shepard, Sharon Lynn Kagan, and Emily Wurtz, eds., “Principles and Recommendations for Early Childhood Assessments,” National Education Goals Panel, February 1999, <http://govinfo.library.unt.edu/negp/reports/prinrec.pdf>; Ann S. Epstein *et al.*, “Preschool Assessment: A Guide to Developing a Balanced Approach,” *Preschool Policy Matters* 7 (July, 2004), <http://www.doe.in.gov/sites/default/files/earlylearning/preschool-assessment-guide-developing-balanced-approach.pdf>; Marcy Guddemi and Betsy Case, “Assessment Report: Assessing Young Children, Pearson, 2004, http://images.pearsonassessments.com/images/tmrs/tmrs_rg/Assessing_YoungChildren.pdf?WT.mc_id=TMRS_Assessing_Young_Children; Valerie Strauss, “Bush Plan to Assess 4-Year Olds’ Progress Stirs Criticism,” *Washington Post*, January 17, 2003; and Michele McNell, “New Race to Top Stresses Pre-K Tests, Early Ed. Program Ratings,” *Education Week*, July 1, 2011, http://blogs.edweek.org/edweek/campaign-k-12/2011/07/to_compete_states_must.html?ga=2.128105100.1239417728.1493166805-1254062810.1426548188 (all links last accessed April 25, 2017).

neighborhoods.² Massive corporations generate hundreds of millions in revenues selling standardized tests to various school districts and state education boards, all while other massive corporations generate hundreds of millions in revenues selling pricey standardized test preparatory guides, courses, and private tutoring programs.³ Standardized testing has even manifested its own unique version of anxiety, leading psychologists to consider the therapeutic and pharmaceutical regimens students could use to weather potentially disabling scenarios.⁴ The anxiety that many feel over standardized testing reflects the disproportionate role it plays in everyday American life.

Many Americans have become so frustrated with the ubiquity of standardized testing that they have developed an entire political culture in response. The contemporary anti-testing movement has its anchorage in a broader politics of the family, which advocates that the rights of

² Lauren Meade, “The Higher the SAT Scores, the More the House is Worth,” *Christian Science Monitor*, April 28, 2005, <http://www.csmonitor.com/2005/0428/p11s01-lihc.html>. Observe as well the routine way in which SAT scores are remarked in real estate profiles of tony areas; see: Elsa Brenner, “Chappaqua, N.Y.: A Hamlet in a Woodsy Setting,” *New York Times*, September 3, 2014, <https://www.nytimes.com/2014/09/07/realestate/chappaqua-ny-a-hamlet-in-a-woodsy-setting.html> (all links last accessed April 25, 2017).

³ The British company Pearson VUE, one of the largest test providers for the United States, posted revenues of over 4 billion pounds in 2016, while Graham Holdings posted over \$280 million in revenues through its test preparation holdings alone; see: Pearson, “Press Release: Pearson 2016 Preliminary Results (Unaudited),” February 24, 2017, https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/news/news-announcements/2017/2016_Full_Year_Results_Press_Release.pdf; Graham Holdings, 1, *2016 Annual Report*, <http://www.gcco.com/phoenix.zhtml?c=62487&p=irol-reportsannual> (all links last accessed April 25, 2017).

⁴ Psychologically speaking, test anxiety is not a distinct diagnostic category in the most recent edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), and is usually located within a broader umbrella of anxiety disorders. Absence in the DSM-5 aside, psychologists have spent decades determining the existence of test anxiety, its physiological and cognitive effects, its relationship to learning disabilities and other anxiety disorders, and its measureable effects on test performance. Test anxiety’s amorphous quality—a subject of considerable psychological analysis, but not an official diagnostic category—ironically makes it easier to circulate as a popular topic among nonprofessional circles. For recent examples of psychological research, see: Dubi Lufi and Lina Darliuk, “The Interactive Effect of Test Anxiety and Learning Disabilities Among Adolescents,” *Educational Research* 43 (2005): 236-249; Markus Sommer and Martin E. Arendasy, “Comparing Different Explanations of the Effect of Test Anxiety on Respondents’ Test Scores,” *Intelligence* 42 (2014): 115-127; and Frank Herzer, Julia Wendt, and Alfons O. Hamm, “Discriminating Clinical from Nonclinical Manifestations of Test Anxiety: A Validation Study,” *Behavior Therapy* 45 (2014): 222-231.

parents take precedent over the responsibilities of the state.⁵ A groundswell of parents, concerned that educators are stuck doing little more than “teaching to the test,” have staged widespread protests of statewide testing programs.⁶ In New York State alone, parents pulled over 200,000 children from statewide standardized tests in 2015—and in some districts, over two-thirds of students refrained from that year’s exams.⁷ Even Massachusetts, which had been at the vanguard of developing high-stakes testing programs that generated more ambitious curricula as well as demonstrable growth in student learning, backed away from proposed multistate testing schemes after considerable public antipathy.⁸ Fueled partially by social media networks,

⁵ For some sense of the background of this contemporary political culture, see: Donald T. Critchlow, *Phyllis Schlafly and Grassroots Conservatism: A Woman’s Crusade* (Princeton: Princeton University Press, 2005); Robert O. Self, *All in the Family: the Realignment of American Democracy Since the 1960s* (New York: Hill and Wang, 2012); and Georgina Denton, “Neither Guns Nor Bombs—Neither the State Nor God—Will Stop us From Fighting For Our Children: Motherhood and Protest in 1960s and 1970s America,” *The Sixties: A Journal of History, Politics, and Culture* 5, no. 2 (November, 2012): 205-228.

⁶ Teaching to the test, by contrast, is often advocated by thinktanks, and reflects a centrist consensus on education in which educational assessment is pegged to social order in a fashion that presumes the deviance of students and teachers who do not willingly submit to such a system. In one confounding example, printed in the prestigious peer-reviewed *Quarterly Journal of Economics*, depicted as a necessary function similar to speed traps or anti-terrorist measures. In this article, later used by The Brookings Institute as the premise for its support for teaching to the test, the author remains unclear about why teachers should be related to speeding drivers or terrorists, and what good comes from an educational apparatus premised on the presumed deviance of educators. (Or, for that matter, why students would be cars or bombs in such economic analogies.) Dismissing such an argument as inane, however, misses the role that criminality and obedience play in many high-stakes testing regimes—namely, that teachers are presumed incapable of effectively guiding students toward learning objective unless threatened with negative consequences based on poor test scores. The test—and more importantly, the threat of occupational sanctions following low test scores—is thus rendered the only acceptable motivation for learning. See: Edward P. Lazear, “Speeding, Terrorism, and Teaching to the Test,” *Quarterly Journal of Economics* 121, no. 3 (August, 2006): 1029-1061; Michael Hansen, “Teaching to the Test: Hype or Reality?” Brown Center Chalkboard blog, Brookings Institute, <https://www.brookings.edu/blog/brown-center-chalkboard/2016/05/16/teaching-to-the-test-hype-or-reality/> (last accessed April 24, 2017); and Roland G. Fryer, Jr., *et al*, “Enhancing the Efficacy of Teaching Incentive Through Loss Aversion: A Field Experiment,” National Bureau of Economic Research Working Paper 18237 (Cambridge: NBER, 2012), <http://www.nber.org/papers/w18237> (last accessed April 24, 2017).

⁷ Elizabeth A. Harris and Ford Fessenden, “‘Opt Out’ Becomes Anti-Test Rallying Cry in New York State,” *New York Times* May 20, 2015, https://www.nytimes.com/2015/05/21/nyregion/opt-out-movement-against-common-core-testing-grows-in-new-york-state.html?_r=0. For a look at parent-led grassroots opt-out efforts in Florida, a state much more beholden to high-stakes standardized testing programs, see, Kristina Rizga, “Sorry, I’m Not Taking This Test,” *Mother Jones* September/October, 2015, <http://www.motherjones.com/politics/2015/08/opt-out-standardized-testing-overload> (links last accessed April 29, 2017).

⁸ Kate Zernike, “Massachusetts’s Rejection of Common Core Test Signals Shift in U.S.,” *New York Times* November 21, 2015, https://www.nytimes.com/2015/11/22/us/rejecting-test-massachusetts-shifts-its-model.html?_r=0. The anti-testing movement has also provided a way for teachers unions to dredge up support in otherwise

this “opt out” movement reflects the distrust many Americans hold for the state (however abstractly or imaginatively defined) and its data-gathering impulse.⁹ Within this capacious political mindset, standardized testing regimes are, at once, an assault on childhood, an imposition of anti-intuitive learning patterns, a devaluation of genuine inquiry and knowledge, an abandonment of traditional educational basics, a form of social control, the triumph of neoliberal educational corporatization, a waste of taxpayers’ dollars, and the reason the United States will fall behind one of any number of countries that either, ironically, use tremendous high-stakes testing mechanisms (China) or barely employ standardized educational testing at all (Finland).¹⁰

unpopular environments, using an issue that effects their livelihood (the punitive use of test scores in annual evaluations) and parents’ fears (an unjust labeling of their child as abnormal or deficient, especially as studies indicated teachers stuck in high-stakes testing environments spent more effort on marginally-proficient students than low-performing pupils); see: Kate Taylor and Motoko Rich, “Teachers’ Unions Fight Standardized Testing, and Find Diverse Allies,” *New York Times* April 20, 2015, https://www.nytimes.com/2015/04/21/education/teachers-unions-reasserting-themselves-with-push-against-standardized-testing.html?_r=0; and Kathleen McGrory, “Gov. Rick Scott Signs Bill That Scales Back Testing,” *Tampa Bay Times* April 14, 2015, <http://www.tampabay.com/news/politics/stateroundup/gov-rick-scott-signs-bill-that-scales-back-testing/2225436>; Chicago Teachers Union, “Debunking the Myths of Standardized Testing: A CTU Position Paper,” CTU website, http://www.ctunet.com/quest-center/research/position-papers/text/CTU_Testing_Position_Brief_web-1.pdf (all links last accessed April 24, 2017); and Derek Neal and Diane Whitmore Schanzenbach, “Left Behind by Design: Proficiency Counts and Test-Based Accountability,” *Review of Economics and Statistics* 92, no. 2 (May, 2010): 263-283.

⁹ Observe, for example, the relative popularity of anti-Common Core Facebook Groups. One can join “CHOOSE TO REFUSE COMMON CORE” (38,178 as of late April, 2017), “AMERICANS AGAINST COMMON CORE” (2,177 members), “Don’t Be Cattle! Fight Common Core!” (7,430)—or, based on state and local allegiances, one can join similar groups united against Common Core in Utah, Rhode Island, Ohio, and East Islip, New York (4,448, 1,606, 2,817, and 485 members, respectively.) In these groups, Common Core is linked to evils of progressive politics, Islam, socialism, charter schooling, big government, and transgenderism under the matrix of “indoctrination.” In turn, anti-Common Core activists—whether or not they oppose actual standardized testing based on Common Core standards or just a shapeless “Common Core” bogeyman—pose themselves as the defenders of the purity of children’s minds and spirits. While state- and local-focused groups seem more opposed to the ideas of needless educational bureaucracy and corporatism, national groups seem overtly aligned with far-right politics; Facebook’s algorithms suggest Trump-fawning pages based on the network-joining patterns for these general anti-Common Core groups.

¹⁰ For brief considerations of whether teaching to the test actually dilutes learning through teachers’ disproportionate focus on content most likely to appear on a high-stakes exam—thus creating test inflation, or the false impression that students have learned more overall rather than select discrete points of content—see: Henry Braun, “Taming Inflation is Never Easy,” *Measurement* 13 (2015): 31-34; and Jennifer L. Jennings and Jonathan Marc Berek, “‘Teaching to the Test’ in the NCLB Era: How Test Predictability Affects Our Understanding of Student Performance,” *Educational Researcher* 43, no. 8: 381-389.

For widespread public commentary against teaching to the test, see: Kelly Gallagher, Commentary: Why I Will Not Teach to the Test,” *Education Week*, November 12, 2010, http://www.edweek.org/ew/articles/2010/11/17/12gallagher_ep_h30.html; Times Editorial Board, “No Child Left Behind: How to End ‘Teaching to the Test,’” *Los*

These parents' anti-testing politics are steeped in the fears and fantasies about the direction of American society, the United States' standing in relation to other countries, and most importantly, the future their own children will face. Opting out is not seen primarily as a retreat from society but a means of self-preservation.

Nonetheless, standardized testing endures, namely because it has served as a vital tool for the post-1945 American state. The postwar state prioritized, on the one hand, the cultivation of intellects resilient enough to sustain American geopolitical supremacy through scientific discovery and technological innovation and, on the other hand, the maintenance of an obedient population that would not disrupt existing social hierarchies. Standardized testing helped the postwar state solve this mind-body dilemma. As a function for social order, standardized testing provided the means for governing bodies to make sense of their citizens—particularly to gauge the skills, knowledge, and ability youth could eventually bring into a labor force. Standardized testing thus makes it clear not just who the so-called best and brightest are, but how well a population of students has adjusted to a set of educational norms the state has deemed necessary for future social success. Standardized testing is thus a way to monitor and enforce educational compliance with projected state and social needs.

Angeles Times, February 23, 2015, <http://www.latimes.com/opinion/editorials/la-ed-testing-no-child-left-behind-20150223-story.html>; and Vince Bertram, "Standardized Testing Review: Let's Make 'Teaching to the Test' a Thing of the Past," *Fox New Opinion*, November 16, 2015, <http://www.foxnews.com/opinion/2015/11/16/standardized-testing-review-lets-make-teaching-to-test-thing-past.html>; (links last accessed April 24, 2017).

For a sense of American education writers' fascination with Finland and its nearly-complete lack of standardized testing, see: Anu Partanen "What Americans Keep Ignoring About Finland's School Success," *The Atlantic*, December 29, 2011, <https://www.theatlantic.com/national/archive/2011/12/what-americans-keep-ignoring-about-finlands-school-success/250564/>; Pasi Sahlberg, *Finnish Lessons 2.0: What Can the World Learn from Educational Change in Finland?* (New York: Teachers College Press, 2014); and William Doyle, "Op-Ed: Why Finland Has the Best School," *Los Angeles Times*, March 18, 2016, <http://www.latimes.com/opinion/op-ed/la-oe-0318-doyle-finnish-schools-20160318-story.html> (links last accessed April 26, 2017).

This dissertation examines how standardized testing became a vital instrument for the postwar state—and, in turn, how this state dependence on standardized testing fueled several larger postwar political cultures. Beginning with World War Two, the American state began engaging with widespread standardized testing in a sustained fashion, with an eye toward broader geopolitical needs.¹¹ The state’s ability to continue large-scale testing depended on its use of independent test manufacturers, which developed standardized exam programs for the military, corporations, and institutions of higher education alike. Testing programs provided a convenient way to harmonize geopolitical, business, and educational interests without breaching commonly agreed-upon tenets of liberalism, capitalism, or decentralization.

This dissertation also focuses on the tension between the ways the state made sense of its citizens and the ways citizens made sense of society through standardized tests. Standardized testing acted as a massive social sorter in postwar America, but with time, many groups of Americans began to question the wisdom of educational reliance on standardized testing, as well as challenge the foundational assumptions about what testing actually measured. African-

¹¹ By “state,” I mean the institutions and actors that give the federal government its shape, scope, and power. This not only includes the executive, legislative, and judicial branches of the federal government, but also the military, governmental bureaucracy, federal agencies, and their employees. The state also includes the policies, rituals, and laws that compel social order among a certain people in a specified territory. What I aim to convey through this dissertation is that the state is buttressed by non-state actors that provide vital technologies for the state—namely, technologies that facilitate social order—but which are not themselves officially part of the state. Hence, standardized testing is used by the state to make meaning of its citizens in a way that the state, by its own devices, cannot. See: Theda Skocpol, “Bringing the State Back In: Strategies of Analysis in Current Research,” in *Bringing the State Back In*, eds. Peter B. Evans, Dietrich Reuschemeyer, and Theda Skocpol (Cambridge: Cambridge University Press, 1985), 3-37; Patrick J. Akard, “Bringing the Economy Back in (Again): Conceptions of the Capitalist State and Their Relevance for Public Policy,” *Mid-American Review of Sociology* 11, no. 2 (Winter, 1986): 75-100; Theda Skocpol, “A Society Without a ‘State’? Political Organization, Social Conflict, and Welfare Provision in the United States,” *Journal of Public Policy* 7, no. 4 (October-December, 1987): 349-371; Desmond King and Robert C. Lieberman, “Finding the American State: Transcending the ‘Statelessness’ Account,” *Polity* 40, no. 3 (July, 2008): 368-378; Eric M. Patashnik and Julian E. Zelizer, “The Struggle to Remake Politics: Liberal Reform and the Limits of Policy Feedback in the Contemporary American State,” *Persepectives on Politics* 11, no. 4 (December, 2013): 1071-1087; Daniel Béland, François Vergnoille de Chantal, and Sarah-Louise Raillard, “The American State: Between Political Invisibility and Institutional Fragmentation,” *Revue Française de Science Politique* 64, no. 2 (2014): 1-14; and Gillian E. Metzger, “Agencies, Polarization, and the States,” *Columbia Law Review* 115, no. 7 (November, 2015): 1739-1787.

Americans, women, and working-class Americans increasingly resisted the way standardized testing was used as both proof of their diminished capabilities and justification for their social disparity. These groups pursued legal, legislative, and academic methods to push back against unjust standardized testing practices. This resistance also often provoked responses from those who had great stakes—whether socioeconomic or corporate—in maintaining the use and meaning of standardized tests. Politicians, parents, consumer advocates, academics, educational reformers, feminists, civil rights activists, entrepreneurs, and white supremacists all interpreted standardized testing scores and trends for larger political ends, often using similar information as their opponents to stake far different positions about race, gender, class, and merit. Yet, as more Americans questioned the legitimacy of standardized testing, they often reinforced standardized testing data's use as a rhetorical tool, ultimately entrenching testing data as a way to make sense of society, even as more and more Americans find testing regimes purposeless.

Finally, this dissertation shows that, by the beginning of the twenty-first century, standardized testing data had become a potent tool for contending who mattered and who did not for America's future. Because debates about standardized testing's social utility often hinged on the meaning of test data—whether it reflected objective truths about the natural distribution of mental aptitudes or, instead, exposed the biases psychologists built into their devices as well as the prejudicial baggage that informed laypeople's interpretation of statistical information—standardized testing itself continues to have value as a political weapon. Whether or not one has a specific policy proposal for the future use of standardized testing, its rhetorical function as a symbol for what is wrong with America will continue to fuel numerous, often oppositional, debates. Until we get rid of standardized testing, we are stuck with political cultures fueled by standardized testing.

Ironically, as much writing exists on the perils and promises of standardized testing, there remains relatively little historical research on the topic. The touchstone for the history of standardized testing remains Nicholas Lemann's *The Big Test*, which not only critically examined the development of American higher education's dependence on the SAT, but also the test's role in fueling the myth of meritocracy—that is, the ideological belief that the most qualified and talented earn a space at the top of society through natural talent and hard work and not profound structural advantages. Meritocracy let prestigious institutions offer the semblance of a path to social comfort without ever acknowledging the broader social forces that kept minorities outside of elite educational environments and, by extension, outside social privilege. The SAT added a technocratic sheen to meritocracy's bootstrapping rhetoric. Those who did well on the SAT—namely, middle-to-upper-class white men in well-funded schools—could attribute their place in prestigious colleges to the fuzzy idea of merit rather than ever question if the SAT was a device better suited for maintaining the existing composition of the social elite. Lemann's work spurred considerable scrutiny and reform among American admissions counselors, and added weight to public displeasure with the college application process, but his work lets the SAT stand in for a far broader system of standardized assessment. By focusing on the attitudes and backgrounds of the men who ran Educational Testing Service, however, Lemann forecloses on the primary role the postwar state played in making standardized testing so commonplace a social sorting device. Standardized testing was not just a tool for technocrats to replicate the Northeastern elite among the masses, but also a way for the state to use new technology for reproducing existing racial and economic power dynamics.¹²

¹² Nicholas Lemann, *The Big Test: The Secret History of the American Meritocracy* (New York, Farrar, Straus, and Giroux, 2000). To see Lemann's lasting influence on shifting some colleges away from SAT-heavy admissions decisions, see: National Association for College Admission Counseling, "Report of the Commission on the Use of Standardized Tests in Undergraduate Admission," September, 2008, <http://files.eric.ed.gov/fulltext/>

Although other useful works exist on the history of standardized testing—in particular, William Reese’s investigation of the origins of testing in mid-nineteenth-century Boston schools—most research on the topic either has a decidedly journalistic bent or is the product of internal corporate histories. While this allows some authors to give compelling portrayals of test-preparation entrepreneurs (as David Owen did with the founders of The Princeton Review in *None of the Above*), or to depict the actual children harmed by harebrained state testing policies (as Anya Kamenetz did in *The Test*), such work is usually pegged to immediate public concerns with standardized testing. Such work often has a deeply foreshortened sense of testing’s history, as well as a clear aim to castigate standardized testing wholesale. By contrast, histories created by standardized testing companies are flush with bureaucratic detail, but devoid of any genuine critical examination of standardized testing’s impact on postwar American society. In these corporate histories, standardized testing is always portrayed as a social good that contributes to an undefined but ever-present idea of Progress. Standardized testing does not contribute to socioeconomic disparities, these corporate histories maintain, but merely reflects the unequal educational experiences that create test score differences between different socioeconomic groups: don’t shoot the messenger. Both journalistic and corporate approaches to this topic avoid a thorough examination of the role standardized testing played in expanding the reach of the postwar state, as well as the intergenerational effect standardized testing data has had on various groups’ socioeconomic mobility.¹³

[ED502721.pdf](#); Andrew S. Belasco, Kelly O. Rosinger, and James C. Hearn, “The Test-Optional Movement at America’s Selective Liberal Arts Colleges: A Boom for Equity of Something Else?” *Education Evaluation and Policy Analysis* 37, no. 2 (June, 2015): 206-223; and Sarah Kaplan, “The Rise and Demise of the Mich-Loathed SAT,” *Washington Post*, July 28, 2015, https://www.washingtonpost.com/news/morning-mix/wp/2015/07/28/how-the-sat-came-to-rule-college-admissions/?utm_term=.aed7af817fd8 (links last accessed April 27, 2017).

¹³ William J. Reese, *Testing Wars in the Public Schools: A Forgotten History* (Cambridge: Harvard University Press, 2013); Anya Kamenetz, *The Test: Why Our Schools are Obsessed with Standardized Testing—But You Don’t Have to Be* (New York: PublicAffairs, 2015); and David Owen, *None of the Above: Behind the Myth of*

The history of standardized testing makes more sense when placed in a broader history of numeracy, quantification, and social science research. Over the course of four centuries, scientists have engaged in the process of relating humans to abstract quantities, whether by surveying communities, forecasting voters' choices, imagining the behavioral patterns of rational actors, or constructing ideal physical shapes. Quantified abstractions mark the division between the pre-modern and modern worlds: they signify when people became populations, when the value of a life became an actuarial sum, when enslaved humans became units for insurance adjustment, and when legal systems encoded fractional bases for race and nativity. The modern state could hold better control over its subjects by treating them as economic figures—and, in turn, could accustom populations to discipline themselves in line with statistical norms. Modernity, then, has involved the confluence of governmental control, economic constraints, and corporeal self-regulation through methods that render humans into numerical data.¹⁴

Scholastic Aptitude (New York: Houghton Mifflin, 1986). For another worthwhile examination of the long history of standardized testing, see: Mark J. Garrison, *A Measure of Failure: The Political Origins of Standardized Testing* (Albany: SUNY Press, 2009). For keen work on the spread of high-stakes standardized testing to Canada, see: Arlo Kempf, *The Pedagogy of Standardized Testing: The Radical Impacts of Educational Standardization in the US and Canada* (New York: Palgrave MacMillan, 2016).

For examples of histories produced internally by test designers see: Michael C. Johanek, ed., *A Faithful Mirror: Reflections on the College Board and Education in America* (New York: College Board, 2001); Manuel Maldonado-Rivera, *The History of the Puerto Rico and Latin American Office*, edited by Michael Farah, <http://about.collegeboard.org/region-offices/puerto-rico> (last accessed April 27, 2014); George H. Hanford, *Life with the SAT: Assessing Our Young People and Our Times* (New York: College Entrance Examination Board, 1991); John A. Valentine, *the College Board and the School Curriculum: A History of the College Board's Influence on the Substance and Standards of American Education, 1900-1980* (New York: College Entrance Examination Board, 1987); and Frank Bowles, *The Refounding of the College Board, 1948-1963: An Informal Commentary and Selected Papers* (New York: College Entrance Examination Board, 1967).

¹⁴ For a sampling of these ideas, see: Ian Hacking, "Biopower and the Avalanche of Printed Numbers," *Humanities in Society* 5 (1982): 279-95; Peter Miller and Ted O'Leary, "Accounting and the Construction of the Governable Person," *Accounting, Organizations and Society* 12, no. 3(1987): 235-265; Joan Wallach Scott, "A Statistical Representation of Work: *La Statistique De L'Industrie à Paris, 1847-1848*," in *Gender and the Politics of History* (New York: Columbia University Press, 1988); Susan Leigh Star and James R. Griesemer, "Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Berkeley's Museum of Vertebrate Zoology, 1907-39," *Social Studies of Science* 19, no. 3 (august, 1989): 387-420; Patricia Cline Cohen, *A Calculating People: The Spread of Numeracy in Early America* (New York: Routledge, 1999); Ian Baucom, *Specters of the Atlantic: Finance Capital, Slavery, and the Philosophy of History* (Durham: Duke University Press, 2005); Ian Hacking, "Making Up People," *London Review of Books*, 28, no. 16 (August 17, 2006), <https://www.lrb.co.uk/v28/n16/ian->

Standardization has been one crucial facet of quantification—and the early history of standardized testing, as discussed below, involved the incremental development of tools that allowed the American state to make sense of its citizens and, in turn, for Americans to make sense of themselves in numerical terms.

The history of standardized testing also benefits from sociological examinations of the mechanics of class formation—both the actual creation of student bodies and the reinforcement of socioeconomic strata—through postsecondary institutions, particularly the research of Jerome Karabel and Mitchell Stevens. Whereas Karabel illustrates the historical dimensions of elite replication at Harvard, Yale, and Princeton, Stevens examines the contemporary admissions process at a small liberal arts college. Karabel and Stevens show that elite colleges are not just producers of the social elite, but respondents to broader economic forces. Moreover, Karabel and Stevens illustrate how both prestigious and less-prestigious educational institutions appeal to select groups of students based on a sense of their own place in the academic pecking order: an small private college or large public research university, cannot market itself in the same way as Ivy League schools do, quite simply because they lack the same clout—a clout increasingly based on quantified guidelines.¹⁵

What, Exactly, Is a Standardized Test?

One irony of the widespread use and criticism of standardized tests is that relatively few people are clear with what, exactly, standardized tests are. In this sense, standardized testing

[hacking/making-up-people](#) (last accessed April 28, 2017); Sarah E. Igo, *The Averaged American: Surveys, Citizens, and the Making of a Mass Public* (Cambridge: Harvard University Press, 2007); Martha Lampland, “False Numbers as Formalizing Practices,” *Social Studies of Science* 40, no. 3 (June, 2010): 377-404; Andrea Mennicken and Peter Miller, “Accounting, Territorialization and Power,” *Foucault Studies* 13 (May, 2012): 4-24; and Caroline Lambert and Eric Pezet, “Accounting and the Making of *Homo Liberalis*,” *Foucault Studies* 13 (May, 2012): 67-81.

¹⁵ Jerome Karabel, *The Chosen: The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton* (Boston: Houghton Mifflin, 2005); and Mitchell L. Stevens, *Creating a Class: College Admissions and the Education of Elites* (Cambridge: Harvard University Press, 2007).

becomes pornographic: we can't exactly define it, but we know it when we see it. A standardized test is, quite simply, an assessment that is designed and administered in a uniform fashion, from how much time test-takers are given, the questions they receive, the tools they may use, and even the instructions they are provided.¹⁶ Although standardized testing may carry high stakes—such as students' promotion from one grade to the next, teachers' pay rates and employment, or schools' funding and operations—not all standardized tests carry such severe consequences. Nor do all standardized tests have the same primary purpose: some gauge the degree to which test-takers have acquired skills up to a certain point, while others assess test-takers' learning strengths or predict readiness for further education. Nonetheless, a standardized assessment's shape and questions first depend upon its stakeholders agreeing upon what its main objectives should be.¹⁷ Based on these objectives, test designers—usually private corporations—create a considerable number of test items, far more than is needed for the actual final exam, and give them a trial run on sample populations. These experimental sample groups are ideally representative of who will be taking the final exam—e.g., a test designed to measure high-school seniors' math skills should not use sixth-grade students for a trial sample. Given how well or poorly these sample groups fared, test designers cull the best-constructed questions and assemble them in a final experimental version. The data from this final experimental version often forms the foundation for the published test's scoring standards.¹⁸ Standardized exams thus go through an elaborate multi-stage process to hone how they will assess test-takers' skills and knowledge.

¹⁶ Educational Testing Service, "Glossary of Standardized Testing Terms," ETS website, https://www.ets.org/understanding_testing/glossary (last accessed April 12, 2017).

¹⁷ As Wayne Au observes, the decision for a standardized test to carry high stakes is above all a policy decision, one that has historically been disproportionately detrimental to minority communities; see: Wayne Au, *Unequal By Design: High-Stakes Testing and the Standardization of Inequality* (New York: Routledge, 2009).

¹⁸ Craig A. Mertler, *Interpreting Standardized Test Scores* (New York: SAGE Publications, 2007); Kimberly O'Malley, "Standardized Testing: What Is It and How Does It Work?" Pearson Research Blog, May 17,

Contemporary standardized tests fall into two broad categories, which each offer considerably different models for adequate and extraordinary behavior. In a norm-referenced test, test-takers' performances are compared to those of a pre-constructed sample group. Such sample groups are created during a test's experimental period, although larger-scale standardized testing programs (such as the SAT) may develop ongoing experimental design to ensure a cohort of test-takers is never disproportionately abnormal compared to the original sample group. Test-takers' scores on norm-referenced tests do not directly reflect how well or poorly they did on questions but, instead, their performance compared to the experimental sample group. A test-taker may answer only half of all questions correctly and still score in the 70th percentile on a norm-referenced test if they performed better than 70 percent of all students in the sample group. In a criterion-referenced test, by contrast, students are measured by how much they know about a specified domain of knowledge. The reference point for test-takers' scores is not an original sample group but, instead, predetermined objectives—what a high-school junior ought to know in English by the end of the school year, or fundamental nursing procedures. These test-takers are not in competition with a sample population, but against an imagined ideal test-taker who knows all of the content.¹⁹ Whatever the basis for reference, standardized test scores do two

2012, <https://www.pearson.com/corporate/news/blogs/research/2012/05/standardized-testing--what-is-it-and-how-does-it-work-.html>; Ruth Mitchell, "A Guide to Standardized Testing: The Nature of Assessment," Center for Public Education website, <http://www.centerforpubliceducation.org/Main-Menu/Evaluating-performance/A-guide-to-standardized-testing-The-nature-of-assessment>; Titania Kumeh, "Education: Standardized Tests, Explained," *Mother Jones*, March 25, 2011, <http://www.motherjones.com/mixed-media/2011/03/NCLB-standardized-tests-explained> (all links last accessed April 29, 2017).

¹⁹ James Popham and T.R. Husek, "Implications of Criterion-References Measurement," *Journal of Educational Measurement* 6, no. 1 (Spring, 1969): 1-9; Linda Crocker and Jeri Benson, "Achievement, Guessing, and Risk-Taking Behavior under Norm Referenced and Criterion Referenced Conditions," *American Educational Research Journal* 13, no. 3 (Summer, 1976): 207-215; Lorrie Shepherd, "Norm-References vs. Criterion-Referenced Tests," *Educational Horizons* 58, no. 1 (Fall, 1979): 26-32; Anthony J. Nitko, "Distinguishing the Many Varieties of Criterion-References Tests," *Review of Educational Research* 50, no. 3 (Autumn, 1980): 461-485; and "Criterion- and Standards-Referenced Tests," FairTest website, <http://www.fairtest.org/criterion-and-standards-referenced-tests> (last accessed April 29, 2017).

things at once: they render individual students into quantities suitable for comparison and, more potently, they personalize quantifiable data. Standardized test scores become a stand-in for capability, and the ineffable aspects of skill and achievement are replaced with numbers referent only to a test's measurement system.

Test standardization also involves the elimination of subjective—that is, human—judgment whenever possible, which allows proponents to more easily frame the argument that standardized tests are fair and reasonable devices. More specifically, standardization involves reducing the number of questions that require considerable time to assess, as fatigued graders are prone to making scoring variations or inserting their own biases. Most Americans commonly associate standardized tests with No. 2 pencils, pastel bubble sheets, and scanning machines. This technology, developed by Reynold Johnson in the early 1930s and honed by IBM soon after, allowed for hundreds of tests to be scored in an hour—making national annual standardized tests feasible in subsequent decades.²⁰ Some test manufacturers have reduced the need for human graders by building computer-based standardized exams. People who take a computer-adaptive test (CAT) are first presented a set of medium-difficulty questions. Based on how well they perform with this initial batch, test-takers are then presented with questions that the testing program has calculated to be moderately difficulty for that particular test-taker. As test-takers progress through a CAT, the weight for each cluster of questions typically decreases; how well or poorly one does on the last five questions often has much less impact on their final score than how well or poorly one does on the first five. CAT designers maintain the results on such exams provide a more accurate depiction of test-takers' range of ability than other versions of standardized testing can provide: high-ability test-takers are not drained of energy answering

²⁰ “Automated Test Scoring: Overview,” IBM 100 website, <http://www-03.ibm.com/ibm/history/ibm100/us/en/icons/testscore/> (last accessed April 29, 2017).

easy questions, and low-ability test-takers did not gain cheap points with lucky guesses on hard questions.²¹ In both paper-based and computer-based formats, however, human adjudication is pored to its requisite minimum.²²

A Brief History of Standardized Testing to the late 1930s

Although standardized testing is often seen as an educational practice, its origins extend well beyond the classroom. The early history of standardized testing is ensnarled in the histories of bureaucracy, intelligence, industrialization, urbanization, immigration, progressivism, and warfare. Each of these historical phenomena generated concepts and practices fundamental to contemporary standardized testing. Briefly considering how standardized testing developed before the Second World War illuminates how, above all, early test developers sought to manage social change. The initial uses for standardized testing reflect an underlying aim to maintain the structural integrity of social institutions all while shaping the direction in which different populations flowed into social systems. Whatever the anchoring ideology, standardized testing has been from its earliest versions a mechanism for social control and order.

From a bureaucratic perspective, standardized testing is nothing new. Indeed, standardized examinations were central to maintaining imperial China's vast civil service.²³ As

²¹ Mary E. Lunz and Betty A. Bergstrom, "The Effect of Review on Student Ability and Test Efficiency for Computerized Adaptive Tests," *Applied Psychological Measurement* 16, no. 1 (March, 1992): 33-40; Tim Davey, "A Guide to Computer Adaptive Testing Systems," Council of Chief State School Officers, 2011, <http://files.eric.ed.gov/fulltext/ED543317.pdf>; and National Council of State Boards of Nursing, "Computerized Adaptive Testing (CAT)," NCSBN website, <https://www.ncsbn.org/1216.htm> (links last accessed April 29, 2017).

²² Not all standardized testing is devoid of human grading. Most notably, the College Board's Advanced Placement exams employ a fleet of temporary graders (mostly educators), who work through a massive amount of AP exams in dozens of subjects at a rapid pace; see: Brittany Shoot: "Filling in the Blanks: The Thousands of Volunteers Who Grade Millions of AP Tests," *Pacific Standard Magazine*, July 9, 2014, <https://psmag.com/filling-in-the-blanks-the-thousands-of-volunteers-who-grade-millions-of-ap-tests-b1a232f6f18f>.

²³ Ichisada, Miyazaki, *China's Examination Hell: The Civil Service Examinations of Imperial China*, trans. Conrad Schirokauer (New York: John Weatherhill, 1976). For a sense of the current Chinese college admission exam system, the *gaokao*, and its relation to both the older imperial form and contemporary standardized college admission exams see: Yuan Feng, "National College Entrance Examinations: the Dynamics of Political Centralism

an American educational practice, however, standardized testing began in earnest during the mid-nineteenth century, as several northern states began to enact mandatory minimum schooling laws. The most notable early case of U.S. standardized testing took place in Boston, where Massachusetts Board of Education Secretary Horace Mann ended the practice of recitation-based examinations administered by visiting committees. Influenced by Whig politics, Prussian pedagogical practices, and phrenology, Mann enacted citywide written examinations in 1845. Rather than have cherry-picked students reply to individualized prompts, such tests required all pupils to answer identical series of questions in the same subjects within a set timeframe at a specific point in the year: standardization meant uniformity. Mann and his reformist allies used the results from these standardized exams to usher in structural reforms for Boston's schools, including the transition from one-room schoolhouses to graded classrooms segregated by students' age. Although written standardized exams would gain widespread use throughout many populous American school districts during the late-nineteenth century, debates over whether to use standardized exams would continue to fuel broader educational arguments over the relationship between states' power and local control.²⁴

in China's Elite Education," *Journal of Education* 181, no. 1 (1999): 39-57; Emily Hannum, Xuehui An, and Hua-Yu Sebastian Cherng, "Examinations and Educational Opportunity in China: Mobility and Bottlenecks for the Rural Poor," *Oxford Review of Education* 37, no. 2 (April, 2011): 267-305; Gareth Davey, Chuan De Lian, and Louise Higgins, "The University Entrance Examination System in China," *Journal of Further and Higher Education* 31, no. 4 (November 2007): 385-396; Brook Lerner, "Inside a Chinese Test-Prep Factory," *New York Times Magazine*, December 31, 2015; and Terry Crawford, "What Students in China Have Taught Me About U.S. College Admissions," *The Atlantic* 6 January, 2015, <http://www.theatlantic.com/education/archive/2015/01/what-students-in-china-have-taught-me-about-us-college-admissions/384212/>.

²⁴ Mark Garrison, *A Measure of Failure: The Political Origins of Standardized Testing* (Albany: SUNY Press, 2009), 59-72; George F. Madaus and Thomas Kellaghan, "Testing as a Mechanism of Public Policy: A Brief History and Description," *Measurement & Evaluation in Counseling & Development* 26, no. 1 (April, 1993); George F. Madaus and Laura M. O'Dwyer, "A Short History of Performance Assessment: Lessons Learned," *Phi Delta Kappan* 80, no. 9 (May, 1999): 688-695; and William J. Reese, *Testing Wars in the Public Schools: A Forgotten History* (Cambridge: Harvard University Press, 2013).

Mann and other early proponents of standardized testing built their arguments on the foundational tenets of nineteenth-century social science, particularly the idea that humanity could be measured, generalized, and discussed in quantified terms. Belgian statistician Adolphe Quetelet's work with normal curves provided a foundation for quantified knowledge about human characteristics. Quetelet observed that people's physical dimensions fell along a normal curve: some had rather small features, other rather large, but most clustered around an average equally distant from both extremes. From these observations, Quetelet devised the ideal "average man"—*l'homme moyen*-- and posited that any individual's deviation from that imagined average could be represented as a probability.²⁵ Quetelet's *homme moyen* not only naturalized abstract ideas about the body, but also served as a potent metaphor for the social body. Quantifiable information about individual bodies, human capacity, and the larger social body served states' political, moral, and economic needs. Such information fueled state power systems that exacted social and corporeal discipline through knowledge production, a process Michel Foucault would later deem *biopolitics*.²⁶ Standardized testing's scoring systems—often built to fit normal curves and designed to quantify abstract information about human capability—are one legacy of these earlier biopolitical developments.

²⁵ Walt Haney, "Testing Reasoning and Reasoning About Testing," *Review of Educational Research* 54, no. 4 (Winter, 1984): 599.

²⁶ Brian P. Cooper, *Family Fictions and Family Facts: Harriet Martineau, Adolphe Quetelet, and the Population Question in England, 1798-1859* (London: Routledge, 2007): 197-244; Nancy Krieger, "Who and What is a 'Population'? Historical Debates, Current Controversies and Implications for Understanding 'Population Health' and Rectifying Health Inequities," *Milbank Quarterly* 90, no. 4 (December, 2012): 634-681; and Jean-Guy Prévost and Jean-Pierre Beaud, *Statistics, Public Debate and the State, 1800-1945: A Social, Political and Intellectual History of Numbers* (London: Pickering and Chatto, 2012).

For a better sense of the relationship between statistics and biopolitics, particularly how the relationship made more nuanced through the work of French theorist Alain Desrosières, see: Keith Hoskin, "The Examination, Disciplinary Power and Rational Schooling," *History of Education* 8, no. 2 (1979): 135-146; Keith W. Hoskin and Richard H. Macve, "Accounting and the Examination: A Genealogy of Disciplinary Power," *Accounting, Organizations and Society* 11, no. 2 (1986): 105-136; Kathy Bussert-Webb, "The Teacher's Testing Panopticon," *Teaching & Learning* 18, no. 3 (Summer, 2004): 98-122; and Ann Rudinow Sætnan, Heidi Mork Lomell, and Svein Hammer, *The Mutual Construction of Statistics and Society* (New York: Routledge, 2011)

During the second half of the nineteenth century, many social scientists became enchanted by developments in biological sciences, and shifted their focus from measuring social bodies to improving them. What emerged from this shift was eugenics, a scientific ideology premised on the belief that humanity could be qualitatively improved by encouraging certain groups to reproduce more and preventing other groups from reproducing at all. British polymath Francis Galton was at the forefront of nineteenth-century eugenics. Inspired by Quetelet's work with normal curves, Galton developed tests to gauge mental faculties relative to an imagined norm. Galton was also inspired by the research of biologist (and cousin) Charles Darwin, but feared that natural selection, if left to its own devices, would not create lasting evolution. Observing the seed size pattern for several generations of pea plants, Galton discovered a phenomenon he deemed "regression to the mean": pea plants that developed from large seeds tended to produce seeds closer to the average seed size of the progenitor plant. Galton presumed that humanity, like pea plants, could lose superlative qualities just as easily—and that the social elite would bear middling sorts without any intervention.²⁷

The mental test was one eugenicist tool psychologists designed to correct this observed course of nature. Researchers honed mental testing practices during the early twentieth century, standardizing numerous methods for assessment and diagnosis. By studying the faculties of French youth, Alfred Binet and Theodore Simon developed the concept of mental age. Binet and Simon grouped children by how well they performed relative to the average performance of children at specific ages on a series of increasingly difficult cognitive tests. A twelve-year old,

²⁷ Joachim Kunert, Astrid Montag, and Sigrid Pöhlmann, "The Quincunx: History and Mathematics," *Statistical Papers* 42 (2001): 143-169; Chris Renwick, "From Political Economy to Sociology: Francis Galton and the Social-Scientific Origins of Eugenics," *British Society for the History of Science* 44, no. 3 (September, 2011): 343-369; Nicholas W. Gillham, "The Battle Between the Biometricians and the Mendelians: How Sir Francis Galton's Work Caused his Disciples to Reach Conflicting Conclusions About the Hereditary Mechanism," *Science and Education* 24 (2015): 62-64; and Sherrin Berezowsky, "Statistical Criticism and the Eminent Man in Francis Galton's *Hereditary Genius*," *Victorian Literature and Culture* 43 (2015): 821-839.

for example, could have a mental age of seven if her performance on these tests was equivalent to the degree to which most seven-year olds would successfully complete them.²⁸ The intelligence scale that developed from these studies—conveniently named the Binet-Simon scale—became the foundation for American psychologists’ sense of intelligence measurement.

The Binet-Simon Scale did not gain its cultural heft, however, until Stanford psychologist Lewis Terman made several alterations. Most substantially, Terman retooled the way the Binet-Simon scale expressed intelligence levels, changing the expression to a ratio between an individual’s mental age and their chronological age. Terman deemed this expression the intelligence quotient, or IQ. Terman’s IQ soon became the psychological and cultural standard for making sense of intelligence.²⁹ The appeal of IQ, as Paula Fass argues, resides in its ability to personalize an abstract concept. Because IQ expresses intelligence as a ratio, it hides the abstract qualities that form the concept of mental age. The relational aspect of IQ rests on the specious premise that physical age has as some direct, linear relationship to normal mental development.³⁰

Because Terman devised norms for IQ by measuring the capabilities of native-born middle-class white children, the intelligence quotient also masked the standards against which ethnic minorities were measured. Through psychological slight-of-hand, IQ recast the economic

²⁸ Corwin Boake, “From the Binet-Simon to the Wechsler-Bellvue: Tracing the History of Intelligence Testing,” *Journal of Clinical and Experimental Neuropsychology* 24, no. 3 (2002): 385-6; John Carson, “Mental Testing in the Early Twentieth Century: Internationalizing the Mental Testing Story,” *History of Psychology* 17, no. 3 (2014): 249-255; For a sense of how Binet’s testing developed to curb the influence of psychiatry, as well as concurrent developments in assessing developmental disabilities, see: Serge Nicolas *et. al.*, “Sick? Or Slow? On the origins of intelligence as a Psychological Object,” *Intelligence* 41 (2013): 699-711; Annette Mülberger, “The Need for Contextual Approaches to the History of Mental Testing,” *History of Psychology* 17, no. 3 (2014): 177-186; and Elisabetta Ciccioia, Renato Foschi, and Giovanni Pietro Lombardo, “Making Up Intelligence Scales: De Sanctis’s and Binet’s Tests, 1905 and After,” *History of Psychology* 17, no. 3 (2014): 223-236..

²⁹ Boake, 388.

³⁰ Paula S. Fass, “The IQ: A Cultural and Historical Framework,” *American Journal of Education* 88, no. 4 (August, 1980): 431-458.

and social advantages of whiteness as the conditions for normal intelligence. The logic of IQ positioned poverty and race-based disparities as social abnormalities—and IQ became, in turn, evidence for the inherent inferiority of African-Americans, immigrant communities, and poor whites. Although intelligence scales would sustain substantial revisions and criticisms throughout the twentieth century, IQ retained cultural power because it posed abstract concepts as natural skill, and anchored relational measurements in individuals' abilities.³¹

As Terman developed a new scale for expressing intelligence, Robert Yerkes invented a method now synonymous with popular depictions of standardized testing: multiple-choice questions. The psychologist built multiple choice apparatuses, partitioned wooden devices with brass keys wired to lamps and buzzers. Yerkes and his assistants had subjects determine which key was connected to a buzzer, rotating switches to ensure each experimental trial used a different key. After a person successfully determined which key had been wired, the researchers asked the subject to articulate how they figured out which key to press, scoring for both the speed at which they came to the right answer and the way in which they described their choices. Yerkes used these contraptions on numerous species in order to measure ideational behavior, or the “series of reactive tendencies that a person displays when confronted with solving a

³¹ For a sense of how IQ became a tool for various jurisdictions to control the livelihood of certain individuals—either on the basis of their disproportionate burden to society or their perceived menace to white society, both threats premised on low intelligence scores—see: Robert L. Hayman, Jr., “Presumptions of Justice: Law, Politics, and the Mentally Retarded Parent,” *Harvard Law Review* 103, no. 6 (April, 1990): 1201-1271; James W. Trent, *Inventing the Feeble Mind: A History of Mental Retardation in the United States* (Berkeley: University of California Press, 1994), 131-183; Benjamin Beit-Hallahmi, “Science, Ideology, and Ideals: The Social History of IQ Testing,” *Centennial Review* 38, no. 2 (Spring, 1994): 341-360; William H. Tucker, “Places, Everyone! IQ Heritability, Ideology, and Education,” *Journal of Educational Thought* 32, no. 3 (December, 1998): 217-240; Steven Selden, *Inheriting Shame: The Story of Eugenics and Racism in America* (New York: Teachers College Press, 1999); Wendi Kline, *Building a Better Race: Gender, Sexuality, and Eugenics from the Turn of the Century to the Baby Boom* (Berkeley: University of California Press, 2001); and Gerald V. O’Brien, *Framing the Moron: The Social Construction of Feeble-Mindedness in the American Eugenic Era* (Manchester: Manchester University Press, 2013).

problem.”³² To Yerkes, ideational behavior was a measurable action rooted in intelligence—and although intelligence itself could not be directly measured, the behaviors that reflect intelligent decisions could. While multiple-choice devices fell out of favor after World War One, the multiple choice method held experimental currency among researchers for quite some time.

Psychologists’ fondness for standardized testing and assessment mirrored late-nineteenth- and early-twentieth-century shifts in American capitalism. As the United States industrialized, businesses began to standardize all aspects of operations: manufacturing, resources, and distribution all began to coalesce around agreed-upon terms. Standardization of rail track gauges, screw sizes, brick dimensions, and thousands of other materials not only lubricated infrastructural development, but also encouraged marketplace cohesion. Yet, as with the tensions that developed through the growth of standardized education testing, industrial capital’s standardization processes revealed a tension between local political power and national economic forces. What set the United States’ industrial growth apart from contemporaneous developments elsewhere, as D. Linda Garcia notes, was the private sector’s preponderant role in establishing bodies for standards development.³³ During the industrial era, standardization was never solely within the purview of the state.

This is not to say the era of industrial capitalism was devoid of any state influence. Rather, the late-nineteenth and early-twentieth-centuries marked substantial growth in the size and scope of U.S. federal governmental bureaucracy—as well as its increasing encroachment on

³² Shae A. Trewin, “Robert Yerkes’ Multiple-Choice Apparatus, 1913-1939,” *The American Journal of Psychology* 120, no. 4 (Winter, 2007): 645. For an example of the early aims to build upon Yerkes’s apparatus, see: Florence Emma Whittell, “Standardization of the Yerkes Multiple Choice Method for Human Adults,” master’s thesis, University of California, 1920. For contemporary developments in testing, see: John Levi Manahan, *A Bibliography of Educational Surveys and Tests* (Charlottesville, VA: The University of Virginia Press, 1916).

³³ D. Linda Garcia, “Standard Setting in the United States: Public and Private Sector Roles,” *Journal of the American Society for Information Science*, (1986-1998 series), 43, no. 8 (September, 1992): 531.

everyday life. The Progressive Era makes most sense, however, when viewed as a loosely bundled series of political movements emanating from the American middle class. White middle-class Progressives feared the drag both the elite and destitute alike placed on societal wellbeing. Through this fear, these Progressives elided being middle class with being righteous and normal—making the very wealthy and the poor statistically, politically, and morally deviant. This slippage between moral virtue, white supremacy, political soundness, and statistical preponderance guided Progressive reformers as they entered politics, seeking to reformat the relationship between society and state.³⁴

The Progressive mission was fueled by the fear that the white middle class—and, by extension, the American state—was at risk of corruption and devolution. Progressives grafted this fear of degeneration onto the language of intelligence, which allowed reformers to express nativist and white supremacist logic in scientific terms. Intelligence was one of many ways progressive reformers anchored the fantasy that American civilization was at risk from blacks, poor whites, and the foreign-born. For many Progressives, however, the most insidious social menaces were those whose intelligence was just below average. Idiots and imbeciles, as intelligence researchers and eugenicists deemed those with mental ages below eight years of age, posed less risk for white Americans' degeneracy than morons: those whose mental ages were deficient compared to the standard adult population, but who were able enough to operate in society undetected. Henry Herbert Goddard, who invented the condition of *moronia* in the early-

³⁴ Glenda Gilmore, *Gender and Jim Crow: Women and the Politics of White Supremacy in North Carolina* (Chapel Hill: UNC Press, 1992); Gail Bederman, *Manliness & Civilization: A Cultural History of Gender and Race in the United States, 1880-1917* (Chicago: University of Chicago Press, 1995); Michael McGerr, *A Fierce Discontent: The Rise and Fall of the Progressive Movement, 1870-1920* (New York: Free Press, 2003); Robert D. Johnston, *The Radical Middle Class: Populist Democracy and the Question of Capitalism in Progressive Era Portland* (Princeton: Princeton University Press, 2003); Greg Downs, *Declarations of Dependence: The Long Reconstruction of Popular Politics in the South, 1861-1908* (Chapel Hill: UNC Press, 2011).

twentieth century, insisted that morons were multigenerational threats to the white race. Because morons could not expand their mental capacity, the state wasted valuable resources on their education—and because feeble-mindedness was a heritable trait, normal whites risked diluting their own family's stock by mating with morons. Further, Goddard and other eugenicists used genetic logic to link feeble-mindedness with criminality, thereby turning moral shortcomings into medical symptoms. Any outward deviation of public social norms became proof of poor genetics, and poor genetics were diagnosable by observed by the lewd and licentious behavior of ancestors. Identifying and curbing feeble-mindedness—by legally sanctioning intelligence testing, institutionalization, and sterilization—was not just sound state policy but racial duty. Progressive logic turned standardization and intelligence testing into moral social obligations: white American vitality depended upon vigilantly monitoring for statistical abnormalities.³⁵

World War One profoundly expanded the scale of standardized mental testing as well as the stakes the American state placed in psychological testing results. Robert Yerkes, Lewis Terman, and dozens of other psychologists collaborated to devise the Army Alpha exam (for those literate in English) and Army Beta exam (for illiterate recruits). The United States Army administered these standardized tests to over 1.7 million recruits over the course of American involvement in World War One. The number of recruits who were discharged or sorted on the basis of these tests was quite small compared to the overall number of recruits. Despite the limited direct application of test scores, the Army Alpha and Army Beta tests showed that massive testing projects could be undertaken, and that test-takers could be categorically sorted on the basis of their results. Further, the state-orchestrated test-design project that created the Army

³⁵ Patrick J. Ryan, "Unnatural Selection: Intelligence Testing, Eugenics, and American Political Cultures," *Journal of Social History* 30, no. 3 (Spring, 1997): 669-685; Michael Willrich, "The Two Percent Solution: Eugenic Jurisprudence and the Socialization of American Law, 1900-1930," *Law and History Review* 16, no. 1 (Spring, 1998): 63-111; and Anna Stubblefield, "'Beyond the Pale': Tainted Whiteness, Cognitive Disability, and Eugenic Sterilization," *Hypatia* 22, no. 2 (Spring, 2007): 162-181.

Alpha and Army Beta exams united dozens of previously detached researchers, sparked a sense of national community among psychologists, and familiarized many Americans to psychology and psychological testing.³⁶

Although the results of the Army Alpha and Army Beta tests had limited immediate application, psychologists and politicians alike used data from recruits' exam scores to reinforce white supremacist logic. Black and white recruits who took the Army Alpha Test had, on average, a significant gap between results—a difference used by both researchers and laypersons as evidence of black mental inferiority. Yet, on closer examination, several factors clearly contradicted the white supremacist explanation for white recruits' superior test performance. Although white recruits had higher average performances on the Army Alpha Test, the median scores for recruits from northern states, whatever their race, were higher than their counterparts in southern states. Northern black recruits frequently performed better on the Army Alpha Test than southern whites. The relationship black and white recruits' scores also correlated strongly with the average amount of time blacks had in school relative to white students: the more time blacks spent in school, the narrower the gap. Rather than consider socioeconomic educational disparities as the root causes for the gap between blacks' and whites' Army Alpha test scores, leading psychologists and laypeople leaned on published intelligence test data to further naturalize racial difference.³⁷

³⁶ Haney, 604-608; Alexandra K. Wigdor and Wendell R. Garner, *Ability Testing: Uses, Consequences and Controversies*, vol. 1 (Washington, D.C.: National Academy Press, 1982): 81-90; Paul Davis Chapman, *Schools as Sorters: Lewis M. Terman, Applied Psychology, and the Intelligence Testing Movement 1890-1930* (New York: NYU Press, 1988): 65-83; Bernard R. Gifford, "The Political Economy of Testing and Opportunity Allocation," *The Journal of Negro Education* 59, no. 1 (Winter, 1990): 58-69.

³⁷ John L. Rury, "Race, Region, and Education: An Analysis of Black and White Scores on the 1917 Army Alpha Intelligence Test," *The Journal of Negro Education* 57, no. 1 (Winter, 1988): 51-65.

During the interwar period, standardized testing became part of the grammar of U.S. education.³⁸ Following the perceived success of the Army Alpha and Beta projects, IQ tests and other psychological assessments became increasingly popular instruments for measuring and sorting students. These standardized tests became instruments for managing the massive demographic changes in public schooling that had begun during the late-nineteenth century. Although immigration quotas essentially ended immigration from Southern Europe, Eastern Europe, and Asia by the mid-1920s, these multigenerational immigrant communities still placed tremendous stress on school districts—not only by the amount of resources required to teach millions of children, but also by the political and ideological battles over how to best incorporate these youth into American society. Meanwhile, child labor laws and minimum education provisions ensured that youth typically stayed in school longer than their counterparts from previous generations. The demands of industrial capital clashed with Progressive educational theory: How could schools provide equal education for all students, while also preparing certain

³⁸ See: David Tyack and William Tobin, “The ‘Grammar’ of Schooling: Why Has it Been so Hard to Change?” *American Education Research Journal* 31, no. 3 (Autumn, 1994): 453-479. Tyack and Tobin define the grammar of schooling as “the regular structures and rules that organize the work of instruction,” which, just as with linguistic grammar, do not “[need] to be consciously understood to operate smoothly” (454). Whether linguistic or institutional, grammars change over time—but at a rate slow enough to instill a sense of naturalness: this is how things ought to be. As Tyack and Tobin suggest, grammars are not rigid, but it is only through radical departures or changes that grammars become apparent. An individual’s particular dialect may spark interest, but the impulse to invoke “proper” grammar more frequently arises when a speaker breaches its rules. Similarly, local variations in education receive less scorn than drastic breaks in national standards.

This analogy, useful as it is, rests on two assumptions. First, one must presume grammars are power dynamics, between those who make grammars and those who use them. Because grammars do not always have rules explicitly stated at all times, the power of a grammar rests in its ability to seem like obvious common custom without much reminder of its invention. Second, this analogy also rests on the idea that decentralized education in largely rural areas—as was the case with expansion of common schools in the Northern and Western U.S. during the nineteenth century—relied on an organizing principle outside of the federal state or industrial capital. Tyack earlier argued that the expansion of public education during the nineteenth century should be seen primarily as a nation-building social movement; the industrial logic of scientific management and governmental bureaucratic oversight latched on later. Public education had a grammar, then, because nationalism had a grammar. See: John W. Meyer, David Tyack, Joane Nagel, and Audri Gordon, “Public Education as Nation-Building in America: Enrollments and Bureaucratization in the American States, 1870-1930,” *American Journal of Sociology* 85, no 3 (November, 1979): 591-613.

segments of students for different types of labor? Standardized testing helped educational bureaucrats broker this paradox by making it easier for schools to sort students into different paths—typically, vocational, normal, and college preparatory tracks. Standardized intelligence testing did not create educational tracking, just as it did not create nativism. But, these standardized psychological and intelligence tests provided a seemingly objective scientific logic that allowed many school districts to reinforce nativist sentiments and eugenic beliefs, all while satisfying industrial demands. If poor whites and the foreign-born scored poorly on standardized exams, it was a sign of natural shortcomings and not socioeconomic conditions, and so it was in national best interest to channel these youth into vocational educational tracks.³⁹

During the early-twentieth century, institutions of higher education also began turning toward standardized testing to address demographic pressures without democratizing the entire system. At the end of the nineteenth century, elite American colleges and universities each had their own particular set of admissions requirements. The materials students composed to apply for admission to Harvard, for example, had little overlap with the materials used to apply to Yale. Several of these prestigious schools formed the College Entrance Examination Board at the turn of the twentieth century to standardize their admissions process—including using a common admissions exam. For this exam, students composed essays in several subjects, which were then read and scored by College Board graders. These standardized essay exams were not originally intended to democratize access to higher education, however. Instead, this early version of collegiate standardized testing provided College Board member institutions a clearer standard to gauge students from different elite preparatory schools, to better ensure that college student bodies were not overrepresented by select feeder academies on the basis of tradition alone. This

³⁹ Chapman, 39-106; Jeannie Oakes, *Keeping Track: How Schools Structure Inequality* (New Haven: Yale University Press, 1985).

insularity continued when the College Board developed the Scholastic Aptitude Test—or SAT—in the late 1920s. Even though the original SAT more closely resembled contemporary standardized tests forms, many of the original 8,000 test takers were East Coast students from top-level preparatory schools applying for entrance into the most prestigious colleges. Not until the 1930s would elite colleges explicitly use the to draw in talented youth from distant Midwestern and Western locations. These early College Board exams broadened educational opportunity only for those who already had a great deal of it.⁴⁰

Although standardized testing became a far more popular educational practice after World War One, some public figures challenged the intelligence testing rhetoric espoused by Lewis Terman and other prominent psychologists. Most famously, the journalist Walter Lippmann questioned the tenets of intelligence testing in a series of essays for *The New Republic* in late 1922. Over six installments, Lippmann argued that intelligence testing, already a flawed instrument, was “in danger of gross perversion by muddleheaded and prejudiced men.”⁴¹ Lippmann accused psychologists of failing to clearly define intelligence, and for designing unjust, unreliable tests that were primarily devices for social sorting. The journalist worried dogmatic educations would “treat people with low intelligence quotients as congenitally and hopelessly inferior.”⁴² Indeed, Lipmann concluded his series by wishing the belief in hereditary intelligence would go “into that limbo where phrenology and characterology and the other Babu

⁴⁰ Michael S. Schudson, “Organizing the ‘Meritocracy’: A History of the College Entrance Examination Board,” *Harvard Educational Review* 42, no. 1 (February, 1972): 32-69; James Hocker Mason, “The Educational Milieu 1874-1911: College Entrance Requirements and the Shaping of Education,” *The English Journal* 68, no. 4 (April, 1979): 40-45.

⁴¹ Walter Lippmann, “The Mental Age of Americans,” *The New Republic*, October 25, 1922: 215.

⁴² Walter Lippmann, “The Abuse of the Tests,” *The New Republic*, November 15, 1922: 297

sciences are to be found.”⁴³ Lewis Terman did not hold Lippmann’s criticisms in high esteem, and rattled off several dismissive, sarcastic replies to *The New Republic*—dragging the public debate well into the following year. This friction, between what psychologists intended to measure with standardized tests and how nonspecialists understood standardized tests, would become the main force for public discourse on standardized testing in the United States in the century following World War One.

American psychologists continued to refine and reconstruct standardized psychological assessments throughout the interwar period. Most notably, the psychologist David Wechsler developed his namesake intelligence scales in 1939 while leading New York’s Bellevue Psychiatric Hospital. The Wechsler-Bellevue scale attempted to clarify adult intelligence measurement by reworking the formula for IQ. Rather than frame IQ as a ratio between mental age and chronological age—a model that, when misinterpreted, suggests either that people become less intelligent as they age, or that people’s intellectual capacity stops somewhere early in adolescence—Wechsler calculated IQ by using standard deviations. Using the Wechsler-Bellevue rubric, a psychologist could calculate a person’s IQ by measuring the distance between their cumulative test scores and the average score for people the same age. Because the Wechsler-Bellevue test measured for both verbal IQ and performance IQ, the exam also drastically altered psychological perceptions of intelligence.⁴⁴ Just as Lippmann’s essays set a foundation for decades of public criticism for intelligence testing, Wechsler’s scale provided a wedge for future psychologists to offer—and fight over—competing models for intelligence.

⁴³ Walter Lippmann, “A Future for the Tests,” *The New Republic* November 29, 1922: 10. Lippmann later tempered this fervor in a profile of British psychologist Cyril Burt, whose “common sense” approach to intelligence testing showed “there can be little or no doubt [that] the I.Q. is a composite measure of native and acquired abilities”; see: Lippmann, “Mr. Burt and the Intelligence Tests,” *The New Republic* May 2, 1923: 264.

⁴⁴ Richard W. Woodcock, “New Looks in the Assessment of Cognitive Ability,” *Peabody Journal of Education* 77, no. 2 (2002): 8; Boake, 396-398.

The early history of standardized testing cannot entirely fit within the history of education because standardized testing has some basis in social measurement and management. Contemporary versions of standardized testing grew out of nineteenth-century social scientists' methods for quantifying human characteristics, early psychologists' quest to measure intelligence, Progressive Era reformers' fear of racial degeneration and moral depravity, industrial capital's tendency to turn human labor into an abstract resource, and American armed forces' aim to sort recruits by their potential usefulness in combat. These branches all had a common biopolitical root: modern states have sustained themselves in part by generating knowledge about individual bodies and the social body. Standardization, including methods of educational assessment, was one of many biopolitical methods. Although the American federal government had relatively weak authority over the course of education, standardized testing allowed greater district, state, and national encroachment into practices traditionally deemed community affairs.

Testing and the Postwar World

My dissertation picks up where this short history leaves off, by unpacking the profound changes standardized testing made to the relationship between state and society during and after World War Two. What changed was not the use of testing itself—after all, standardized testing had been part of the American educational landscape for decades by the time the United States entered conflict—but the scale and stakes of testing data. War created an opportunity to use standardized tests to gather an unprecedented amount of data about the American populace's capabilities. This data, already reflective of a very specific and narrow understanding of aptitude, became a tool for defining the American state's postwar brainpower needs. Standardized testing validated postwar visions that premised democratic vitality on efficient social sorting. The

postwar state, in turn, validated the widespread use of standardized testing by using test data as a key metric for spotting and harnessing brainpower. What developed, however, was a gap between the limited technical applications for standardized testing data and what many Americans, both non-specialists and experts alike, asserted the data truly meant. Standardized testing data thus became a tool in broader arguments over who truly contributed to—and thus who deserved to benefit from—the postwar American state.

This dissertation also depicts how standardized testing served a vital function for the postwar American state without being internal to the American state. Until the late 1970s, the American state lacked a separate cabinet-level entity for education. Yet, as much as education was traditionally a responsibility of state and local governance, it was also entangled with scientific discovery, technological prestige, and military innovation—all issues central to the postwar state. The American state needed a way to harness brainpower without disrupting, at least on the surface, the traditional divisions of power for education. Standardized testing, in its various forms, filled that need. Standardized testing helped channel students into certain educational and vocational paths—steering to some degree their life course—without creating bureaucratic heft. The American state instead relied on outside manufacturers to research and produce standardized tests, thus allowing for vital state functions without warping the shape of the state. Even when the Department of Education eventually took on a broader mandate decades after its creation through No Child Left Behind, it acted as the supervisor for statewide test-based accountability programs. Standardized testing—and especially its creators—allowed the postwar American state to be strong where it was structurally weak.

Over the following five chapters, this dissertation tracks the way the American state came to rely on standardized testing for its wartime and postwar needs—and, eventually, how various

groups of Americans reacted against the use of standardized testing as a social sorting device. In Chapter One, “Manpower from the Masses,” I examine the vital role standardized testing served in the developing logic of the postwar national security state. The American state’s dependence on standardized testing began in earnest as it drafted millions of men during World War II. This dependence grew as the state reintegrated servicemen back into civilian life, and solidified as new test-development organizations undertook enormous projects for the federal government. Standardized testing provided the American state the mechanism for making manpower a measurable national resource. Both during and after the war, manpower served as a key unit of measurement for American military, business, and scientific needs. The rhetoric of manpower allowed policymakers to create measures—most notably, the GI Bill—that rewarded servicemen with social benefits without setting them apart as a privileged class of citizens. Manpower instead designated all members of the war effort, soldier and civilian alike, as citizen-workers, whose labor was both a civic duty and an abstract quantity serving various state needs. By prioritizing manpower, policymakers could endorse postwar plans that expanded the scale of the state’s educational responsibilities without changing its fundamental federal shape. Standardized testing, refined and massively employed during the war, provided the state a tool to continue quickly sorting millions of Americans by their abilities, aptitudes, and interests, all in the effort to extract the greatest possible manpower for postwar projects.

The second chapter, “American Talent Show,” examines how the state used standardized testing as a geopolitical tool. Well before the Eisenhower Administration responded to *Sputnik*, the federal government, extragovernmental educational organizations, and corporations began building a network dedicated to catching the United States’ brightest youth and steering them toward scientific and engineering careers that would secure American geopolitical supremacy.

Sputnik did not spark a commitment to science and mathematics so much as solidify the direction multiple stakeholders, both private and public, took in harnessing brainpower for geopolitical means. Test-based scholarship programs allowed corporations to use standardized tests to cull talent while promoting the merits of American capitalism. Standardized testing also allowed the federal government to dodge accusations of centralization at a moment when many alleged it had superseded traditional local control over education. Rather than restructure U.S. education, standardized testing became the netting that bound all vested parties together. Standardized testing provided heft and fuel for geopolitical ambitions that the rhetoric of American democratic brainpower could not alone.

Chapter Three, “Managing Marginal Men,” unpacks the way the federal government used standardized testing as a way to maintain oversight and control over black Americans’ social mobility. Over a several-year period following the Second World War, the United States Armed Forces integrated its troops—and in the process, replaced its “racial quota” with one based on the Armed Forces’ standardized aptitude entrance exam: the General Classification Test (GCT). In swapping the racial quota for a GCT quota, policymakers used standardized testing to forge a postwar social order that, on the surface, prioritized ability over race, but in effect used testing as a means to avoid disrupting too much of the existing white supremacist social order. By allowing some viable channel for deeper black participation in the armed forces, however limited, the Truman Administration could satisfy its aim to integrate and modify the military without creating too much discontent within military leadership at a vital moment in American statecraft. This pattern of the state disproportionately marginalizing black men through standardized testing continued through the advent of minimum competency testing and test-based special education tracking. Conversely, quota systems designed to increase the presence of minorities in graduate

programs came under legal fire during the 1970s, as white students alleged the looser test score requirements for minority candidates created a double standard that discriminated against qualified whites. Standardized testing thus not only reinforced the postwar racial order, but attempts to create racially equitable educational environments were met with resistance from whites who felt their test scores genuinely reflected why they merited a space in an increasingly competitive postgraduate environment.

This dissertation's fourth chapter, "Bad at Math," analyzes the emergence and development of dueling arguments about the root cause of American women's lack of presence in science and mathematics during the late-twentieth century. Although American girls tended to do better than boys in math classes, far more men wound up in engineering and laboratory positions than women. This anomaly persisted even as women entered higher education and the American labor force at unprecedented levels. Standardized testing data fueled two opposing arguments for this phenomenon. One group, influenced by second-wave feminist politics, maintained that girls were negatively socialized into women who avoided scientific careers. The other group, building off of sociobiological models, maintained that as a matter of biological capacity, females typically lacked the exceptional quantitative and scientific skills found among top male researchers. Both groups nonetheless relied on standardized testing data to build contrasting arguments about sex, gender, and women's role in the state. By the time the federal government formally launched its own investigation of the "gender gap," the debate had already begun to become the topic of mainstream debate, in which identity politics clashed with media depictions of "pink" and "blue" brains.

In the final chapter, "Birth of the Student-Consumer," I examine how parents, students, and pundits voiced their discontent with standardized testing during the 1970s and 1980s—and,

by securing reforms, unintentionally wed standardized testing to marketplace practices. Spurred by Ralph Nader's Public Interest Research Groups, students and parents voiced their frustrations with standardized testing from the position of consumers entitled rights—and accused certain testing companies of widespread consumer fraud. In securing transparency legislation, consumer advocates and student-consumers stripped Educational Testing Service and other test manufacturer of their longstanding protections as nonprofit organizations. To remain financially solvent—in light of not only laws that mandated the publication of test answers, but also a growing and federally legitimized test preparation industry—test manufacturers embraced a new business model. Test producers packaged old tests for commercial release and sued test-prep companies for copyright infringement. Ultimately these companies staged public battles with test-prep providers over who genuinely had the American middle class's best interests in mind. By framing their grievances as dissatisfied customers, student-consumers helped reinforce the broader marketplace aspects of standardized testing—and indeed, of higher education at large.

This dissertation concludes with a brief consideration of the political culture borne from No Child Left Behind and sustained through Common Core. The No Child Left Behind Act of 2001 (NCLB) established a profound shift in the federal government's relationship to education, one rooted in widespread standardized testing. Under the watchword of "accountability," NCLB mandated schools show quantifiable improvement in students' reading and mathematical abilities or face drastic consequences. Voucher programs and the threatening rhetoric of "school choice" compelled many public schools to overhaul their curricula and impose strict test preparation regimens. But what factors influenced the way American ultimately responded to these testing regimes? What broader forces shaped the ways they framed their discontent? And why does standardized testing seem to be more embedded that ever as an educational practice and a mass

tool for staking social identity, even as a growing number of Americans find the practice detrimental to genuine learning or social success? Ultimately, the fear of what could replace standardized testing—and the social and economic investment it would take to develop something more worthwhile—may be what will keep it a widespread educational practice.

CHAPTER ONE: MANPOWER FROM THE MASSES: STANDARDIZED TESTING, STATE LEGITIMACY, AND POSTWAR ORDER

As the Second World War revved the American economy into overdrive, many policymakers worried about postwar stability. War had relocated millions of men to defense positions, placed white women in the workforce at unprecedented levels, sparked a second wave of black internal migration, transformed manufacturers into materiel suppliers, and rerouted domestic spending patterns.⁴⁵ Yet the war could not go on indefinitely, nor could the men and women aiding the war effort. Whether or not the Second World War genuinely wrested the United States out of the Great Depression, the end of war threatened a calamitous economic transition. The peacetime national economy had to accommodate millions of returning service personnel and displaced defense workers or risk fermenting social distress and deep political unrest.

Policymakers made manpower—and in effect, standardized testing of all sorts—central to postwar socioeconomic plans. Both during and after the war, manpower served as the unit of

⁴⁵ Karen Tucker Anderson, “Last Hired, First Fired: Black Women Workers during World War II,” *Journal of American History* 69, no. 1 (June, 1982): 82-97; Ruth Milkman, “Redefining ‘Women’s Work’: The Sexual Division of Labor in the Auto Industry during World War II,” *Feminist Studies* 8, no. 2 (Summer, 1982): 336-372; Mark Aldrich, “The Gender Gap in Earnings During World War II: New Evidence,” *ILR Review* 42, no. 3 (April, 1989): 415-429; Sherrie A. Kossoudji and Laura J. Dresser, “Working Class Rosies: Women Industrial Workers during World War II,” *Journal of Economic History* 52, no. 2 (June, 1992): 431-446; Lee E. Ohanian, “The Macroeconomic Effects of War Finance in the United States: World War II and the Korean War,” *American Economic Review* 87, no. 1 (March, 1997): 23-40; Brian Wansink, “Changing Eating Habits on the Home Front: Lost Lessons from World War II Research,” *Journal of Public Policy & Marketing* 21, no. 1 (Spring, 2002): 90-99; Terrence H. Witkowski, “World War II Poster Campaigns: Preaching Frugality to American Consumers,” *Journal of Advertising* 32, no. 1 (Spring, 2003): 69-82; Kendall Hoyt, “Vaccine Innovation: Lessons from World War II,” *Journal of Public Health Policy* 27, no. 1 (2006): 38-57; and Lisa Krissof Boehm, *Making a Way Out of No Way: African American Women and the Second Great Migration* (Jackson: University of Mississippi Press, 2009).

measurement for military, business, and scientific needs. Manpower, as a national resource, could be harnessed and refined, ensuring the state used Americans effectively for broader geopolitical goals.⁴⁶ The logic of manpower allowed policymakers to create measures—in particular, the GI Bill—that rewarded servicemen with social benefits without setting them apart as a privileged class of citizens. Manpower instead placed all members of the war effort, soldier and civilian alike, as citizen-workers, whose labor was both a civic duty and an abstract quantity serving various state needs. By prioritizing manpower, policymakers could endorse postwar plans that expanded the scale of the state’s educational responsibilities without changing its fundamental federal shape. Standardized testing, battle-tested during the world war, provided the state the tools to continue quickly sorting millions of Americans by their abilities, aptitudes, and interests, all in an effort to extract the greatest possible manpower for postwar projects.

Through wartime military entrance testing, the American state solidified its use of standardization as a method for social order. Although standardization had been a feature of the American economy and society since the mid-19th century, the state adopted standardization in the 20th century as a way to regulate social attitudes, behaviors, and practices. This governmental approach to standardization typically took judicial, legislative, and regulatory forms. The General Classification Test (GCT) allowed the armed forces—and by extension, the American state—to incorporate standardization as a matter of protocol. Millions of men experienced the

⁴⁶ I use the term “national resource” here and throughout the chapter, not only in reference to the National Resources Planning Board—a federal resource-management agency which spanned the late 1930s and early 1940s and thus set the stage for wartime approaches to citizen manpower management—but also to signify that human resources (manpower, brainpower, man-hours, etc.) are vital to the state in a way distinct from natural resources (e.g., oil, bauxite, cotton, etc.). Human resources do double duty: they are found within the territory of the state while being a distinct creation of the nation. Human resources are thus not territorial goods—people aren’t harvested as one does oranges or almonds—but social goods. Manpower thus reflects the structures, institutions, policies, and mores that generate quantifiable goods that allow favorable comparisons to other nation-states. See: National Research Council, *Research—A National Resource, vol. II—Industrial Research* (Washington, D.C.: U.S. Government Printing Office, 1941); National Resources Planning Board, *Our National Resources: Facts and Problems* (Washington, D.C.: U.S. Government Printing Office, 1941).

GCT as part of the routine for military induction—and, in turn, saw their lives in the armed forces shaped in their performance on this exam. The high stakes the GCT carried, then, established a social precedent that standardized testing was a serious practice that merited genuine effort. The personal consequences for not taking this standardized test seriously were too steep not to comply.⁴⁷

Wartime standardized testing also set a foundation for the postwar expansion of guidance counseling, which led Americans to use testing data as a means for constructing their sense of self and relationship to society. Federal postwar planning emphasized the role higher education had to play in sustaining American vitality—and through the GI Bill, the state facilitated the veterans' postsecondary training. The state's reliance on advanced training enabled a massive expansion of American higher education during the immediate postwar period. Colleges and universities needed personnel who could not only help veterans adjust to their new role as students, but help these veteran-students select a path of study that would, at once, satisfy their interests, tap into their intellectual strengths, fulfill the state's geopolitical needs, and not drain school resources. Standardized testing—here in the form of guidance inventories and personality assessments—allowed guidance counselors to efficiently determine the path of best fit for their

⁴⁷ Desmond King and Marc Stears, "How the U.S. State Works: A Theory of Standardization," *Perspectives on Politics* 9, no. 3 (September, 2011): 505-518. King and Stears make an important development in the socio-historical study of the American state—namely, they turn away from the tedious slog over whether the American state has historically been a weak or strong entity. The socio-historiographical claim that the American state was unique in its thinness was, at its core, an aim to locate America's exceptional quality in the presumed weakness of the federal state. Attempts to pose the American state as historically strong, however, typically missed those aspects of governance where local control plays a prevailing counterforce, particularly education. Stears and King steer their analysis toward standardization as a method of statecraft, making it possible to illustrate those mechanisms (here, standardized testing) that have allowed the American state to be both convincingly weak and strong, centralized and decentralized, bureaucratic and diffuse. See also: J.P. Nettl, "The State as a Conceptual Variable," *World Politics* 20, no. 4 (July, 1968): 559-592; William J. Novak, "The Myth of the 'Weak' American State," *American Historical Review* 113, no. 3 (June, 2008): 752-772; Desmond King and Robert C. Liberman, "Finding the American State: Transcending the 'Statelessness' Account," *Polity* 40, no. 3 (July, 2008): 368-378; Laura S. Jensen, "Government, the State, and Governance," *Polity* 40, no. 3 (July, 2008): 379-385; and Julia Adams, "The Puzzle of the American State...and Its Historians," *American Historical Review* 115, no. 3 (June, 2010): 786-791.

charges. This type of diagnostic standardized testing produced a self-reflexive process by which people made sense of sense of their intellect, talents, and personality in uniform categorical fashions. For such testing to be successful in the postwar period, test developers and administrators developed methods for communicating about testing in ways that allowed individuals to personalize and personify certain psychometric principles.

Harnessing Manpower: Resource Planning, Vocational Guidance, and Wartime Education

Well before officially entering conflict, the Roosevelt Administration began sketching its ideal postwar condition. The vision that emerged among policymakers acknowledged the United States' irreversible role influencing international affairs, but refused to accept historic precedent for peacetime economic busts. Were the federal government to shy away from direct involvement in the postwar economy, as it did immediately after the First World War, a deep, sharp recession seemed inevitable. Private enterprise, left to its own devices, could not maintain wartime-level demand in peacetime conditions. Yet, for the sake of democracy and capitalism, a peacetime government could not control specific facets of the economy the way a wartime government could. To combat potential economic calamity, the Roosevelt Administration used the idea of manpower to mobilize American business, labor, military, and education in line with both its wartime demands and postwar visions.

In late 1940, President Roosevelt assigned the National Resources Planning Board (NRPB) the hefty task for harnessing manpower: consider “all constructive plans for significant public and private action in the post-defense period [...using] the natural and human resources of the Nation.”⁴⁸ The key for postwar vitality, NRPB determined, was government-sustained demand. As NRPB forecasted, war would create full employment and unprecedented gross

⁴⁸ National Resources Planning Board, *After Defense—What?* (Washington, D.C.: U.S. Government Printing Office, 1942), 9.

national product. By shifting massive governmental investment into public works projects, industrial partnerships and security programs, the American state could ensure peacetime private enterprise had incentive to absorb the estimated 20 million Americans left unemployed by demobilization—all while cushioning the American economy from periodic boom-bust cycles. In this vision, the federal government would not engage in central economic planning. Instead, the NRPB would serve as a coordinating agency that ensured public and private entities cooperated in the shared goal of sustained American postwar vitality. In its post-defense plans, NRPB established cooperation and coordination as actions that would protect federal, democratic, capitalist traditions against socialist and totalitarian threats.⁴⁹

The Roosevelt Administration enabled these plans by shifting American economic policy toward personnel management tactics that treated human effort as a national resource. As the United States began mobilizing for conflict, Roosevelt increasingly embraced Keynesian economic theory, which advocated large-scale public investment and deficit spending to maintain national economic stability. Keynesian theory cast the state as a manager whose involvement served as a multiplier for economic growth. Keynesian formulas treated national income as a dependent variable through which policymakers could gauge the effectiveness of certain economic policies.⁵⁰ By extension, embracing Keynesian methods required the state to

⁴⁹ NRPB, *After Defense—What?*; NRPB, *Post-War Planning* (Washington, D.C.: U.S. Government Printing Office, 1942); NRPB, *Demobilization and Readjustment: Report of the Conference on Postwar Readjustment of Civilian and Military Personnel* (Washington, D. C.: U.S. Government Printing Office, 1943); NRPB, *Progress Report, 1940-1941* (Washington, D.C.: U.S. Government Printing Office, 1941); NRPB, *National Resources Development Report for 1943—Part II: Wartime Planning for War and Post War* (Washington, D.C.: U.S. Government Printing Office, 1943); Charles E. Merriam, “The National Resources Planning Board: A Chapter in American Planning Experience,” *American Political Science Review* 38, no. 6 (December, 1944): 1075-1088; and Landon G. Rockwell, “The Planning Function of the National Resources Planning Board,” *Journal of Politics* 7, no. 2 (May, 1945): 169-178.

⁵⁰ Elba K. Brown-Collier and Bruce E. Collier, “What Keynes Really Said about Deficit Spending,” *Journal of Post Keynesian Economics* 17, no. 3 (Spring, 1995), 341-355; Patrick Renshaw, “Organized Labour and the United States War Economy, 1939-1945,” *Journal of Contemporary History* 21, no. 1 (January, 1986): 3-22; Patrick Renshaw, “Was There a Keynesian Economy in the USA between 1933 and 1945?” *Journal of*

treat human labor as a quantifiable unit—manpower—that could be measured and centrally managed. If national income was the way to measure sound economic policy, Keynesian approaches to manpower became the means to ensure national income remained high, both in the short and long term. Personnel management shifted the priority in labor away from the rights of workers and toward the needs of managers, whether the workplace was a shop floor or the U.S. Armed Forces.⁵¹

As a national and natural resource, manpower did not limit itself to a certain type of American. Every worker, whatever their vocation, possessed manpower. Manpower had some internal limits, namely a person's intellectual and physical capabilities. But the state could encourage measures that ensured individuals refined their manpower through training and received employment in areas that made the most of their capabilities. Using Keynesian logic, large-scale public investment in manpower cultivation would not only help protect the national economy from future economic instability—it would also give the state multiplied benefits when it reinvested refined manpower into public works projects. What mattered to the state, by extension, were the ways Americans could strengthen their manpower through training. Higher education, vocational institutions, and the military all provided convenient venues for boosting manpower. Distinctions between wartime educational, labor, and military needs, then, resided in the methods the state could employ for harnessing and distributing manpower.⁵²

Contemporary History 34, no. 3 (July, 1999): 337-364; Richard P. Adelstein, "'The Nation as an Economic Unit': Keynes, Roosevelt, and the Managerial Ideal," *Journal of American History* 78, no. 1 (June, 1991): 160-187.

⁵¹ Sanford M. Jacoby, *Employing Bureaucracy: Managers, Unions, and the Transformation of Work in the 20th Century* (Mahwah, NJ: Lawrence Erlbaum Associates Publishers, 2004).

⁵² For a sense of both how legislators, pundits, and managers made sense of manpower at the time, see: United States Senate, 77th Congress, 2nd Session, *Hearings Before a Subcommittee of the Committee on Education and Labor on S. Res. 291, part 1* (Washington, D.C.: U.S. Government Printing Office, 1942); United States Senate, 78th Congress, 1st Session, *Hearings Before a Subcommittee of the Committee on Appropriations on Investigation of Manpower, part 1* (Washington, D.C.: U.S. Government Printing Office, 1943); United States House of Representatives, 78th Congress, 1st Session, *Hearings Before the Committee on Military Affairs on H.R. 2239, H.R.*

The Roosevelt Administration created federal means for managing manpower, thus addressing military and economic demands through the same unit of measurement. By executive order, President Roosevelt established the War Manpower Committee in April 1942 in order to coordinate “the recruitment, vocational training, and placement of workers to meet the needs of industry and agriculture.”⁵³ Chaired by Federal Security Agency Administrator Paul V. McNutt, the War Manpower Commission also gauged shifting military and civil service manpower needs, and recommended policies to ensure the American labor force did not gravitate toward more alluring lines of work. The War Manpower Commission aimed to ensure government agencies and industries alike adhered to geographic and numerical hiring restrictions. Although limited in power by partisan pushback, the War Manpower Commission reinforced the idea that, as a matter of self-preservation, the state had the right to measure and manage human resources as it would natural resources.⁵⁴ Good citizens—human or corporate—did not take jobs that ran against the state’s immediate needs, because such jobs robbed the state of manpower.

Some industries concerned with personnel management turned to standardized testing to maximize manpower. One such company, Wichita’s Coleman Lamp and Stove Company, employed the personnel guidance company Search-Service, Incorporated to develop a battery of standardized vocational tests to determine who among their workers would make worthwhile shop foremen. To do so, Search-Service first gathered managers’ composite ratings for current

1742, H.R. 1728 and H.R. 992—*Bills Relating to the Full Utilization of Manpower* (Washington, D.C.: U.S. Government Printing Office, 1943); Harold W. Metz, *Is There Enough Manpower?* (Washington, D.C.: The Brookings Institution, 1942); Research Institute of America, *Manpower: Solving the Problems of Labor Shortage* (New York: Research Institute of America, 1943); and Asa S. Knowles and Robert D. Thomson, *Management of Manpower* (New York: The MacMillan Company, 1943).

⁵³ Executive Order 9139, “Establishing the War Manpower Commission,” April 18, 1942, <http://www.presidency.ucsb.edu/ws/?pid=16248>.

⁵⁴ Gladys M. Kammerer, *Impact of War on Federal Personnel Administration, 1939-1945* (Lexington: University of Kentucky Press, 1951); United States Senate, Subcommittee on Military Affairs, Executive Summary, S. 607, April 17, 1943.

foremen's employee relations and budgetary efficiency. Search-Service then subjected these same foremen to a series of standardized vocational aptitude tests to determine which exams produced the clearest and most distinct relationships between test-takers' responses and those two qualities. The psychologists Search-Service used these correlation coefficients to devise a master rating system—a T-score—that provided the Coleman Stove and Lamp Company a standardized, efficient way to select future leaders. Workers whose combined vocational guidance test results cleared a certain T-score threshold could be placed in “a ‘pool’ for future replacement and upgrading.”⁵⁵ As mathematically complex as this setup was, companies that hired psychological consultants such as Search-Service believed the net result would be an efficiently-managed workforce: no manpower would be wasted training men with insufficient skills for leadership positions, nor would manpower be squandered holding workers with “natural” skills down in lower positions. Psychologists portrayed efficient corporate personnel management as a matter of national security, too. As search-Service representatives asserted, “the problem of selecting personnel for industry on the home front [was] second only to the efficient utilization of human resources in the armed forces.”⁵⁶ Victory hinged on the testing, training, and enriching, American manpower, for the shop floor as well as the battlefield.

During wartime, the federal government implemented programs that steered both workers and employers toward embracing the idea of harnessing manpower through education. In this approach, education was not postsecondary training, but job-specific instruction designed to

⁵⁵ Irvin T. Shultz and Bentley Barnabas, “Testing for Leadership in Industry,” *Transactions of the Kansas Academy of Science* 48, no. 2 (September 1945): 160. See also: Jay L. Otis, Carroll L. Shartle, and William H. Stead, *Occupational Counseling Techniques: Their Development and Application* (New York: American Book Company, 1940); S.D. Chamberlain, “The Use of ‘Aptitude Tests’ in Sales Management,” *Journal of Marketing* 8, no. 2 (October, 1943): 159-164. For a sense of aptitude tests and research available for industrial and corporate use at the time, see: Hazel C. Benjamin, *Employment Tests in Industry and Business: A Selected, Annotated Bibliography* (Princeton, Princeton University, 1945).

⁵⁶ Shultz and Barnabas, “Testing for Leadership in Industry,” 160.

facilitate personnel management. Overseen by the War Manpower Commission, the Training Within Industry program reoriented labor toward viewing personnel management as a matter of resource refinement. To better refine the manpower needed for wartime industrial jobs—skilled manufacturing jobs that many millions of men had little prior experience given the previous decade of economic crisis—Training Within Industry packaged short job instruction training programs that shop foremen used to provide workers intensive supervised guidance. In time, Training Within Industry also developed and refined job relations training programs for foremen and supervisors. By instructing workers and supervisors how to interrogate the methods of production in a way that increased efficiency without inflaming workplace antagonisms, Training Within Industry reinforced the idea that manpower, as a unit of both liberal democracy and capitalism, depended upon subsuming personal needs to those of the state.⁵⁷

The National Resources Planning Board likewise set education as a cornerstone of its “post-defense” plans for democratic and capitalist vitality.⁵⁸ Asserting education of all types “would greatly assist the manpower readjustment process,” the NRPB sketched a dozen general provisions for postwar education in its 1943 plans for demobilization and readjustment.⁵⁹ At the base level, education would serve as a general counterbalance for employment levels: the fewer available jobs, the more the federal government would promote education as an opportunity to

⁵⁷ Training Within Industry, *Job Instruction: A Manual for Shop Supervisors and Instructors* (Washington, D.C.: War Production Board, 1942); Training Within Industry, *Second Annual Progress Report* (Washington, D.C.: War Production Board, 1942); Training Within Industry, *Job Methods Training: Training Sessions Outline and Reference Material* (Washington, D.C.: War Manpower Commission, 1943); Training Within Industry, *Training Within Industry Materials: Bulletins Issued by Training Within Industry and Outlines of the Training Within Industry Programs for War Plants and Essential Services* (Washington, D.C.: War Manpower Commission, 1945); and Mary Brigh, “‘We Cannot Afford to Hurry’: Training Within Industry Applied to Nursing,” *American Journal of Nursing* 44, no. 3 (March, 1944): 223-226; and William J. Breen, “Social Science and State Policy in World War II: Human Relations, Pedagogy, and Industrial Training, 1940-1945,” *Business History Review* 76, no. 2 (Summer, 2002): 233-266.

⁵⁸ NRPB, *Post-War Planning*, 6.

⁵⁹ NRPB, *Demobilization and Readjustment*, 42.

refine manpower for skills useful to the postwar state. Men who returned from military service would be allowed to resume interrupted vocational training—and, once fully demobilized, industries would be encouraged to admit “an appropriate quota of ex-service men” for peacetime apprenticeships.⁶⁰ Beyond vocational and secondary education, the NRPB also proposed coordination between the federal government and higher education institutions to devise plans that incorporated veterans more fully into American colleges and universities. In this vision, education was an important for demobilization, and the means by which military manpower could be reconverted to peacetime manpower in ways that created social good as well as generated profit.

As the National Resources Planning Board and War Manpower Commission tethered manpower to specific kinds of workplace training, the U.S. Office of Education (USOE) coordinated vocational training efforts among the nation’s disparate educational systems. Under the auspices of the Federal Security Agency (FSA), USOE oversaw a nationwide vocational development program for various industrial, mechanical, engineering, and managerial defense positions, training 1.75 million Americans between mid-1940 and Pearl Harbor.⁶¹ FSA Administrator Paul McNutt and USOE Commissioner John W. Studebaker insisted that the United States’ decentralized approach to educational oversight did not hinder vocational training efforts but instead allowed for a rapid rate of defensive vocational training. The officials portrayed shared federal, state, and local oversight on vocational education as the most democratic approach to defensive training—but this shared investment also created a complex

⁶⁰ NRPB, *Demobilization and Readjustment*, 45.

⁶¹ Federal Security Agency, U.S. Office of Education, “17 Months Defense Training Parallels World War Effort,” press release, ~10 December, 1941 housed in RG 12: Records of the Office of Education—Federal Security Agency: Office of the Commissioner: Office File on Wartime Educational Programs, 1940-1945, Box no. 1, Entry 8 (Advisory to Defense) National Archives, II, College Park MD.

network of responsibilities. The chain of authority over defensive vocational education connected various departments and intersected several levels of governance: the president, a state-level bureaucrat, and a local school board member all had a share in the direction of vocation education. Plans to preserve democracy by cultivating manpower also had to defer to traditional decentralized divisions of educational governance.⁶²

USOE created a wartime educational agenda that stressed manpower development over military drill in secondary schools. Maintaining that defense was not strictly a combat issue—nor was military preparedness simply a matter of marching exercises—the office urged secondary schools to pivot coursework toward equipping students with broad defensive readiness. Schools could develop “robust toughened bodies” by boosting inexpensive contact sports and exhausting activities.⁶³ Beyond preparing future combat troops, schools could aid the war effort by helping students develop technical skills with uses beyond the battlefield: youth with auto maintenance experience or radio repair skills would be valuable national resources both during and after the war. The Office proposed that even courses with no direct connection to defense—such as English and social studies—could be reoriented toward developing communication and contextual skills vital for a fully equipped citizenry. As USOE policymakers insisted, these courses were “military training in as real a sense as is military drill.”⁶⁴ By focusing on defense

⁶² Paul V. McNutt and John W. Studebaker, “Graphic Memorandum on Organization for Defense Training Through Vocational Schools,” Federal Security Agency--Office of Education, March 1941, housed in RG 12: Records of the Office of Education—Federal Security Agency: Office of the Commissioner: Office File on Wartime Educational Programs, 1940-1945, Box no. 1, Entry 8 (Advisory to Defense) National Archives, II, College Park MD.

⁶³ U.S. Office of Education, March 1942[?], “The Best Kind of High-School Training for Military Service.” p. 2. U.S. Office of Education (Washington D.C.); housed in RG 12: Records of the Office of Education—Federal Security Agency: Office of the Commissioner: Office File on Wartime Educational Programs, 1940-1945, Box no. 1, Entry 8 (Advisory to Defense) National Archives, II, College Park MD.

⁶⁴ USOE, “The Best Kind of High-School Training,” p.1.

rather than specific combat need, USOE made clear all youth could contribute to the war effort, so long as schools fully developed all aspects of their manpower.

The federal government collaborated with colleges differently, carving out overt military programs for post-secondary study. As USOE urged secondary schools to prepare students for well-rounded defense needs, the War Department rolled out Enlisted Reserve Corps programs for institutions of higher education. These programs, as well as their Air Force equivalents, sought to transform “a certain number of college students possessing superior qualification” into reservists prepared for sudden combat demands.⁶⁵ Yet, the War Department was less interested in making sure intellectually capable men found their way to college and more interested in ensuring their college training happened speedily. If men’s status as reservists depended upon completing their college education—and, in turn, their ability to lead proven by having earned a degree—then the armed forces needed speedier ways for college-goers to earn their credentials. Simply getting a man into educational channels appropriate for his intellectual capacity would not do: he also had to be trained as quickly as possible. As the Army and Navy jointly stressed, “[t]he country can no longer afford to have young men proceed with their education at a moderato tempo.”⁶⁶ Manpower was a time-sensitive resource, and educational institutions put the nation at risk if they refined it too slowly.

The Federal Security Agency also devised plans to improve American teachers’ manpower so they could better serve the state’s wartime needs. Although the American Armed Forces had developed the Engineering, Science, and Management Defense Training program (EMSDT) to equip young Americans with the skills needed for advanced air-based warfare, it

⁶⁵ Frank Knox, Army and Navy Joint Press Release, May 14, 1942, housed in RG 12: Records of the Office of Education—Federal Security Agency: Office of the Commissioner: Office File on Wartime Educational Programs, 1940-1945, Box no. 1, Entry 8 (Advisory to Defense) National Archives, II, College Park MD.

⁶⁶ *Ibid.*

also realized too few teachers had substantial training in mathematics, physics, aviation, or physical education. With no projected end to war in sight, officials projected the strain on educators would only worsen. Hence, in late spring 1942, USOE Commissioner John Studebaker solicited college and university academic directors to indicate the amount of summer sessions their respective institutions planned to offer in crucial subject areas. As Studebaker stressed, summer study programs would not only equip college teachers with the scientific know-how, but also deepen the bench of who could prepare reservists and defense personnel. With adequate summer intensives, high school and retired teachers could also do their share.⁶⁷ As the state made unprecedented manpower demands, it also had to devise new programs to ensure enough people could actually harness and refine young adults' potential skills.

By the end of the United States' first year of combat, officials in the Roosevelt Administration pursued two seemingly contradictory paths: lowering the draft age to the edge of adulthood, all while preparing for a postwar society in which youth delayed their independence. Although the Selective Service Act of 1940 required all men between 18 and 65 to register for the draft, men under the age of 20 were not conscripted when the United States first entered combat. By mid-1942, however, the War Department changed tack, and began publicly petitioning the president to lower the draft floor. Selective Service Director General Lewis Hershey argued that existing age restrictions could not keep pace with forecasted manpower demands: the army and navy needed millions more young men unencumbered by spouses or families. The Selective Service had already deferred millions of available men from the draft

⁶⁷ Memo from John W. Studebaker, U.S. Commissioner of Education, to Directors of Summer Sessions in Colleges in University, re: "Urgent Need of Certain Summer Session Courses, May 1, 1942, Federal Security Agency—U.S. Office of Education (Washington D.C.); housed in RG 12: Records of the Office of Education—Federal Security Agency: Office of the Commissioner: Office File on Wartime Educational Programs, 1940-1945, Box no. 1, Entry 8 (Advisory to Defense) National Archives, II, College Park MD. See also: Henry H. Armsby, *Engineering, Science, and Management War Training: Final Report* (Washington, D.C.: U.S. Government Printing Office, 1946).

because they had dependents—and men under 20 who voluntarily enlisted in the armed forces neither made up for that shortage nor consistently offered “A-1”-level capabilities. Following some congressional debate, including a failed Senate amendment requiring teenage draftees to spend a year in military training before facing overseas combat, President Roosevelt signed draft revisions into effect in November, 1942. Over two million men suddenly became eligible for conscription. Buoyed by public support, the federal government emphasized that the most valuable manpower came from young American men unburdened by any other obligations except those to the state.⁶⁸

Before the draft age officially dropped, however, policymakers and educational experts began formalizing practices that ensured youth would have more fully-developed manpower whether they stopped schooling after high school or continued onto college. As one War Department official explained in a fall, 1942 letter to USOE Commissioner John Studebaker, the armed forces needed more men with “preliminary preparation of a fundamental scientific and technical character,” who could more easily acclimate to specific combat positions.⁶⁹ Given these manpower needs, the department enacted a Pre-Induction Training Section within the

⁶⁸ Between exclusive coverage and Associated Press wire services, the *New York Times* provides a thorough timeline of governmental measures to lower the draft age floor. See: “Draft Registration June 30 for Youths from 18 to 20,” *New York Times* May 23, 1942: 1; “Roosevelt to Sift 18-19 Draft Plan,” *New York Times* May 27, 1942: 8; “3,000,000 Youths Register Today; Most, 18 and 19, Not Callable Yet,” *New York Times* June 30, 1942: 1, 6; “Wadsworth Seeks Permanent Draft,” *New York Times* July 20, 1942; “Senators Split on 18 to 19 Draft,” *New York Times* September 4, 1942: 15; George Gallup, “Draft of Youths Backed by Public,” *New York Times*, September 6, 1942: 19; “No 18-19 Draft Call Before Jan. 1, In the Opinion of the President,” *New York Times* September 12, 1942: 1, 11; Frederick R. Barkley, “House by 345 to 16 Lowers Draft Age to Take Boys of 18,” *New York Times* October 18, 1942: 1; “Youth Draft Delay Hits Married Men,” *New York Times*, November 8, 1942: 31; “Text of 18-19 Draft Bill,” *New York Times* November 11, 1942: 18; “Senators Abandon 18-19 Draft Fight,” *New York Times* November 12-1942: 16; “President Signs 18-19 Draft Bill,” *New York Times* November 14, 1942; “22,453 in ‘Teen Age Added to Draft,” *New York Times*, January 1, 1943: 12; Louis Stark, “New Draft Order Puts All 18 to 38 in Service in ’43,” *New York Times* April 13, 1943: 1; and “Draft Age to Stay at 18,” *New York Times* April 16, 1943: 11.

⁶⁹ James P. Mitchell, letter to John W. Studebaker, September 5, 1942, pg. 1; housed in RG 12: Records of the Office of Education—Federal Security Agency: Office of the Commissioner: Office File on Wartime Educational Programs, 1940-1945, Box no. 1, Entry 8 (Advisory to Defense) National Archives, II, College Park MD.

Manpower Branch of its Services of Supply Civilian Personnel Division. Yet, consistent with its earlier dismissal of demands for military drill in high schools, the War Department stressed that pre-induction training programs would not take a narrow focus. For high schools with fewer than 300 pupils—by far the greatest proportion of institutions at the time—the Pre-Induction Training Section recommended offering courses in physics and mathematics, shop, physical fitness, and “pre-flight aeronautics,” and encouraged larger high schools to provide classes in radio, typewriting, anatomy, and first aid.⁷⁰ Because defensive needs would change as the war continued, the US Armed Forces needed men and women who could be pressed into a variety of tasks. The war effort also required participation beyond combat, and universities designed programs of study to ensure women as well as men deemed unfit for service also met wartime agricultural, educational, scientific, and industrial demands. Some, such as the Ohio State University, offered a variety of study tracks—from accounting clerk, to cartography, to farm management—that could be completed within two years.⁷¹ Wartime educational policy ensured that American youth had enriched their manpower in ways applicable to shifting defensive needs.

Yet educators also took the early war period to speculate about the future of youth. While considering postwar education, the Washington-based Educational Policies Commission forecasted postwar economic demands would encourage “expanded conceptions of youth” that

⁷⁰ Rall I. Gigsby, Memo to USOE John W. Studebaker, re: “Pre-induction Training,” August 3, 1942, pg. 4; housed in RG 12: Records of the Office of Education—Federal Security Agency: Office of the Commissioner: Office File on Wartime Educational Programs, 1940-1945, Box no. 1, Entry 8 (Advisory to Defense) National Archives, II, College Park MD.

⁷¹ For one example of higher education pre-induction study tracks, see: The Ohio State University, War-Time Training for Young Men (Columbus: The Ohio State University, c. 1943).

would keep many youth in educational institutions and public agencies through the age of 20.⁷² The commission also predicted the end of general and vocational educational tracks. Following the trends of wartime educational demands, postwar students would be expected to develop a wide array of practical and scientific skills that could be applied to numerous peacetime challenges. By drawing 18- and 19-year olds into combat, the American state may have also created conditions that encouraged prolonged transitions into adulthood in postwar society. If the concept of manpower bound educational, military, and workplace trajectories—and if the state favored the cultivation of manpower for its wartime needs over simply gathering bodies for combat—then postwar planners may have envisioned a social structure where the transition to adulthood was marked by thorough cultivation of an individual’s own manpower. Before it could fully figure out what to do with men in postwar society, however, the American state had to first figure out how to sort men already poised for military service.

Measuring the Masses: The General Classification Test and Preparatory Guides

The concept of manpower made it easier for the American state to unify its wartime needs, but the military had particular time constraints that required a quicker way to sort men by their potential capabilities. Men and women who took accelerated college programs or on-the-job vocational training had time to develop their manpower before giving it to the American defense effort. The tradeoff in these cases was straightforward: development now for services rendered later—and, theoretically, for well-paying jobs afterward that could serve the postwar state. The Army and Navy, by contrast, needed to locate men for particular roles without

⁷² William G. Carr, et. al, memorandum to Commissioner of Education, USOE, and FSA, re: “Part II of the report on regional conferences on ‘the Implications of the Defense Training Programs,’ held at Baltimore, Md., Atlanta, Ga., Boston, Mass., Chicago, Ill., and Kansas City, Mo, in June, 1942,” October 1, 1942,” pg. 2; housed in RG 12: Records of the Office of Education—Federal Security Agency: Office of the Commissioner: Office File on Wartime Educational Programs, 1940-1945, Box no. 1, Entry 8 (Advisory to Defense) National Archives, II, College Park MD.

extending the time it took to create soldiers and sailors. The military needed a snapshot of servicemen's existing manpower capabilities before training even began.

The solution ultimately embraced by the state—mass standardized testing as part of the induction process—would find widespread postwar use outside of the military, and helped legitimize the field of psychology to many who encountered military entrance exams. War brought psychologists and their objective tests an unprecedented audience: over ten million men encountered the Army General Classification Test (AGCT).⁷³ Through its widespread use of general classification tests, the armed forces instilled a high social value in standardized testing. AGCT test scores were meaningful because they carried such high stakes. Not only did psychologists assert that the test indicated a recruit's overall aptitude, but AGCT scores also set the foundation for servicemen's military careers, influencing the potential positions men could earn and the perils they might face. The AGCT exam set a precedent for American social value in standardized testing as a mechanism that could profoundly alter an individual's life course—and as such, ushered in a wave of commercial preparatory material. These preparatory materials reinforced the state's manpower needs: men who understood the AGCT could provide a better sense of their manpower capabilities and would thus be of better to the nation, in war and beyond. Using the AGCT to link recruits' aptitude to military order and the preservation of democracy, military psychologists set rhetorical connections between individuals' mental skills, institutional efficiency, and the needs of the American state which endured well beyond war's end.

⁷³ The AGCT was also used by the navy, and as such, was sometimes called the Navy General Classification Test, the Army-Navy General Classification Test, or just the General Classification Test. But, for the sake of consistency, I will focus mainly on the Army's use of the General Classification Test and, as such, will use the acronym AGCT.

In order to create a military entrance exam, the U.S. War Department had to forge bureaucratic space for psychological personnel well before the United States' formal entry into World War II. By spring 1940, the department made several organizational adjustments to ensure a psychological system for classifying recruits and inductees would be in place before any massive military escalation—in particular, creating a Personnel Testing Section within the department's Adjutant General's Office the year prior. (The division would later be called the Personnel Research Section.) The department also tapped into existing scientific institutions created during the First World War, requesting that the National Research Council form a Committee on Classification of Military Personnel. To counterbalance this bureaucratic hodgepodge, the war department appointed Walter Bingham as both head of the new National Research Council committee and Chief Psychologist of the Army. Under Bingham's watch, the Personnel Testing Section and the Committee on Classification of Military Personnel had leeway to plan various standardized objective tests.⁷⁴ This leeway, however, came with the tradeoff of psychology having a disjointed space within the state.

Once assembled, the psychologists under Bingham's supervision had to create a test for incoming recruits that measured complex thinking skills without unfairly rewarding educational advantages.⁷⁵ An entrance exam that relied too heavily on specific subjects would only reflect

⁷⁴ Joseph Zeidner and Arthur J. Drucker, *Behavioral Science in the Army: A Corporate History of the Army Research Institute* (Alexandria, VA: United States Army Research Institute for the Behavioral and Social Sciences, 1988), 22-34.

⁷⁵ For a sense of how the US Armed Forces continued to cull psychological expertise during wartime—not only fully-trained psychologists with doctorates, but those with undergraduate and graduate training who could still benefit the war effort—see: Stueart Henderson Britt, "Psychology and the War," *Psychological Bulletin* 39, no. 4 (April, 1942): 255-260; Stueart Henderson Britt, ed., "Psychology and the War," *Psychological Bulletin* 39, no. 6 (June, 1942): 369-410; Stueart Henderson Britt, "Office of Psychological Personnel—Report for the First Six Months," *Psychological Bulletin* 39, no. 9 (November, 1942): 773-793; Stueart Henderson Britt, "Occupational Deferment of Psychologists and Psychologists in Training," *Psychological Bulletin* 39, no. 10 (December, 1942): 873-879; Roger M. Bellows and Marion W. Richardson, "Training in Military Personnel Psychology: Minimum Requirements for College Courses," *Psychological Bulletin* 40, no. 1 (January, 1943): 39-47; American Psychological Association, "College Curriculum Adjustments in Psychology to Meet War Needs," *Psychological*

the coursework a soldier had taken, neglecting more fundamental verbal and mathematical reasoning skills—the difference, for example, between questions about geometry properties and questions about spatial reasoning. Yet psychologists also couldn't make test questions too abstract. An exam that didn't "appeal to the average officer and soldier as sensible" would also elicit inaccurate results.⁷⁶ Researchers spent the spring and summer of 1940 creating thousands of questions and designing experimental versions of the test, cutting out overly verbal items and revising the test form to expand its range of difficulty. By August, Army psychologists completed version 1A of the Army General Classification Test. In its final form, the AGCT was a 40-minute, 150-question exam that alternated between clusters of multiple-choice vocabulary, arithmetic, and block-counting questions, becoming more difficult the further test-takers advanced.⁷⁷ Although researchers would create several other versions of the AGCT during the war, each test form sought to strike a balance between measuring how quickly recruits could answer general knowledge questions and how deep their verbal and quantitative aptitudes ran.⁷⁸

In order for the AGCT to have any meaningful use for the army, however, psychologists actually had to standardize the exam. Test standardization is a multi-step process, and its

Bulletin 40, no. 7 (July, 1943): 528-535; Dael Wolfle, "The Army Specialized Training Program Course in Personnel Psychology," *Psychological Bulletin* 40, no. 10 (December, 1943): 780-786; and Donald G. Marquis, "The Mobilization of Psychologists for War Service," *Psychological Bulletin* 41, no. 7 (June, 1944): 469-473.

⁷⁶ Staff, Adjutant General's Office Personnel Research Section, "The Army General Classification Test," *Psychological Bulletin* 42, no. 10 (December 1945): 761; Staff, Adjutant General's Office Personnel Research Section, "Personnel Research in the Army: 1. Background and Organization," *Psychological Bulletin* 40, no. 2 (February, 1943): 129-135

⁷⁷ Adjutant General's Office Personnel Research Staff, "The Army General Classification Test."

⁷⁸ Psychologists in the US Armed Forces also developed other aptitude tests during the course of the war, all used to assess incoming soldiers' set of skills. The GCT, however, remained the broadest test of skills and the largest sorter of recruits. See: Staff, Adjutant General's Office Personnel Research Section, "Testing As a Part of Military Classification," *Science* 97, no. 2526 (May 28, 1943): 473-478.

fundamentals illustrate how the AGCT would set a profound precedent for widespread American social constructs of normalness, intelligence, and national worth:

. Standardized tests measure individual performances in relation to a larger group. After taking a standardized test, a person's performance is translated to a raw score, and that raw score is then compared to other test-takers' performances. This comparative performance is typically represented as a percentile—e.g., “So-and-so performed better on this exam than 73% of test-takers.” Finally, this percentage is translated to a scaled score. Scaled scores—almost always the number that people refer to when discussing standardized test scores—represent the raw score's location on a statistical model. This model is typically a normal distribution. In a normal distribution, about two-thirds of all values are within one standard deviation from the mean, 95% fall within two standard deviations, and virtually all values fall within 3 standard deviations.⁷⁹ This normal distribution is set to an arbitrary score range—for the AGCT, practically all scores fell between 40 and 160, with a median score of 100—and divided into evenly spaced increments. Because standardized tests are often given multiple times over many years, these score scales are often based on a representative sample originally gathered during a test's experimental phases. Although these master scales are often adjusted and recalibrated, especially if a test becomes very popular or if the original sample group no longer reflects the current group of test-takers, they remain the backbone for how test-takers understand their performance and how institutions assess test-takers' worth. The statistics behind the final score ultimately mean that it usually takes far more effort for a person to raise their score from slightly below-average to slightly above-average than it would to raise their score from above-average to well above-

⁷⁹ A standard deviation is the mean distance of all values within a set of numbers from the mean—in other words, the average distance from the average. The smaller a standard deviation, the more tightly clustered the values sit around the mean, and vice versa.

average. Whether a speed or power test, standardized assessments groups a wide range of performances under the idea of “average” —and in turn, standardized test scores become shorthand for social usefulness.⁸⁰

While standardizing the AGCT, army psychologists had to reduce the friction between the state’s needs and the exam’s limitations. Namely, researchers had to standardize an exam designed for armed force recruits when no such population existed. Researchers developed the earliest version of the AGCT months in advance of the Selective Service Act of 1940, which mandated all male citizens and resident aliens between the ages of 21 and 36 to register for the draft.⁸¹ Lacking an available pool of draftees, the test’s designers essentially invented a sample population, administering the AGCT to over 4000 men already serving in the Army and Civilian Conservation Core.⁸² Using these results, army psychologists built a score scale using a 5-grade system: the higher the scaled score a test-taker received, the higher their classification. Yet army officials remained unsatisfied with the number of recruits who received the AGCT’s lowest possible classification—causing the test’s designers to make its lowest grade much narrower and the next-lowest grade much wider. Readjusting the boundary between Grade IV and Grade V classification may not have made hundreds of thousands of incoming recruits any more or less capable for duties in the armed forces—but it reduced the stigma these men would have faced during their service, opened opportunities that would have otherwise been closed, and allowed

⁸⁰ In psychometric theory, a speed test measures, quite simply, how fast a test-taker can complete questions, while a power test measures a test-taker’s rate of correct responses. There are rarely, if ever, tests that are entirely speed or power tests—and AGCT Form-1a, which had a 40-minute time limit for 150 questions arranged in spiral-omnibus format, would certainly be between these two poles. To see some of the theory developed about these types of tests during this time period, see: Harold Gulliksen, “Speed versus Power Tests,” in *Theory of Mental Tests* (New York: John Wiley & Sons, 1950): 230-244.

⁸¹ “Selective Service Act of 1940,” Sec 3 (a); Sec 5 (f), 1:2.

⁸² Walter V. Bingham, “Personnel Classification Testing in the Army,” *Science* 100, no. 2596 (September 29, 1944): 276.

the armed forces a new rationale for their use. By inventing a representative sample where none existed, psychologists gave the armed forces the ability to quickly compare millions of men against an imagined average—and with a couple basic adjustments, established the boundaries for which men were more or less useful for the state.

Psychologists for the Adjutant General's Office made repeated efforts to show the AGCT was a valid examination, revealing in the process the complicated logic needed to demonstrate standardized tests' effectiveness.⁸³ Some studies focused on the relationship between test-takers' scores and their civilian occupations. Researchers concluded that the greater degree of training a man needed for his job, the higher his AGCT score. Accountants, chemists, personnel clerks, and engineering students typically earned far higher scores than miners, farmers, and lumberjacks. Yet, these results only validated the AGCT by suggestion: surely the test measured men's intellectual capacities accurately, because the men who scored highest typically had highly skilled, training-intensive occupations.⁸⁴

Researchers also studied AGCT results to see if the test helped or hurt the military's policies on illiteracy. Widespread illiteracy complicated American wartime manpower needs. Illiteracy was not proof in itself of a man's mental faculties—but the military could not afford to refine all illiterate draftees' manpower when time and money could instead be spent on men with basic reading skills. Yet the state also could not turn away too many illiterate men, especially as

⁸³ Test validity is most commonly defined as the relationship between the content on a test and the actual ability the test aims to measure or forecast. While convenient, this definition of validity blurs the many different types of validity, as well as eliminates the ethical and ontological aspects of whether a test is truly valid. See: Samuel Messick, "Test Validity and the Ethics of Assessment," (Princeton: Educational Testing Service, 1979); Denny Boorsboom, Gideon J. Mellenbergh, and Jaap van Heerden, "The Concept of Validity," *Psychological Review* 111, no. 4, 1061-1071; and College Board, "Test Validity," College Board website, <https://research.collegeboard.org/services/aces/validity/handbook/test-validity> (last accessed May 12, 2017).

⁸⁴ Adjutant General's Office, "Relationship Between Main Civilian Occupation and AGCT Standard Score, Part II" (Washington, D.C: Adjutant General's Office, 1945): 12-14; Adjutant General's Office, "Relationship Between Main Civilian Occupation and AGCT Standard Score, Part III" (Washington, D.C.: Adjutant General's Office, 1945).

the American war effort continued to escalate: combat required bodies.⁸⁵ This dilemma created wildly shifting recruitment policies for illiterate men during the war, and placed pressure on army psychologists to determine the actual benefit of special training units geared toward streamlining illiterate recruits into the general draftee population after intensive reading training. Analyzing the test-taking patterns of illiterate recruits in special training units, researchers at Fort Leavenworth, Kansas discovered a pattern: many boosted their AGCT scores by focusing on test question with little written content. In particular, such recruits gravitated toward block-counting questions, which made test-takers determine how many blocks a three-dimensional form would contain based off of a two-dimensional representation. By favoring text-less questions, many men reached the Grade IV AGCT classification needed to graduate from their special training units and enter general basic training.⁸⁶ These findings revealed a tension between standardized testing and manpower needs. Men who gamed the AGCT by avoiding verbal question may have cost the state specific types of skills—but by excelling at block-counting questions, they demonstrated other kinds of aptitude that could have been harnessed to benefit wartime needs. Standardized tests revealed recruits’ capabilities whether or not men took them as intended.

Army psychologists also analyzed the relationship between AGCT scores and race, repudiating certain racist logic in order to build legitimacy for their own field. Decades after the United States employed psychologists to design entrance exams for the First World War, differences between blacks’ and native-born whites’ scores on Army Alpha and Beta Tests continued to fuel racialized theories of intelligence. The AGCT’s designers sought to avoid racial

⁸⁵ Gladyce H. Bradley, “A Review of Educational Problems Based on Military Selection and Classification Data in World War II,” *Journal of Educational Research* 43, no. 3 (November, 1949): 161-165; Deborah Brandt, “Drafting U.S. Literacy,” *College English* 66, no. 5 (May, 2004): 485-502.

⁸⁶ “Differential Patterns of Item Attempts on the Army General Classification Test Exhibited By Grade IV and V Men Tested at the Reception Center, Fort Leavenworth, Kansas, 1944,” AGO Report No. 519: 2.

differences in test scores becoming additional evidence for prevailing arguments about “native capacity.”⁸⁷ As one 1942 Army report stressed, black and white recruits’ AGCT scores were not fully comparable because no existing samples could account “[for] economic, social, and cultural factors [and] with respect to educational opportunities and background.”⁸⁸ No adequately sized sample for similarly equipped men existed. Even college-educated black men—who scored similarly to white recruits with some high school education—lacked economic resources and social safeguards enjoyed by white inductees. Yet the Army’s repudiation of racist intelligence theories was, above all, an aim to distance itself from inappropriate uses of statistical data. The report’s emphasis on incomparable subgroups—namely, the implausible conditions under which white and black recruits’ AGCT scores could be soundly judged against each other—was primarily a lesson in statistical logic. Psychologists, working on a project of unprecedented scale, wanted to ensure their newly-gained legitimacy was not just as quickly lost, particularly as they aimed to pave over the enduring cultural memory of the Army Alpha and Beta Tests. Their concern was less about social inequities themselves (the armed forces, after all, remained segregated throughout the war) and more about the misuse of statistical psychological data to maintain racialized theories of difference.

Psychologists’ difficulty attributing meaning to score differences may have also been complicated by imbalanced access to AGCT preparatory material. Although the armed forces maintained secure testing conditions at its induction centers, recruits with even a little financial leeway could acclimate to the Army and Navy entrance exams by purchasing a guidebook. Guidebooks not only prepared men for the general classification tests, but also enforced the

⁸⁷ “Interpretations of AGCT Test Scores of Negro and White Selectees,” AGO-PRS Report No. 307[2?]: 7 April 1942: 1.

⁸⁸ “Interpretations of AGCT Test Scores of Negro and White Selectees,” 8.

notion that testing was central to efficient personnel management. Recruits may not have known every occupation the army and navy needed to fill, but guidebooks offered the promise that, with some practice, men could find a role among thousands of job classifications that made the most of their education, skills, and potential.⁸⁹ These guidebooks used the idea of aptitude to connect personal responsibility to military victory. As one guidebook published by Arco asserted, “the most pronounced aptitude may be 2 ½ and [sic] 3 times as strong as the least pronounced” within any particular individual—and recruits needed to spend time determining what was the best they had to offer the state.⁹⁰ Guidebook publishers thus justified their products as a service rather than a strictly commercial item by providing the means for potential recruits to make the best display of their manpower capabilities on military entrance exams.

To encourage sales, guidebook authors repeatedly stressed to readers that the AGCT was not an intelligence test. Intelligence tests, these authors noted, measured relatively fixed mental capacities, and if an individual’s general native intelligence could not change, then there would be no need to purchase a guidebook to improve one’s AGCT score. Guidebook authors also had to delicately balance the relationship between AGCT scores and education. Prep guide writers routinely emphasized that training for the AGCT should not be confused for formal education—but, to remain enticing to readers, guidebooks also emphasized that military entrance exams assessed a wide range of aptitudes, most of which could be strengthened with diligent short-term training. Some prep books, such as one printed by Pergande, used anecdotal data to suggest that the most educated men did not always do best on the AGCT:

⁸⁹ Arco Publishing Company, *Practice for the Army Tests* (New York: Arco Publishing Company, 1942), 4. See also: Alvin Harold Kaplan, *Prepare for the Official U.S. Army-Navy Tests* (New York: Capitol Publishing Company, 1942); William Vilmos, *Instruction for Tests in the Armed Forces* (New York: Grosset & Dunlap, 1943); Current Book Company, *Air Corps Practice Tests* (New York: Current Book Company, 1942); and Practice Publishing Company, *Practice for the Soldier’s IQ Test* (Cincinnati: Practice Publishing Company, 1943).

⁹⁰ Arco, *Practice for the Army Tests*, 3, 4.

"In a Milwaukee examination for the air corps, a young man who had completed only the eighth grade made a rating of 120 points. For this test 80 was the passing grade and three hours were allowed. On average 50% of those who applied here failed to pass. In one group four out of five college students failed. The highest score ever made in Milwaukee was 142 and the lowest 23."⁹¹

Without confirming whether the man with an eighth-grade education or the highest-scoring Milwaukee recruit had ever used a prep guide, Pergande severed the relationship between schooling and test scores. The state needed men of many different talents—including those ignored by schools.

The surefooted, precise way prep guides familiarized their audiences with the military induction process contrasted sharply with their motley assortment of AGCT practice material. Prep guides walked readers through the initial steps of the induction process. These guides familiarized potential recruits with soldier qualification cards, stressing that men should prepare for the biographical information required on the 8 ½ -by-11 cardstock form. One guide went so far as to tell readers the types of clothing recruits would receive upon induction.⁹² But specific AGCT material was often harder to come by. Some guides, such as one printed by Arco, bombarded students with nearly thirty pages of vocabulary practice, asserting that “[i]t may be stated on sound, psychometric authority that the vocabulary test is one of the most important single means of testing intelligence”⁹³—without ever indicating whether the words printed in the prep guide would ever appear on the AGCT. Mathematics practice often took the form of elementary arithmetic and algebraic review. Readers were assured that the inductions tests

⁹¹ Pergande Publishing Company, *Full-Length Practice Examinations for U.S. Army-Navy Classification Tests* (Milwaukee, Pergande Publishing Company, 1943), 59.

⁹² Pergande, *Full-Length Practice Examinations*, 6.

⁹³ Arco, *Practice for the Army Tests* 13.

“measure ability and intelligence with a very high degree of accuracy”⁹⁴. One company even included additional testing material that had little to do with preparing recruits for the AGCT, such as general intelligence tests for civil service licensures; within a span of three questions, test-takers were asked what the definition for “versatile,” Bizet’s most famous work, and the primary use for a trombone.⁹⁵ In the absence of reliable information about actual AGCT questions, prep guides replaced quantity for quality.

Neither the gravity of war nor imprecise preparatory material stopped some publishers from reminding readers they were, in fact, businesses with other goods for sale. Pergande encouraged men with lackluster vocabularies to purchase its vocabulary and spelling guidebook, “which contain[ed] 4000 specifically selected words” suitable for review.⁹⁶ Arco reminded readers that the AGCT was likely not the only standardized test they would face in their lifetime—“After the War—You’ll Want a Good Position!”—and listed the printing and pricing information for dozens of other licensing exam prep guides.⁹⁷ Publishers also built off of their wartime prep guide consumer base when the United States maintained a peacetime draft. Some postwar guides, such as one written by former chief of the Adjutant General Office’s Manpower Analysis Section Reuben Horchow, aimed to capitalize on the author’s wartime experience and expertise. Horchow’s guide, *How to Get Ahead in the Armed Forces*, posed military service as a near-certainty—and whether a man enlisted or received a draft notice, he had plenty of opportunities to make the most of his required service. Horchow stressed that a soldier’s test records existed far beyond the superiors with whom he had direct contact: a man’s

⁹⁴ Pergande, *Full-Length Practice Examinations*, 7.

⁹⁵ Pergande, *Full-Length Practice Examinations*, insert n6.

⁹⁶ Pergande, *Full-Length Practice Examinations*, 13, 60.

⁹⁷ Arco, *Practice for the Army Tests*, ad [np]

“Qualifications Record (tests and all) stands for you and furnishes the chief basis on which action is taken [or] selection is made.”⁹⁸ By this logic, purchasing and using a prep guide was the only way for a soldier to ensure that any official who came in contact with his permanent record would know his true potential.

By the time Horchow’s guide was published in the early 1950s, the link between standardized test scores, military efficiency, and the state’s needs had already been set, even as the armed forces transitioned from the GCT to the Armed Forces Qualification Tests (AFQT). Incoming soldiers who resisted their role in this system by intentionally botching their standardized entrance exam faced grave consequences. Horchow asserted that the “personnel people at induction stations have a pretty good idea about who should pass the test, and they also have ways of finding out if a man is ‘faking.’”⁹⁹ Even a nervous man who unintentionally thwarted a reliable measurement of his aptitudes could undermine his entire military career; personnel who selected soldiers for promotions and advanced training preferred candidates with consistent score records to men with “tests that are ‘out of line’ with each other.”¹⁰⁰ Horchow’s warnings were not just vague threats; psychologists had conducted several studies to see if “malingerers” displayed any generalizable test-taking patterns.¹⁰¹ Regardless of whether

⁹⁸ Reuben Horchow, *How to Get Ahead in the Armed Forces* (New York: Reuben Horchow, 1951), 24.

⁹⁹ Horchow, *How to Get Ahead*, 24.

¹⁰⁰ Horchow, *How to Get Ahead*, 25.

¹⁰¹ Harry Goldstein, “A Malingering Key for Mental Tests,” *Psychological Bulletin* 42, no. 2 (February, 1945): 104-118; C.N. Cofer, June Chance, and A.J. Judson, “A Study of Malingering on the Minnesota Multiphasic Personality Inventory,” *Journal of Psychology* 27, no. 2 (1949): 491-499; Penelope P. Pollaczek, “A Study of Malingering on the CVS Abbreviated Individual Intelligence Scale,” *Journal of Clinical Psychology* 8 (January, 1952): 75-81; Miriam E. Crowley, “The Use of the Kent EGY for the Detection of Malingering,” *Journal of Clinical Psychology* 8 (October 1952): 332-337; and Jane McReynolds, “Development of motivation keys for the Armed Forces Qualification Test Forms 3 and 4 (AFPTRC-TN-56-60),” May 1956, cited in Esther Barlow, *Abstracts of Personnel Research Reports: VIII. 1954-1968 (AFHRL-TR-68-124)* (Lackland Air Force Base, Texas: Air Force Systems Command, 1968).

Horchow's words of caution had any truth behind them, such prep guides reinforced the notion that the only thing worse than a man with little aptitude to give the state is a man unwilling to give their skills to the state. This sentiment would be sustained by the federally backed push to readjust veterans to the state's postwar educational and occupational demands.

Postwar Readjustment: The GI Bill, Guidance Counseling, and Personality Inventories

Well before World War II ended, American policymakers began developing legislative measures that would ensure returning soldiers readjusted to peacetime society, all while continuing to contribute their manpower to the postwar state. However, the main congressional provision for veteran readjustment—the GI Bill—revealed the limits to which the federal government could use education as a means for postwar social planning. For veteran readjustment to truly succeed, educational and corporate institutions also had to embrace psychological developments in guidance counseling and personality inventories. This burgeoning psychological field added new dimensions to standardized assessments, and further encouraged Americans see themselves as measurable categorical types. Through standardized guidance inventories, veterans and civilians alike facilitated postwar manpower planning by internalizing the ways various institutions made sense of student bodies and labor pools.

Many policymakers realized midway through the war that hodgepodge federal provisions for returning servicemen would provide insufficient support, but momentum for comprehensive veteran benefits ultimately came from outside efforts. The National Resource Planning Board had drafted initial plans for veterans' postwar reintegration, but politicized disdain for the NRPB and Rooseveltian governance in general had weakened the board's ability to push forward policy. (Congress made the issue moot by eliminating the board's funding in 1943.) The American Legion, a relatively conservative veterans organization, ultimately did what the NRPB

could not: draft legislation and organize a massive public relations campaign to ensure Congress would address the issue of veterans' readjustment in a timely manner. The Legion drew upon existing state-level provisions for veterans to craft legislation for an expansive set of entitlements, including unemployment benefits, home and farm loan benefits, and educational aid. For the Legion, an all-encompassing package seemed more likely to withstand congressional bandying than the existing piecemeal bills languishing in Congress. Boosted by sympathetic press coverage from publishing magnate William Randolph Hearst's network of newspapers—as well as the evocative title “The GI Bill of Rights”—the ambitious legislation made its way to the 78th Congress carrying popular support, particularly for having been developed by former military men rather than government bureaucrats.¹⁰²

As much clout as the GI Bill of Rights carried when it entered Congress, the legislation faced several key detractors who opposed the way the bill would recalibrate the relationship between the government, citizens, and soldiers. As the legislation worked through the Senate Finance Committee in early 1944, lawmakers gathered testimony from leaders of several rival veterans' interest organizations, who believed that expansive federal benefits for all veterans would harm disabled veterans already underserved by the government. Several state school board chancellors also offered the subcommittee their concerns that broad federal provisions would create bureaucratic duplication that undermined the traditional subfederal control of education. Even the bill's chief sponsor, Mississippi Senator John Rankin, turned against the legislation after realizing the wide array of benefits available to returning soldiers could disrupt the white supremacist system that kept black citizens at the bottom of the socioeconomic order

¹⁰² R.B. Pitkin, “How the First GI Bill Was Written” (part 1 of 2), *American Legion Magazine* January, 1969: 24-28, 51-53; R.B. Pitkin, “How the First GI Bill Was Written,” (part 2 of 2), *American Legion Magazine* February, 1969: 22-26, 48-51; Kathleen Frydl, *The GI Bill* (Cambridge: Cambridge University Press, 2009).

bereft of political power. These detractors all expressed a concern that the postwar order would weaken the very power networks kept the existing social structure afloat.¹⁰³

Despite its many discontents, the GI Bill became law on June 22, 1944, retaining the delicate balance of power between federal and state governments while striking a new relationship between education and citizenship. Title II of Public Law 346 granted veterans, whatever their physical condition upon discharge, numerous educational opportunities. The bill defined educational or training institutions to include:

“all public or private elementary, secondary, and other schools furnishing education for adults, business schools and colleges, scientific and technical institutions, colleges, vocational schools, junior colleges, teachers colleges, normal schools, professional schools, universities, and other educational institutions, [and] business or other establishments providing apprenticeship or other training on the job”¹⁰⁴

By staking such an expansive concept of education, the GI Bill ensured that veteran education remained valuable insofar as it enabled former servicemen to expand their manpower capabilities in ways that best suited both the individual and the state. The law entitled soldiers with at least ninety days’ service and “discharged or released [...] under conditions other than dishonorable” to at least one year of education, with up to four years available depending on a veteran’s length of service.¹⁰⁵ The GI Bill allowed veterans to continue their education at any institution

“whether or not located in the State in which he resides,” but insisted that the federal government

¹⁰³ United States Senate, 78th Congress, 2nd Session, “Hearings Before a Subcommittee of the Committee on Finance, United States Senate, 78th Congress, 2nd Session, on S. 1617: A Bill to Provide Federal Government Aid For the Readjustment in Civilian Life of Returning World War II Veterans,” January 14, 15, 21, 24, February 11, 14, 28, March 8, 10, 1944; United States Senate, 78th Congress, 2nd Session, “Hearings Before the Committee on World War Veterans’ Legislation, House of Representatives, 78th Congress, 2nd Session, on H.R. 3917 and S. 1767: To Provide Federal Government Aid for the Readjustment in Civilian Life of Returning World War II Veterans,” January 11, 12, 13, 17, 18, February 24, and March 9, 10, 27, 28, 29, 30, 31, 1944”; Frydl, *The GI Bill*.

¹⁰⁴ Public Law 436, Title II, Sec. 400.

¹⁰⁵ *Ibid.*

would not “exercise any supervision or control, whatsoever, over any State educational agency, or State apprenticeship agency.”¹⁰⁶ By placing education alongside home loans and unemployment payments that veterans could—but by no means were required—to take, Public Law 346 kept federal expansion in check by deferring to existing bodies of governance and using the language of choice.

Not all veterans benefitted equally from the new educational provisions. The men who gained the most were those who already did well on the general classification tests: younger white men who had completed their schooling. This cohort was considerably more likely to complete high school and at least some college than nonveterans in the decade following their service. Black veterans, by contrast, were largely written out of the GI Bill’s educational benefits—and by extension, out of the promises of the postwar state. By deferring educational appropriations to the states, Congress ensured the GI bill wouldn’t disrupt traditional racist educational power structures. Although black veterans in the north had limited positive educational experiences, southern black veterans were offered little. Lackluster administrations for southern black vets—many of whom were disqualified from the GI Bill’s generous unemployment provisions for refusing to accept white VA officer’s job placement offers for menial labor—encouraged many to relocate to northern urban areas. While the GI Bill may not have explicitly excluded benefits to veterans on grounds of race, the execution of the bill’s provisions ensured that black veterans would not be able to use their benefits to upend America’s racial social order, nor that the bill’s benefits would contradict existing race-based policies in education, housing, or employment.¹⁰⁷

¹⁰⁶ *Ibid.*

¹⁰⁷ The debate over who did—and more pointedly, who didn’t—benefit from the GI Bill has gone on for decades. Underneath this debate is a broader question over whether educational demographics would have continued

As part of its educational benefits, the GI Bill also authorized the Veterans' Affairs Administrator to "arrange for educational and vocational guidance to persons eligible for education and training."¹⁰⁸ Such guidance would not only include informing veterans about "the need for general education," but also encouraging returning servicemen to pursue job-specific training "in the various crafts, trades, and professions."¹⁰⁹ Though slight, this provision established the idea that guidance was a necessary component of readjustment. Veterans, left to their own devices, may not make occupational or educational choices that would maximize what manpower they had to offer the state. Guidance offered a conduit between individual vocational needs and state manpower demands.

To accommodate the postwar wave of veteran-students, universities created guidance programs that both situated servicemen to college life and adjusted their institutions to postwar demands. Through new guidance policies and procedures, higher education institutions sought to

changing as they had in the late 1930s in the absence of war: Was the GI Bill deeply restructure educational access, or merely compensate for ongoing patterns interrupted by war? This debate typically fails to fully consider factors of race and region, which complicate judgments of how transformative the law was: Did the legislation create new opportunities for millions or merely increase the distance between those accepted into the idea of the mass public and those written out of it?

See: Charles B. Nam, "Impact of the 'GI Bills' on the Educational Level of the Male Population," *Social Forces* 43, no. 1 (October, 1964): 26-32; Davis R.B. Ross, *Preparing for Ulysses: Politics and Veterans During World War II* (New York: Columbia University Press, 1969); Keith W. Olson, "The G.I. Bill and Higher Education: Success and Surprise," *American Quarterly* 25, no. 5 (December, 1973): 596-610; Joshua Angrist and Alan B. Kruger, "Why do World War II Veterans Earn More than Nonveterans?" *Journal of Labor Economics* 12, no. 1 (January, 1994): 74-97; David H. Onkst, "'First a Negro...Incidentally a Veteran': Black World War Two Veterans and the G.I. Bill of Rights in the Deep South, 1944-1948," *Journal of Social History* 31, no. 3 (Spring, 1998): 517-543; Michael C.C. Adams, "Who Didn't Use the G.I. Bill? Notes on a Lingering Question," *Studies in Popular Culture* 23, no. 2 (October, 2000): 65-74; Sarah Turner and John Bound, "Closing the Gap or Widening the Divide: the Effects of the G.I. Bill and World War II on the Educational Outcomes of Black Americans," *Journal of Economic History* 63, no. 1 (March, 2003): 145-177; Marcus Stanley, "College Education and the Midcentury GI Bills," *The Quarterly Journal of Economics* 118, no. 2 (May 2003): 671-708; Christopher P. Loss, "'The Most Wonderful Thing Has Happened to Me in the Army': Psychology, Citizenship, and American Higher Education in World War II," *The Journal of American History* 92, no. 3 (December, 2005): 864-891; Edward Humes, "How the GI Bill Shunted Blacks into Vocational Training," *The Journal of Blacks in Higher Education* no. 53 (Autumn 2006): 92-104; and Ira Katznelson and Suzanne Mettler, "On Race and Policy History: A Dialogue About the G.I. Bill," *Perspectives on Politics* 6, no. 3 (September, 2008): 519-537

¹⁰⁸ Public Law 436, Title II, Sec. 400.

¹⁰⁹ *Ibid.*

comply with federal mandates without diluting their existing prestige. The University of Wisconsin system steered veteran-students toward guidance services; in the absence of recent educational measurements, a battery of guidance test could more accurately capture a veteran's set of aptitudes and abilities. Universities such as Princeton felt a responsibility to their graduates as well as potential freshmen, pledging to help alumni with job guidance. Psychologists, like veterans, also had to readjust, and City College of New York launched a four-week program to acclimate psychologists to the guidance needs of veteran-students. Although over sixty percent of those who served in the Army and Navy during the Second World War had not completed high school, the number of veterans who did have a high school education could have placed unbearable strain on American colleges and universities had institutions not made adjustments to most of them used GI Bill benefits to pursue higher education ¹¹⁰

Worried about the maintenance of their own internal standards, some universities conducted internal studies to determine whether veteran-students maintained the same academic rigor as their civilian counterparts. In the absence of standards, the University of Michigan worried "government assistance might induce a good many veterans to try college work," using resources at the expense of otherwise qualified civilians.¹¹¹ Without any previous bases for comparison, Michigan turned to measurements they could control and measure: grade point averages. Yet researchers at the university could not discern a consistent, let alone statistically significant, difference between the grade of veterans and non-veterans. (A noticeably positive

¹¹⁰ Wisconsin State Department of Public Instruction, *The Wisconsin Educational Program: A Handbook for World War II Veterans—War Workers and Counselors* (Madison, WI: Veterans Recognition Board, 1944); John Gray Peatman, "Special Training Program for Counselors of War Veterans,"; Joseph Hillis Miller and John S. Allen, *Veterans Challenge the Colleges: The New York Program* (New York: King's Crown Press, 1947); Francis J. Brown, *Educational Opportunities for Veterans* (Washington, D.C.: Public Affairs Press, 1946): 42.

¹¹¹ University of Michigan, "Academic Achievement of Veterans and Non-veterans in Six Colleges of the University of Michigan" (Ann Arbor: University of Michigan Institute of Human Adjustment, 1946), 1.

difference in grade point averages existed for veterans who returned to school after leaving for combat.) When confronted with a new legal basis for educational access—one that did not allow for much prediction—colleges and universities turned to internal standards to ensure their own educational prestige remained undiluted.

Psychologists also used wartime conditions and postwar planning to lay the foundation for guidance counseling. Guidance, as understood by experts, involved adjusting an individual's personal appraisal and aspirations to fit with broader social needs through objective measurements—in other words, getting people to make sound life plans after some self-reflection and a few standardized tests.¹¹² Guidance counseling, by extension, helped individuals personalize broader social norms and values. Schools had long been institutions for cultural alignment—replicating hierarchies and mores found outside the classroom—but wartime defense planning introduced a new dimension to social adjustment. War had turned labor into manpower, and postwar economic stability depended in part upon Americans' ability to view their efforts and ambitions through this abstract sense of work. Vocational guidance experts, in turn, devised mechanisms that would personalize postwar planning, particularly for those Americans most crucial for sustained successful economic adjustment: returning soldiers, college students, and adolescents. Although vocational counseling and personality inventories existed before the Second World War, defense planning ushered in an approach to guidance counseling that married the language of self-discovery and theory behind standardized objective testing with the logic of manpower.

¹¹² Put another way by experts from the era: "Guidance may be defined as the assistance of the individual, through the use of scientific procedures and ascertained facts, to plan and pursue his education and life in the light of his aptitudes, past achievement, and interests, and also, in the light of attested and recognized values." See: David G. Ryan, *The First Step in Guidance: Self-Appraisal: A Report of the 1940 Sophomore Testing Program* (New York: American Council on Education Cooperative Test Service, 1941), 5-6.

Before vocational guidance experts could steer Americans into postwar jobs, they first had to convince the public well in advance that there would be a worthwhile economy to participate in after the war. This burden of proof was especially high for enlisted men, whose defense jobs were largely redundant once war ended. One pamphlet, printed in 1943 by the nonprofit Public Affairs Committee, assured soldiers the federal government would not shrug off its obligations as it had after the last war with “a \$60 bonus and a ticket home.”¹¹³ Temporary unemployment—on the government’s dime as a paid furlough—would be the best way to ensure that the transition to a peacetime economy would occur in a graduated manner. As experts suggested, if the federal government could be trusted to reposition millions of Americans into wartime defense positions, it could be trusted with ensuring postwar economic success.

Clear as their message may have been, there simply weren’t enough vocational guidance experts during or after the war to provide direct counseling to Americans in non-combat positions. As much as psychologists asserted the public good of their field by publicizing their wartime efforts, they were still a relatively small group—many of whom directly involved in the war effort. Most vocational counselors, particularly those working outside of a military environment, also lacked professional training. In 1944, the War Manpower Commission estimated that the United States had only 6,000 vocational counselors, nearly half of whom “only partially trained in the professional sense.”¹¹⁴ Comparatively few high school guidance counselors worked in the Southeast, Pacific Northwest, or New England—and those schools with

¹¹³ Maxwell S. Stewart, *When Will I Get Out? When Will I Find A Job?* (New York: Public Affairs Pamphlets, 1943), 9.

¹¹⁴ War Manpower Commission Bureau of Training, *The Training of Vocational Counselors* (Washington, D.C.: War Manpower Commission, 1944), 1.

guidance systems often assigned hundreds of students per counselor.¹¹⁵ While dozens of colleges and universities scrambled to develop vocational guidance training programs, the federal government leaned on existing community pillars to steer youth into making sound workforce decisions. Even if the government could curb some vocational pathways, as it did by ending voluntary military enlistment midway through the Second World War, community institutions had to shoulder primary responsibility for how well youth grappled with their choices. As the war drew a growing number of employable minors into the workplace, the federal government needed to address a growing dilemma: how to provide guidance to millions of teenagers and young adults without having widely-dispersed trained personnel.¹¹⁶

Psychologists tried to solve this problem by encouraging readers to understand themselves in standardized terms through self-appraisal and self-assessment. Guidebooks provided tests, charts, and questionnaires to help readers understand their own skills and aptitudes categorically. As authors often stressed, most people could succeed in several vocations—but individuals had a personal responsibility to determine what objective criteria those jobs held in common, as well as what capabilities made them a better fit for certain lines of work. Some guides used the language of accounting to help readers standardize their sense of self, advising youth to create a “Personal Balance Sheet,” in which job seekers soberly assessed their physical, mental, sartorial, vocational, and psychological “assets” and “liabilities.”¹¹⁷ Other

¹¹⁵ U.S. Office of Education Federal Security Agency, *Guidance Problems in Wartime--Education and National Defense Series Pamphlet No. 18* (Washington, D.C.: US Government Printing Office, 1942), 1. The Federal Security Agency based these numbers on figures attained in 1937 and 1938.

¹¹⁶ Federal Security Agency, *Guidance Problems in Wartime*, 3-5. Employment certifications for youth between 14 and 17 years of age nearly doubled in the first year of mobilization, and only continued to escalate as the United States formally entered war. See also: Alfred J. Cardall, *A Wartime Guidance Program for Your School* (Chicago: Science Research Associates, 1943), 5-14.

¹¹⁷ George J. Lyons and Harmon C. Martin, *The Strategy of Job Finding* (New York: Prentice-Hall, 1944), 48-49; Ernest Henry Suerken, ed., *Find Your Vocation*, revised edition (New York: Hobson Book Press, 1941).

guides standardized the dimensions of personality, asking veterans to assess their level of friendliness, initiative, dependability, and “pep and enthusiasm” on a five-point scale.¹¹⁸

Vocational guides that avoided tests still encouraged readers to categorize themselves in broad occupational archetypes: mechanically-inclined people should stay away from jobs best suited for those inherently gifted at communications and sales.¹¹⁹ By familiarizing readers to self-assessment, guidance experts sought to convince Americans that vocational planning was “a *process*, not an *event*,” and that process began with understanding one’s strengths and limits in standardized ways.¹²⁰

Through vocational guidebooks, student readers learned they held primary responsibility in maintaining postwar economic health. Adolescents who failed to develop skills and habits would lack the mental dexterity needed to make informed choices about how they entered the workforce. In this framework, people possessed different levels of intelligence—and while environmental factors could enhance skills, good training and a supportive environment would do little to help an individual who was naturally poor at a skill. As these guides suggested, accepting natural variations in aptitude, whether in reference to general intelligence or broad skill sets, was the necessary first step in lining individuals with jobs that best suited their predispositions. To this end, these guidebooks suggested that adulthood came when a person faced the choice of how to develop and expend their manpower. As one wartime vocational guide suggested, “freedom from fear” was not only one of the Atlantic Charter’s core tenets, but

¹¹⁸ Frank S. Endicott, *How to Find and Succeed in Your Post-War Job* (Scranton, PA: International Textbook Company, 1946), 32.

¹¹⁹ Mark Morris, ed., *Career Opportunities* (Washington, D.C.: Progress Press, 1946); Burchard W. Yale, *Your Job: Selecting It, Securing It, Succeeding in It* (New York: Barnes & Noble, 1947), 12-17.

¹²⁰ Endicott, *How to Find Success in Your Post-War Job*, 6 [emphasis in original]; Henry Eugene Watters, *Youth Makes the Choice: Vocational and Social Self-Guidance* (Nashville, TN: Broadman Press, 1943), 109-148;

a learned response to the uncertain. Those who neglected to overcome their fears through well-reasoned conditioning risked a future of uncertainty—not only for themselves, but also the nation.¹²¹ Regardless the future job market’s strength, young Americans had a civic—and moral—responsibility to understand how they could best serve the postwar economy.

Realizing adolescents often weren’t their own best guidance counselors, psychologists and school boards also developed materials that allowed teachers to guide students through the process of test-based personality assessment. For the psychologists who designed personality inventories, vocational aptitude test results became meaningless when educators neglected to learn fundamental test mechanics. Many assessments had complex scoring systems—such as the Strong Vocational Interest, which allowed for over 40,000 possible weighted responses in 36 career fields. Educators risked wasting their students’ potential manpower with sloppy, halfhearted calculations.¹²² Some school boards, meanwhile, went beyond single assessments and developed semester-long courses in vocational guidance. Developed in 1941, the School Board of Chicago’s Self-Appraisal and Careers course devoted several weeks toward honing students’ understanding of psychological tests, statistical concepts, and test interpretation. By helping students assess test data without feeling “that some mysterious finality is inherent in a result because it is expressed numerically,” teachers could help youth understand their role in the

¹²¹ Blanche Paulson, *The Magic of the Mind* (Chicago: Chicago Board of Education Bureau of Child Study, 1943).

¹²² John G. Darley, *Clinical Aspects and Interpretation of the Strong Vocational Interest Blank* (New York: The Psychological Corporation, 1941), 9. The Strong assessment was not the only test of its era with complex scoring mechanisms. Proprietary differences may have altered the number of vocational fields or the range of possible responses, but other such assessments had similarly complicated and tedious scoring mechanisms. In the era before widespread electronic scoring technologies, psychologists wrote detailed manuals for counselors and educators to interpret the results by hand. See: G. Frederic Kuder, *Revised Manual for the Kuder Preference Record* (Chicago: Science Research Associates, 1946).

post-defense order by attuning to think of their capabilities as improvable abstract qualities.¹²³

Becoming an even-keeled adult not only meant finding a career, then, but using psychological tests to understand how one would fit in best in a rapidly changing world.

Some researchers devoted their time to developing standardized examinations for personality types that were not only easily internalized, but would in time prove highly profitable. During the war, Katharine Briggs and her daughter Isabel Briggs Myers constructed the earliest form of the Myers-Briggs Type Indicator, namely as an alternative method for measuring potential workplace strengths for women who lacked job experience. Inspired by Jungian archetypical models, Briggs and Myers argued that personalities could be categorized along four oppositional spectrums. The two researchers theorized that individuals favored either extraverted or introverted behavior, sensing versus intuitive perception, thinking versus feeling judgment, and general judgment versus perception. Briggs and Myers devised a forced-choice questionnaire that compelled test-takers to indicate their preferences along these spectrums. The duo categorized results into sixteen distinct personality types, stressing that none was inherently better than the other. Introverted, sensitive, thinking, judging individuals were as useful to the workforce as extraverted, intuitive, feeling, perceptive people—so long as they knew their personality dynamics and had an awareness of which jobs suited what personality types best. This archetypical approach to workplace personalities would become far more popular decades as corporations and organizations purchased the Myers-Briggs Type Indicator from Educational Testing Service (and later Consulting Psychological Press) beginning in the 1960s.¹²⁴

¹²³ Grace E. Munson and Lester J. Schloerb, *High School Course in Self-Appraisal and Careers: Teacher's Manual* (Chicago: Chicago Board of Education Bureau of Child Study, 1941), 52.

¹²⁴ Following the sale of MBTI's exclusive publishing rights to Consulting Psychologists Press in the mid-1970s, the test gained global corporate and popular attention. A great deal of this allure lay not only in its memorable classification system—ENTJs and ISFPs can easily carry their designation as a mark of how they fit best in different types of workplaces and careers—but also the tightly controlled, highly profitable method of discovering

Nonetheless, its widespread use and familiarity by the end of the twentieth century illustrates the culmination of what emerged during and immediately after World War II: the corporate embrace of vocational guidance testing as a way to make the American labor force personalize and absorb the responsibility of workplace adjustment.

As professionalized guidance counseling and vocational assessments became more widespread practices, psychologists attuned Americans to a worldview in which it was their civic responsibility to find a job that fulfilled the nation's manpower needs. Self-assessments ingrained in Americans new patterns for relating to the workplace: quantifying your personality type made it easier to figure out the line of work that would best serve you and your workplace. A poorly adjusted workforce could not provide the state robust manpower, even if jobs were fully available. Human resources, then, not only involved extracting and refining manpower, but also guiding individuals toward educational and vocational choices that would create the highest quality manpower for the state. Participatory democracy meant embracing a system in which citizens determined their capacity as workers for the good of the state. Standardized guidance tests, vocational batteries, and personality assessments allowed psychologists to continue

one's personality type. Free online MBTI simulations aside, the only way to get a genuine Myers-Briggs personality designation is to sit for the actual test, as administered by someone who paid \$1700 to earn their certification. Consulting Psychologists Press's exclusive control over MBTI publishing and certifications has translated to an estimated \$20 million in annual profits. What complicates the relationship between the MBTI and other tests is that neither Myers nor Briggs were trained psychologists, and their namesake exam remains largely ignored by the academic psychological community despite decades of use. Still, standardized vocational guidance remains a widespread and profitable business—perhaps especially when unchained from academic constraints.

See: Katherine Bradway, "Jung's Psychological Types" *Journal of Analytical Psychology* 9, no. 2 (July, 1964): 129-135; Isabel Briggs Myers, *The Myers-Briggs Type Indicator: Manual* (Palo Alto, CA: Consulting Psychologists Press, 1962); Mary H. McCaulley, "The Myers-Briggs Type Indicator: A Measure for Individuals and Groups," *Measurement and Evaluation in Counseling and Development* 22, no. 4 (January, 1990): 181-195; Douglas P. Shuit, "At 60, Myers-Briggs is Still Sorting Out and Identifying People's Types," *Workforce Management* (December, 2003): 72-74; Katherine D. Myers, "An Extended History of the Myers-Briggs type Indicator Instrument" *MBTI Type Today* website, 2011, <http://www.mbtitoday.org/downloads/An-Extended-History-of-the-MBTI.pdf>; and Lillian Cunningham, "Myers-Briggs Personality Test Embraced by Employers, but not by Psychologists," *Denver Post* 18 December, 2012: 3C.

justifying the importance of their methods in the postwar order while adjusting millions of Americans to their upcoming responsibilities.

CHAPTER TWO: AMERICAN TALENT SHOW:

STANDARDIZED TESTING AND THE COLD WAR SCRAMBLE FOR BRAINPOWER

As the Cold War developed, American lawmakers and educationists obsessed over the development of brainpower. If the United States did not harness and refine this national resource, experts warned, the country would squander its ability to maintain geopolitical supremacy. The nation needed a way to determine which youth had the greatest potential, guide them toward rigorous high school coursework, and encourage them to pursue science and engineering tracks in college—all while ensuring youth who weren't naturally gifted were also steered toward paths that fulfilled the nation's needs and fostered personal satisfaction. There was, however, one catch. Despite the fear the United States would slide behind the Soviet Union in scientific and technological innovation (and despite the burgeoning administrative state), the federal government remained reluctant to overextend its role in the shape of American education, a system historically built on decentralized authority and populated by non-state administrative actors.¹²⁵ Together, these various nodes of oversight created a complex postwar educational

¹²⁵ Paul E. Peterson, *The Politics of School Reform, 1970-1940* (Chicago: University of Chicago Press, 1985); Lloyd P. Jorgensen, *The State and the Non-Public School* (Columbia: University of Missouri Press, 1987); Douglas J. Slawson, *The Department of Education Battle, 1918-1932: Public Schools, Catholic Schools, and the Social Order* (Notre Dame: University of Notre Dame Press, 2005); Nancy Beadie, *Education and the Creation of Capital in the Early American Republic* (Cambridge: Cambridge University Press, 2010)

For work on the administrative state, see: Rebel E. Schiller, "The Administrative State, Front and Center: Studying Law and Administration in Postwar America," *Law and History Review* 26, no. 2 (Summer, 2008): 415-427; Karen M. Tani, "Welfare and Rights Before the Movement: Rights as a Language of the State," *Yale Law Journal* 122, no. 2 (November, 2012): 314-383, esp. 368-379; Meg Jacobs, "'How About Some Meat?': The Office of Price Administration, Consumption Politics and State Building from the Bottom Up, 1941-1946," *Journal of American History* 84, no. 3 (December, 1997): 910-941.

administrative network and political culture: American education had numerous governors, often at cross-purposes and with intersecting zones of control. If the federal government no longer had vested authority or reason to massively test tens of millions of citizens, as it had during the Second World War, and if massive swaths of white Americans actively resisted any federal interference in the form or pace of education, how could the state find and train the brains it so desperately needed?

Standardized testing companies solved this dilemma by providing products for identifying brainpower that validated both traditional divisions of educational oversight and served the geopolitical needs of the Cold War administrative state. Although standardized testing companies existed earlier in the twentieth century, they were often small-scale affairs typically devoted to the intelligence and personality inventories. After the Second World War, psychologists, emboldened by their wartime contributions to the state, collaborated with educational leaders to consider how standardized assessments could continue to be useful to American society. Philanthropic organizations collaborated with psychologists and leaders of elite universities to form Educational Testing Service (ETS), the largest of several postwar standardized test manufacturers. A private not-for-profit enterprise, ETS could provide testing material to the armed forces without bloating federal bureaucracy—and federal contracts

For work on non-state educational administrative actors (textbook committees and designers, teaching unions, disciplinary organizations, and parent-teacher associations, among others), see: Andrew Hartman, *Education and the Cold War: The Battle for the American School* (New York: Palgrave Macmillan, 2008), esp. 55-72; Marjorie Murphy, *Blackboard Unions: The AFT and the NEA, 1900-1980* (Ithaca: Cornell University Press, 1992), 175-195; Allan A. Needle, "Project Troy and the Cold War Annexation of the Social Sciences, in *Universities and Empire: Money and Politics in the Social Sciences During the Cold War*, ed. Christopher Simpson (New York: The New Press, 1998), 3-38; Peter B. Dow, *Sputnik Politics: Lessons from the Sputnik Era* (Cambridge: Harvard University Press, 1991), 1-31; Susan Levine, *School Lunch Politics: The Surprising History of America's Favorite Welfare Program* (Princeton: Princeton University Press, 2011), 1-10; Judith Kafka, "Shifting Authority: Teachers' Role in the Bureaucratization of School Discipline in Postwar Los Angeles," *History of Education Quarterly* 49, no. 3 (August, 2009): 323-346; Robyn Muncy, "Cooperative Motherhood and Democratic Civic Culture in Postwar Suburbia, 1940-1965," *Journal of Social History* 38, no. 2 (Winter, 2004): 285-310; Christopher J. Phillips, "The New Math and Midcentury American Politics," *Journal of American History* 101, no. 2 (September 2014): 454-479.

provided ETS financial solvency and legitimacy beyond its initial philanthropic funding. As American higher education expanded, and as the federal government grew more concerned with the quality of American scientific expertise, psychometric standardized examinations became more valuable for national goals. Because school districts and colleges remained clients rather than legally mandated users, early postwar test manufacturers nationalized educational assessments and admissions policies while maintaining the spirit of voluntary association. In effect, ETS and other early postwar test companies standardized the type of information various state- and non-state actors wanted to know about young American citizens. By the time both sputniks entered the atmosphere, American education had already begun to transition toward a nationwide talent hunt, in which standardized tests helped determine which brains were most geopolitically useful—and which Americans held the greatest marketplace value.¹²⁶

Harnessing Brainpower: Educational Testing Service, the Korean War, and the SSCQT

The United States enabled its Cold War pursuit of science and technology through a power-sharing arrangement of private and public funding. Prior to the Second World War, the largest funders for scientific research had been private philanthropic organizations—namely the Rockefeller Foundation. War compelled the American state, and in particular the military, to invest in scientific research in order to accelerate technological development. After the Second World War, the military and Rockefeller Foundation joined other federal, philanthropic, and academic organizations to jointly sustain funding for scientific endeavors. This shared custody of scientific and technological funding—neither fully separate from nor entirely bound to the state—meant that the American scientific Cold War effort took on several seemingly-

¹²⁶ Hartman, *Education and the Cold War*; Mary Dudziak, *Cold War Civil Rights: Race and the Image of American Democracy* (Princeton: Princeton University Press, 2010); Barbara Barksdale Clowse, *Brainpower for the Cold War: The Sputnik Crisis and National Defense Education Act of 1958* (Westport, CT: Greenwood Press, 1981).

contradictory characteristics: a commitment to applied research while continuing to fund science for its own sake; visions of a global democratic future while producing tools for an American hegemonic agenda; and researchers who debated the ethics of their creations while agency heads committed to destructive anticommunist agendas. To satisfy the numerous interests of their federal, business, philanthropic, academic, and military funders, postwar scientists (especially postwar social scientists) wedded knowledge work to the promotion of industry and democracy, all while glossing over the difference between experimental approaches to knowledge and the notion of objective truth. The truth, in this approach, was clearly demonstrable—and what was true would also be good for democracy because it was also good for business. Not beholden to a single source or type of funding, American Cold War science glossed over any tensions between the promotion of democracy and a belief in free market values.¹²⁷

Scientists, business leaders, educational experts, and policymakers all cultivated the idea of brainpower, gradually building a case for how the United States could best use its human

¹²⁷ Paul N. Edwards, *The Closed World: Computers and the Politics of Discourse in Cold War America* (Cambridge, MIT Press, 1996); Jessica Wang, *American Science in an Age of Anxiety: Scientists, Anticommunism, and the Cold War* (Durham: University of North Carolina Press, 1999); Toby A. Appel, *Shaping Biology: The National Science Foundation and American Biological Research, 1945-1975* (Baltimore: Johns Hopkins University Press, 2000); Joel Isaac, “The Human Sciences in Cold War America,” *The Historical Journal* 50, no. 3 (September, 2007): 725-746; Zuoyue Wang, *In Sputnik’s Shadow: The President’s Science Advisory Committee and Cold War America* (New Brunswick: Rutgers University Press, 2008); Theodore M. Porter, “Positioning Social Science in Cold War America,” and Hamilton Cravens, “Column Right, March! Nationalism, Scientific Positivism, and the Conservative Turn of the American Social Sciences in the Cold War Era,” in *Cold War Social Science: Knowledge Production, Liberal Democracy, and Human Nature*, eds. Mark Solovey and Hamilton Cravens (New York: Palsgrave MacMillan, 2012): ix-xv, 117-135; Rebecca Slayton, “From a ‘Dead Albatross’ to Lincoln Labs: Applied Research and the Making of a Normal Cold War University,” *Historical Studies in the Natural Sciences* 42, no. 4 (September, 2012): 255-282; Mark Solovey, *Shaky Foundations: The Politics-Patronage-Social Science Nexus in Cold War America* (New Brunswick: Rutgers University Press, 2013); and Jamie Cohen-Cole, *The Open Mind: Cold War Politics and the Sciences of Human Nature* (Chicago: University of Chicago Press, 2014)

Part of the story about postwar social scientists’ involvement in the postwar national security state is their eventual discomfort, disinvestment and public marginalization; see in particular, for a case study on the discipline of sociology’s trajectory, David Paul Haney, *The Americanization of Social Science: Intellectuals and Public Responsibility in the Postwar United States* (Philadelphia: Temple University Press, 2008).

resources to secure geopolitical advantage.¹²⁸ Early postwar definitions of brainpower established it as a national resource—the innate “ability of the human brain to process raw facts and impressions into creative thought,” as the educationist J. W. Brouillette put it—that could be refined and used in the workforce for the benefit of both state and society.¹²⁹ As industrial leaders inserted themselves into discussions of postwar power sharing, however, brainpower also became defined in terms of profit. Speakers at the 1947 National Electronics Conference, for example, made the case that public funding alone for scientific research would never lead the United States to greatness: only free market incentives could spur the best brains into action.¹³⁰ A decade later, participants at the same conference would frame the national brainpower pool as more than human output alone, but “the eventual vastly greater part contributed by synthetic intelligence devices and systems.”¹³¹ Industrial scientists, meanwhile, expanded the idea of brainpower throughout the early Cold War to include not only the trained brains that made

¹²⁸ A note on usage: one indication, however slight, that ‘brainpower’ gained popularity as a concept was its use vis-à-vis slightly different renditions of the word—namely “brain power” and “brain-power.” This elision—analogue to how “electronic mail” became “e-mail” and then “email”—suggests a growing willingness to accept a concept as an entity unto itself rather than a variation on a theme. In this sense, ‘brain-power’ or ‘brain power’ would suggest that brains could generate power in a fashion similar to and similarly measurable as other sources of power (e.g., ‘steam-power’, ‘horse-power’, etc.), whereas ‘brainpower’ would indicate the unique directions humanity could go based on the development of brains, whereas Google’s NGram Viewer gives some sense of how “brainpower” became increasingly popular in usage throughout the twentieth century. While “brain-power” had crested in its popular usage around 1880, and “brain power” peaked around 1920, “brainpower” began to spike in usage in the early 1950s. Although “brain power” remained the most popular phrasing in subsequent decades, its usage was eclipsed by “brainpower” by the turn of the twenty-first century. For the ngram model, see: `<iframe name="ngram_chart" src="https://books.google.com/ngrams/interactive_chart?content=brainpower%2Cbrain-power%2Cbrain+power&year_start=1800&year_end=2010&corpus=15&smoothing=3&share=&direct_url=t1%3B%2Cbrainpower%3B%2Cc0%3B.t1%3B%2Cbrain%20-power%3B%2Cc0%3B.t1%3B%2Cbrain%20power%3B%2Cc0" width=900 height=500 marginwidth=0 marginheight=0 hspace=0 vspace=0 frameborder=0 scrolling=no></iframe>`

¹²⁹ J. W. Brouillette, “Brain Power in Education.” *Peabody Journal of Education* 23, no. 2 (September 1945): 77, 78. See also: Vannevar Bush, *Science: The Endless Frontier* (Washington, D.S.: U.S. Government Printing Office, 1945).

¹³⁰ “Electronics Pool of Brains is Urged” *New York Times* November 4, 1947: 23.

¹³¹ Simon Ramo, “Space or Electronics—which Will Dominate the Century?” in *Proceedings of the National Electronics Conference*, vol. XIV (Chicago: National Electronics Conference, Inc., 1959), 16.

machines, but also the trained machines that freed up time and energy for trained brains.

Brainpower, as a national security resource, came to be portrayed as a profitable creative force with technological byproducts that could exist independent of human action.

Bureaucrats within the Office of Education shared the belief that, if monitored carefully, education could play a key role in preserving postwar democracy. In an address to the National Council for the Social Studies, Office of Education Commissioner John Studebaker asserted that education served as a “trustee for posterity,” helping “preserve, protect, develop, and transmit to each succeeding generation the glorious heritage of freedom and democracy that is ours as a people.”¹³² Communism, as democracy’s polar opposite, threatened to undermine American stability and prosperity. The key to democratic vitality, as Studebaker maintained, was through education system that, among other benefits, gave youth “a true appreciation [...] of their glorious American heritage of freedom wrung from the bitter struggle of centuries” and “an understanding and appreciation of the ethical and spiritual values, as well as the material benefits, of the American Way of Life.”¹³³ Tapping into progressive educational sentiment from the early twentieth century, Studebaker declared that revitalized civics education would “inspire [youth] with the resolve and with the zeal to do their full part in helping to improve the working of democracy”¹³⁴

Yet some within the Office of Education also worried that the very hallmark of democratic American education—local control—could also be its greatest threat.

Superintendents often ran school districts with little mind for democratic process, and

¹³² John W. Studebaker, “Communism’s Challenge to American Education,” address to National Council for the Social Studies, St. Louis, November 26, 1947, Record Group 12, Office Filed of the Commissioner of Education, Box 13, National Archives, College Park, MD.

¹³³ Studebaker, “Communism’s Challenge to American Education,” 4, 5.

¹³⁴ Studebaker, “Communism’s Challenge to American Education,” p. 5 (emphasis in original).

researchers within the office worried that students who encounter autocratic governance in schools would favor autocratic rule as adults. Dismayed that decentralized education could generate citizens unable to “withstand any form or type of insidious foreign ideology,” the office created a report in 1949 outlining seven principles that defined democratically run school districts.¹³⁵ These principles reinforced a functionalist view of society: A democratic society depended upon school districts in which roles for every member, from the superintendent to teachers to the community to grade school students, were clearly delineated and each member used their role to promote democratic principles. To ensure readers understood the importance of democratic process, the report included a checklist that school officials could self-administer to determine if their district lay prone to autocratic rule. Because the Office of Education could not directly enforce administrative procedure—and such centralized dictates would be in themselves anti-democratic—the subdepartment could only provide behavioral guidance for school boards and superintendents that would ideally reverberate in the classroom.

Because the federal government lacked the capability or authority to identify brainpower by itself, it had to rely heavily on external standardized testing research and development. The state found its greatest supplier in Educational Testing Service (ETS), which would soon design a wide array of standardized tests for the government, military, higher education institutions, secondary schools, and occupational licensing boards, among others. ETS spawned from the merger of three key testing and educational organizations—by the American Council on Education, the College Entrance Examination Board and the Carnegie Foundation for the Advancement of Teaching—and was granted an absolute charter by the New York State Board of Regents in December 1947. ETS was not an educational institution in itself, but the Board of

¹³⁵ Edwin H. Miner, Foreword to “Some Criteria for Evaluating American Democracy in School Administration,” by Clare B. Cornell, Office of Education, Federal Security Agency, January 1949, RG 12, Box 13.

Regents had the ability to grant such charters to not-for-profit corporations.¹³⁶ Although the corporation held its main office in midtown Manhattan, ETS also quickly opened offices in Princeton, New Jersey, and Berkeley, California, establishing a cross-country presence even in its infancy. More than any other testing organization, ETS provided the federal government and defense industries the testing tools required to harness Cold War brainpower, making standardized testing much more than a pedagogical practice: it was a Cold War ideological tactic. ETS, in turn, operated less like a nonprofit and more like a state quasi-apparatus.

ETS absorbed a complicated network of new responsibilities from its parent companies. While the American Council on Education relinquished control over its Cooperative Test Service, National Teacher Examination, and American Council Psychological Examination, the Carnegie Foundation gave ETS oversight of the Graduate Record Office. The College Entrance Examination Board, meanwhile, made ETS responsible for all of its testing services.¹³⁷ (ETS devoted a great deal of its early time and materials to the tests it inherited from the College Entrance Examination Board; by 1950, 75,000 students sat for at least one College Board exam.¹³⁸) Bolstered by a \$750,000 grant from the Carnegie Corporation of New York, these parent companies also provided financial backing for the newborn corporation.¹³⁹ Although ETS took over testing research and development from its parent organizations, they still held

¹³⁶ “New Service Gets Regents’ Charter,” *New York Times*, December 20, 1947; “Our Heritage,” Educational Testing Service, <https://www.ets.org/about/who/heritage/> (last accessed May 28, 2017). For more on New York State’s Regents charters for educational nonprofits, see: “Regents Charters and Regents Certificates of Incorporation,” New York State Education Department Office of Counsel website, http://www.counsel.nysed.gov/charters/regents_charters; and “Selected Sections of the Education Law Relating to Education Corporations (Section 210, Sections 214-219), New York State Education Department Office of Counsel website, <http://www.counsel.nysed.gov/pamphlet9/appenda210#a216> (both last accessed May 28, 2017).

¹³⁷ “School News Digest,” *The Clearing House* 22, no. 6 (February 1948): 367, 382.

¹³⁸ Educational Testing Service, *Annual Report to the Board of Trustees, 1949-50* (Princeton: Educational Testing Service, 1950), 42

¹³⁹ Lawrence P. Blum, “Research News and Communications,” *Journal of Educational Research* 41, no. 7 (March, 1948): 560.

considerable sway over the new organization's early leadership and mission. The chairmen of each parent organization not only served as *ex officio* members of ETS's twelve-member Board of trustees, but also appointed the remaining nine board members. Each parent organization also mandated that ETS not encroach upon their non-testing activities, such as the College Board's mandate to set higher education policy guidelines.¹⁴⁰ Beyond improving certain functions of existing testing organizations, ETS would have to develop its own distinct mission to the postwar state.

Through its earliest annual reports to its board of trustees, Educational Testing Service developed the rhetoric it would use to describe its mission to the both the public and the state. In the corporation's first report, published in 1949, ETS President Henry Chauncey stressed how misguided the general American public was about the costs of test manufacture. Although early developers had historically provided psychological tests for nominal costs, the days of small-scale test-craft were over. Americans, by Chauncey's judgment, were shortsighted in their view of modern testing, assuming that test developers could still eat the costs of development and operations while providing an increasing number of services. Charging more for tests was not only a matter of recouping expenses for the extensive time and effort put into test development, but also a reflection of necessary expenses in postwar society.¹⁴¹ The company may not have been publicly traded, "but ETS, though non-profit, is not eleemosynary. Tests cannot be provided at a nominal cost. Fair and reasonable prices must be charged."¹⁴² This proclamation, stentorian and graceless, would define the relationship between ETS and the postwar public:

¹⁴⁰ "The Educational Testing Service," *High School Journal* 31, no. 2 (March-April, 1948): 95.

¹⁴¹ Educational Testing Service, *Annual Report to the Board of Trustees, 1948-1949* (Princeton: Educational Testing Service, 1949), 36.

¹⁴² Educational Testing Service, *1948-1949 Annual Report*, 36.

because ETS defined test users as the organizations who collaborated with the corporation and purchased its tests, its primary role was in increasing institutional efficiency. ETS served the public to the extent that it provided services to organizations in instances where the state could not (or lacked the means) to step in. Their primary responsibility was thus to the American state and not its public.

By the following annual report, Chauncey seemed even more surefooted about his nonprofit's sense of mission. Likening the preceding year to a naval "'shakedown cruise,'" Chauncey asserted that ETS had weathered the turmoil of consolidation and was geared to engage in long-range planning—most notably, creating "an articulated series of test batteries," standardizing a useful system of scores and norms, and developing new types of aptitude and personality tests.¹⁴³ The nonprofit's ambitions, the president declared, were crucial for American vitality. To retain its strength, the United States had to "devise ways of utilizing the nation's human resources skillfully and with a minimum of wastage."¹⁴⁴ For this effort, Educational Testing Service sought to develop a "census of human abilities."¹⁴⁵ Such a census would serve as "an inventory of the characteristics required for the performance of important tasks," and "bring to light the aptitude areas in which our talent shortages are most critical and would provide a basis for realistic planning on a nationwide scale, in peace or war."¹⁴⁶ By building off of preexisting test batteries, Educational Testing Service could thus harness American brainpower into the most effective pathways, regardless of geopolitical climate. This census of

¹⁴³ Educational Testing Service, *1949-1950 Annual Report*, 7,8.

¹⁴⁴ Educational Testing Service, *1949-1950 Annual Report* 11.

¹⁴⁵ *Ibid.*

¹⁴⁶ *Ibid.*

abilities—much like other national censuses—was not an instrument for Americans to understand themselves, but for the American state to extract resources.

Sensing the frightening potential implications for a census of human abilities, Chauncey softened this imagery by discussing the ideal uses for standardized testing. Chauncey assured readers that standardized tests should not be “considered the end-all and be-all” for any course, and that teachers should “not feel under compulsion to adapt their courses to the content of the test.”¹⁴⁷ Indeed, Chauncey asserted that “by providing teachers with folios of test items,” ETS could provide teachers the choice (or at least, the semblance of choice) in testing material while ensuring students would be assessed on a breadth of skills.¹⁴⁸ Finding “the field of testing [...] exceedingly broad and the opportunities for ETS almost limitless,” Chauncey posed ETS’s mission in urgent, optimistic, terms.¹⁴⁹ Giving ETS the widest possible mandate—asserting “it is appropriate for ETS to use its facilities in any situation where tests and related techniques can be of assistance”—Chauncey staked the nonprofit’s mission as both educational and governmental.¹⁵⁰ This mission held faith that testing could “aid in the guidance of students and in their self-understanding.”¹⁵¹ Students were not given tests in order to foster self-actualization, but instead, to demonstrate their worth to the state: testing broadly and frequently, could help channel youth into educational tracks, from elementary to post-secondary testing.

Educational Testing Service asserted its mission included measuring abstract skills useful for both the classroom and the workplace, and by its second year, the organization began

¹⁴⁷ Educational Testing Service, *1949-1950 Annual Report* 12

¹⁴⁸ Educational Testing Service, *1949-1950 Annual Report* 13.

¹⁴⁹ Educational Testing Service, *1949-1950 Annual Report*, 14.

¹⁵⁰ *Ibid.*

¹⁵¹ *Ibid.*

developing a battery of guidance tests for high school and college students that blurred the lines between measuring academic and non-academic qualities. Building off existing standardized tests, ETS aimed to create a battery for gauging “permanent aptitude and achievement factors,” including typical examinations for “verbal, quantitative, spatial, mechanical, visualization, and induction reasoning skills,” as well as tests for carefulness, “perceptual speed, finger dexterity, eye-hand coordination, numerical computation, fluency of expression, [and] speed of judgment.”¹⁵² ETS supplemented this quest for an omnibus guidance battery with plans to design individual tests for aspiration, overstatement, and persistence. For these single-topic exams, ETS detached concepts from strictly educational parameters, instead defining aspiration as “the way an individual adjusts his goals following success and failure on various tasks,” and persistence as the “tendency to continue a task despite monotony, fatigue, or discouragement.”¹⁵³ As with the guidance battery, these exams focused less on students’ self-actualization than with assessing their readiness for the labor market—for determining how well a young adult would function within a workplace.

The Korean War heightened the relationship between federal defense objectives and higher education institutions—and also confirmed ETS’s function as a broker for essential state needs. Weeks after the war began, U.S. Commissioner of Education Earl J. McGrath sent letters to higher education administrative officers, assuring them that, while the National Security Resources Board had no set plans at the moment, advisory committees had already been assembled to consider “the part education is to play in the present emergency.”¹⁵⁴ By early

¹⁵² Educational Testing Service, *1949-1950 Annual Report*, 22

¹⁵³ Educational Testing Service, *1949-1950 Annual Report*, 28, 29.

¹⁵⁴ Earl J. McGrath, letter to chief state school officers, August 3, 1950; see also: Earl J. McGrath, letter to administrative offices of Institutions of Higher Education, July 26, 1950; and R.W. Gregory, letter to state directors of Vocational Education re: “Vocational education in the national emergency,” August 3, 1950; all in RG 12:

August, Selective Service System Director Lewis Hershey drafted working guidelines for student draft deferrals: men with at least sophomore standing who had planned to enroll for subsequent coursework in a timely fashion, and whose “scholastic standing placed him among the upper half of his class” qualified to have their enlistment in the armed forces postponed.¹⁵⁵ But the parameters for collegiate draft deferment, disseminated to higher education administrators in the *Defense Information Bulletin*, continued to change throughout the early part of war, as the state continued to recalculate which human resources were too vital to the American brainpower needs to send into conflict.¹⁵⁶ By September, 1950, the Selective Service expanded educational deferment quotas to include “male optometry, premedical, preosteopathic, preveterinary, preoptometry and predental students.”¹⁵⁷ McGrath stressed such deferments were “an effort to carry out the desire of the Congress to provide the fullest possible utilization of the Nation’s technological, scientific, and other critical manpower resources.”¹⁵⁸ Men in college—particularly those on scientific fields of study—were a privileged class, as their trained skills turned them into a valuable human resource necessary for American Cold War military efforts.

Standardized tests were central to how the state sought to preserve brainpower through draft deferments. Psychologists argued that the Army General Classification Tests (AGCT),

Records of the Office of Education, Office Files of the Commissioner of Education, 1939-1980, Box 15, National Archives, College Park, MD.

¹⁵⁵ Lewis B. Hershey, “Operations Bulletin No. 1,” August 8, 1950, RG 12, Box 15.

¹⁵⁶ Earl J. McGrath, “Emergency Supplement No. 2,” August 17, 1950, RG 12, Box 15.

¹⁵⁷ Earl J. McGrath, “Amendment to Selective Service Act of 1948,” *Defense Information Bulletin* 6, September 11, 1950, RG 12, Box 15.

¹⁵⁸ Earl J. McGrath, “Draft Regulations Affecting College Students,” *Defense Information Bulletin* 9, October 3, 1950, RG 12, Box 15. See also: Earl J. McGrath, “New Basic Policy for College Student Enlistments,” *Defense Information Bulletin* 16, January 26, 1951, RG 12, Box 15. Through these bulletins, McGrath makes clear the difference between postponement—in which a man may delay his induction into the Armed Forces until the end of an academic year—and deferment, which involved a separate military classification and potential renewal upon expiration of the deferment period.

which had been designed to sort millions of men by intellectual capabilities during the Second World War, could be used a template for determining which segments of the population would be most useful to the nation as trained brains. As researchers maintained, those with the most socially valuable careers—and presumably the greatest amount of brainpower—tended to score in the uppermost level on the AGCT. Why not, then, administer a similar scholastic aptitude exam to collegiate men to determine who should have their military induction postponed?¹⁵⁹ This line of reasoning gained official state support after President Truman issued an executive order in late March, 1951 authorizing the Director of Selective Service “to prescribe such qualification test or tests as he may deem necessary” to ensure that a “registrant's activity in study may be considered to be necessary to the maintenance of the national health, safety, or interest.”¹⁶⁰ Citing insufficient manpower reserves in comparison to Soviet capabilities, the Selective Service believed it had to make the most of “our superiority in scientific and technical know how” by using the “approximately 1,000,000 male, nonveteran students in colleges at present” in the most efficient manner.¹⁶¹

The Selective Service entrusted Educational Testing Service with creating a flexible sorting device that would maximize both military manpower and American brainpower. The Selective Service asserted ETS was a trustworthy gatekeeper because it was a “nonprofit, nonstick organization [. . . which gave] more than 1,100,000 tests during a typical year,” many to governmental agencies—in other words, it was already too big and too entwined in the workings

¹⁵⁹ E. Lowell Kelly, “The Facts,” in “Plan for Brainpower,” *The Science News-Letter* 58, no. 27 (December 30, 1950): 421-422.

¹⁶⁰ Executive Order 10230, 1, (b), (3); 1, (c), 31 March, 1951, ; text available <https://www.trumanlibrary.org/executiveorders/index.php?month=&year=1951>

¹⁶¹ Earl J. McGrath, “Plans for a Nation-Wide Test by the Selective Service Systems,” *Defense Information Bulletin* 29, April 2, 1951, RG 12, Box 15. See also: Earl J. McGrath, “Tentative Projection of Number of Full-Time Male Undergraduate Enrollees and Four-Year Male College Graduates, 1950-51 to 1954-55,” *Defense Information Bulletin* 21, February 5, 1951, RG 12, Box 15.

of the state to avoid.¹⁶² On April 2, 1951, the Selective Service System announced ETS's test for determining which college men would receive deferments: the Selective Service College Qualification Test (SSCQT). Rather than thin out college enrollments, ETS created the SSCQT under the premise that most college men would possess a level of scholastic aptitude consistent with average scores for college men on earlier standardized tests.¹⁶³ By standardizing the SSCQT for a general population, most draft-eligible college men would conceivably clear the minimum threshold for academic deferment. What would fluctuate, then, was not the quality of men in college but the degree to which the American state needed to trade off brainpower development for military manpower demands. Although the considerable majority of collegiate men were already ineligible for the draft—having served in the Second World War, disqualified by their 4F qualification, or already bound to service through their enrollment in the Reserve Officers' Training Corps—the SSCQT standardized the level of brainpower the state deemed worthy for postponing military service.¹⁶⁴ Responding to criticism that this deferment system undemocratically privileged a select few, U.S. Commissioner of Education Earl J. McGrath

¹⁶² McGrath, "Plans for a Nation-Wide Test by the Selective Service Systems," 2.

¹⁶³ "One More Year of College," *Science News-Letter* 59, no. 14 (April 7, 1951): 211. The SSCQT was modeled off of the AGCT, but the test's developers created a new scoring scale. A score of 120 on the AGCT was equivalent to a 70 on the SSCQT—but both scores indicated that a test-taker performed one standard deviation above the mean, or in the top 16% of all performances. Although changing the scoring scale isn't entirely cosmetic; this maneuver helps signal to the public that a new test has a distinct purpose—in this case, determining deferment eligibility among collegiate men. See: Warren G. Findley, "The Selective Service College Qualification Test," *American Psychologist* May, 1951: 181-183; and William G. Mollenkopf, "Development and Application of Tests of Special Aptitude," *Review of Educational Research* 23, 1 (February, 1953): 33-34.

¹⁶⁴ Earl J. McGrath, "Test Scores and Class Standings Required for Deferment of College Students," *Defense Information Bulletin* 32, April 6, 1951; Earl J. McGrath, "ODM Statement on Deferment of College Students," *Defense Information Bulletin* 33, April 6, 1951; Earl J. McGrath, "Issues in Current Discussions of Deferment of Students," *Defense information Bulletin* 34, April 6, 1951: 3; Earl J. McGrath, "Further Details of College Student Deferment Plan," *Defense information Bulletin* 36, April 12, 1951; Earl J. McGrath, "Deadline for Selective Service College Qualification Test," *Defense Information Bulletin* 39, May 2, 1951; Earl J. McGrath, "Reports on Professional Students to the Selective Service System," *Defense Information Bulletin* 44, June 5, 1951; and Earl J. McGrath, "Educational Provisions of the Universal Military Training and Service Act," *Defense Information Bulletin* 46, June 19, 1941; all RG 12, Box 15.

insisted the fairest treatment involved using “the services of each man at his own highest level of competence.”¹⁶⁵ By administering an exam with a threshold that could be raised or lowered depending on military demands, the armed forces devised a method for the state to manage brainpower as it did other vital national resources.

ETS familiarized college students and administrators with the logic of the SSCQT by releasing an explanatory bulletin in advance of the first series of test dates. The bulletin instructed men who hoped to qualify for deferment to detach and send a postcard application (postage not included) to Princeton, New Jersey, complete with Selective Service Number and preferred examination locations. ETS would then sent applicants an admission ticket, for which men would sit for one of three test tests in May or June, 1951. ETS assured hopefuls that the SSCQT examined their “ability to read with understanding and to solve new problems by using [their] general knowledge,” foundational skills that were “necessary for success in fields which require advanced training.”¹⁶⁶ Because the company designed the SSCQT to measure scholastic aptitude rather than any specific field of college-level learning, sample questions focused on the ability to make inferences, draw comparisons, and summarize data. Because the textual content in SSCQT questions was arbitrary to the actual skills tested, test candidates would find little success mining reading passages and charts for later recall; the Armed Forces did not care whether a man knew about ancient philosophers or earthquakes, but they did care whether he could make sense of a short passage or grammar question involving some random topic. This type of official preparatory material had to fulfill two seemingly contradictory roles: reduce anxiety by assuring test-takers the exam measured overall thinking skills while, at the same time,

¹⁶⁵ McGrath, “Issues in Current Discussions of Deferment of Students,” 2.

¹⁶⁶ Educational Testing Service Selective Service Examining Section, *Selective Service College Qualification Test Bulletin of Information*, March-June, 1951, 8, RG 12, Box 15.

suggesting that there was no way to cram for such a test. Such an approach to SSCQT test familiarization, in which the test's designers gave potential test-takers just enough practice for a test one could not specifically practice for, would become the standard approach for test designers over the following several decades.

The armed forces tempered their use of the SSCQT as the means for determining draft deferments by also considering class rank. Collegiate men who tested poorly on the SSCQT but ranked high enough in their class could qualify for a deferment. How well a man needed to rank or test depended on his year in college. Because military officials presumed that graduate school required considerably higher scholastic aptitude than undergraduate training, graduating college seniors had to produce the highest SSCQT minimums for deferment. Freshmen, meanwhile, needed better class standings for deferments, as most men who successfully completed their first year of college tended to complete their degree: a deferment in this case would ensure a man had a chance to develop and hone his brainpower uninterrupted. But since the correlation between men's class rankings and their SSCQT scores were not exact—many displayed higher scholastic aptitude than their grades suggested, and others performed stronger in school than their test scores predicted—this dual provision allowed far more draft-eligible collegiate men to qualify for deferments. Except for the East South Central United States, at least 60 percent of draft-eligible men qualified for deferments either by their class standing or SSCQT score; for men majoring in engineering, physical sciences, mathematics, or biological sciences, the deferment qualification rate was greater than 70 percent.¹⁶⁷ Using class rank alongside SSCQT scores gave

¹⁶⁷ Henry Chauncey, "The Use of the Selective Service College Qualification Test in the Deferment of College Students," *Science* 116, no. 3004 (July 25, 1952): 78, 79. See also: Earl J. McGrath, "New Series of Selective Service College Qualification Tests is Announced," *Defense Information Bulletin* 54, September 27, 1951; Earl J. McGrath, "Selective Service Qualification Test Applications," *Defense Information Bulletin* 67, May 8, 1952; both RG 12, Box 15.

the armed forces greater latitude in preserving brainpower without discrediting the traditional and more straightforward method of assessing academic achievement.

By providing a uniform measure for determining merit, the SSCQT helped the American state reorient the relationship between education and civic duty. Because the number of men suitable for military service was always much smaller than the millions of men legally required to register for the draft, the armed forces had to reconsider the grounds upon which a man could legitimately seek deferment. The underlying logic of educational deferments—that a man’s education belonged to the nation before it belonged to him—made it easier for the state to back away from paternity deferments, which placed a man’s family duties before his civic responsibility. By allowing higher education deferments, the military could guarantee men with demonstrable brainpower would not be squandered on the battlefield, all while avoiding responsibility for the factors that prevented many black men and white southern men from attaining such deferments. Advocates of educational deferments conceded that, of course, the SSCQT or any other scholastic aptitude test would fail to detect all of the nation’s best brainpower: American society had already shuttered many young men’s educational opportunities.¹⁶⁸ The solution to this broader brainpower problem, psychologists and psychometrists maintained, was to design and administer many more tests.

Use of the SSCQT continued after the Korean War ended; see: Rall I. Grigsby, “Dates for Selective Service College Qualification Tests,” *Defense Information Bulletin* 81, September 25, 1953, RG 12, Box 15. However, ETS was not the sole administrator of the SSCQT; between 1954 and 1959, and briefly in 1966, Science Research Associates outbid ETS for a defense contract to provide the exam. See: Thomas J. Frusciano, “Student Deferment and the Selective Service College Qualification Test, 1951-1967” (ETS-RM-83-1), Princeton: Educational Testing Service, 1980: 43-46.

¹⁶⁸ By the middle of 1952, slightly over 1 million men eligible for the draft qualified for a Class III-A (dependency deferral) designation, while only about 200,000 qualified for a Class II-S (collegiate deferral) designation; see: U.S. Selective Service System, *Report of the Director of Selective Service for the Fiscal Year 1953 to the Congress of the United States* (Washington, D.C.: U.S. Government Printing Office, 1954): 9-21. See also: Educational Testing Service, *Statistical Studies of Selective Service Testing, 1951-1953* (Princeton: Educational Testing Service, 1955); and National Manpower Council, *Student Deferment and National Manpower Policy* (New York: Columbia University Press, 1952).

Test Early, Test Often: Guidance Counseling, Gifted Youth, and Scholarship Contests

Standardized testing gained traction at a moment when certain types of smartness came under suspicion. Although intellectualism had long existed in friction with mainstream white American society—which often extolled genius and ingenuity while maintaining an anti-elitist political culture and evangelical popular culture—Cold War politics added additional rationale for distrusting trained experts. Academics became objects of suspicion, and critics used the image of the egghead as shorthand for intellectual fecklessness: too smart for his own good, too caught up fussy theoretical details to see reality as it was, too fragile to get the work done himself. In the fervid torrent of McCarthyist anticommunist rhetoric, expertise became linked to femininity, homosexuality, blackness—and, above all, disloyalty. Brainpower, left unbridled, could be used to disastrous ends. Standardized test developers insisted that their products could be used to steer youth toward meaningful and productive lives in service to state and society. In effect, standardized tests could create loyal experts. The trick, these developers maintained, was ensuring school systems employed a robust testing program, so that guidance counselors received steady objective information about their charges. By testing early and often, school systems could cultivate a great deal of brainpower with enough built-in surveillance to allay the fear talented students would turn into eggheads.¹⁶⁹

¹⁶⁹ There is a lineage of research on American anti-intellectualism, both academic and popular. Most of it owes its existence to Richard Hofstadter's *Anti-Intellectualism in American Life* (New York: Alfred A. Knopf, 1963). For a glimpse at how anti-intellectualism informed early Cold War political culture, see: G. Eric Hansen, "Intellect and Power: Some Notes on the Intellectual as a Political Type," *Journal of Politics* 31, no. 2 (May, 1969): 311-328; Robert D. Cross, "The Historical Development of Anti-Intellectualism in American Society: Implications for the Schooling of African Americans," *Journal of Negro Education* 59, no. 1 (Winter, 1990): 19-28; K.A. Cuordileone, "'Politics in An Age of Anxiety': Cold War Political Culture and the Crisis of Masculinity, 1949-1960," *Journal of American History* 87, no. 2 (September, 2000): 515-545; Jessica Wang, "Scientists and the Problem of the Public in Cold War America, 1945-1960," *Osiris* 2nd series, 17 (2002): 323-247; Aaron S. Lecklider, "Inventing the Egghead: The Paradoxes of Brainpower in Cold War American Culture," *Journal of American Studies* 45, no. 2 (May, 2011): 245-265; and Dane S. Claussen, "A Brief History of Anti-Intellectualism in American Media," 97, no. 3 (June, 2011): 8-13.

Educationists encouraged guidance counselors to treat standardized testing as part of a holistic approach for steering youth toward fulfilling careers and productive citizenship. Aiming to quickly professionalize the growing national pool of school counselors, specialists created guidebooks to counsel practitioners on testing methods. These books often advised guidance counselors to learn foundational statistical concepts, as well as actually take the tests they administered to students; without a working knowledge of aptitude tests or test design, counselors could misuse aptitude test results to disastrous ends. Some guides even provided readers case studies. One 1950 guide led readers through the extensive casework of two fictional counselors. Using a combination of standardized tests and psychological insight, Miss Brandon and Mr. Erlandson helped dozens of maladjusted young adults overcome their dysfunctions. Whether guiding Charles to consider switching his engineering major to a pre-law track, or referring Jane to a reading specialist to improve her poor grades, Miss Brandon and Mr. Erlandson always incorporated results from standardized aptitude tests and vocational batteries into their diagnoses. As such guidebooks suggested, standardized test batteries and assessments—whether the Strong Vocational Interest Blank, Minnesota Vocational Test for Clerical Workers, Bell Adjustment Inventory, Bennett Test of Mechanical Comprehension, or Ruder Preference Record—provided insights that student’s grades or stated aspirations could not. Through these guidebooks, specialists established the idea that millions of Chucks and Janes were far more likely to find success and happiness when guidance counselors tempered their professional opinions with the wisdom of test scores.¹⁷⁰

¹⁷⁰ Shirley A Hamrin and Blanche B Paulson, *Counseling Adolescents* (Chicago: Science Research Associates, 1950, 15th printing 1965). See also: C. Gilbert Wrenn and Willis E. Dugan, *Guidance Procedures in High School: Some Recommended Procedures Based Upon a Survey of Present Practices in Minnesota* (Minneapolis: University of Minneapolis Press, 1950); Lester D. Crow and Alice Crow, *An Introduction to Guidance: Principles and Practices* (New York: American Book Company, 1951); Lester Nicholas Recktenwald, *Guidance and Counseling with Psychometric Principles: A Basic Treatment* (Washington, D.C.: Catholic University of America, 1953); Raymond N. Hatch and Paul L. Dressel, *Guidance Service in the Secondary School* (Dubuque:

Gifted students, who had the most potential brainpower to give to the state but were most prone to squander it, proved a particular obsession for guidance counselors, psychometric researchers, and standardized test designers. The College Entrance Examination Board (CEEB), one of the leading developers of academic standardized tests, collaborated with the National Science Foundation (NSF) in 1955 to analyze how the American education system could reduce the number of gifted and talented youth who avoided higher education. The resulting report, *Encouraging Scientific Talent*, distanced intelligence from intellectualism: smart youth did not have to grow up to become eggheads. Indeed, intellectual strength did not come at the cost of other abilities. CEEB noted uncommonly bright individuals—whether speaking about, in their terms, the categorically superior, intellectually gifted, or true geniuses—also typically possessed greater emotional poise and physical strength than their average-minded peers. But, the board warned, the talents of society’s most intelligent members also faded with age. Leaning on previous studies that examined when people in mentally-demanding careers made their greatest discoveries, CEEB suggested the American educational system kept scientifically-gifted individuals in school too long to contribute to society as much as they could have if left to progress through courses at their own pace. To CEEB, the gifted were precious social assets who “contribute[d] more than their share to the welfare of society”—but only if allowed to academically achieve.¹⁷¹ Whereas gifted youth with high levels of academic achievement became “industrious, patient, honest, and conscientious” adults, those who did not reach college

W.C. Brown Co., 1953); and Robert Henry Knapp, *Practical Guidance Methods for Counselors, teachers, and Administrators* (New York: McGraw Hill, 1953).

¹⁷¹ Charles C. Cole, Jr., *Encouraging Scientific Talent: A Study of America’s Able Students Who are Lost to College and of Ways to Attracting Them to College and Science Careers* (New York: College Entrance Examination Board, 1956), 28.

turned “autocratic, blustery, stubborn, and conceited.”¹⁷² For CEEB, the bigger danger was not a man with his nose in a book, but a legion of intelligent Americans soured by wasted potential.

To illustrate how many intelligent youth remained unmotivated and underserved by the American educational system, CEEB built its findings on standardized models of intelligence. Asserting “high intelligence is one of the most important requisites for a scientific career,” CEEB employed several IQ models to calculate how few qualified Americans fulfilled their full intellectual potential.¹⁷³ Using the most generous existing estimates for academic giftedness—an IQ of 125, with a model that estimated 6 percent of all young adults met or passed that threshold—CEEB determined that about 130,000 people born in any given year had the mental capability to attain a PhD. Yet just short of 9000 people had earned a doctorate in the 1953-1954 academic year. Unless those with stakes in the state of education made efforts to understand why relatively few highly intelligent American youth achieved as much as they could, the United States would continue to squander scientific brainpower.

To address these concerns, CEEB authorized a national standardized test and questionnaire to determine American high school students’ level of scientific aptitude and educational motivation. On behalf of CEEB, Educational Testing Service developed and disseminated the National Study of High School Students and Their Plans to over 32,000 twelfth-grade students at nearly 500 American high schools in early 1955. The survey contained two parts: a fifteen-minute, twenty question aptitude test, followed by a thirty-minute questionnaire about college plans. The aptitude test contained statistically reliable math and verbal questions from earlier Educational Testing Service examinations—and how well a student

¹⁷² Cole, *Encouraging Scientific Talent*, 30.

¹⁷³ Cole, *Encouraging Scientific Talent*, 22.

did on the first test determined whether Educational Testing Service would analyze the second. Out of the entire pool of survey respondents, Educational Testing Service focused on roughly 9,700 who had answered at least twelve of the aptitude questions correctly, assessing these students' college plans to see what factors currently discouraged them from higher education. For most high-scoring students, the cost of college remained a potential dissuading factor; over half of all seniors reported that school expenses could keep them from pursuing higher education. What the questionnaire also revealed, however, is that a great deal of high-performing students, particularly boys, could be swayed toward pursuing a degree in science or engineering if offered a scholarship. Based on this questionnaire, CEEB estimated that, among that year's 1,265,000 public high school seniors, somewhere between 55,000 and 90,000 highly intelligent youth would have attended college had scholarships been available.¹⁷⁴ More damning was CEEB's calculation that an additional 100,000 capable high school seniors "with similar ability lack the interest or motivation of college," regardless of any potential award program.¹⁷⁵ By pegging its nationwide survey to a standardized test, however brief, CEEB standardized not only the methods for diagnosing the cause of America's brainpower shortage but also the prescribed cure.

CEEB placed scholarships at the center of its final recommendations for boosting American scientific brainpower—and, in doing so, redoubled its case for more wide-scale standardized testing programs. For any national scholarship program to have any heft, CEEB stressed, it must first be supported by "more widespread early identification and encouragement of gifted students," particularly if such monitoring would allow individuals to complete their

¹⁷⁴ Cole, *Encouraging Scientific Talent*, 139-169.

¹⁷⁵ Cole, *Encouraging Scientific Talent*, 184.

education at an accelerated clip.¹⁷⁶ Although the board did not provide a specific timetable for tests, it noted that existing studies suggested aptitude and achievement tests could be successfully used to scout scientific talent as early as the eighth or ninth grade.¹⁷⁷ When paired with a network of well-trained guidance counselors, standardized tests could ensure that the scientifically and mathematical inclined are guided toward challenging classes.

Scholarships provided an ideal vehicle for promoting the complex joint partnership that emerged in the early Cold War hunt for brainpower. Scholarships represented the idea that an alliance between the public good and private interests was not just attainable, but ideal for American education. Any widespread scholarship program would require schools supported by federal and state funds, corporations willing to sponsor awards, and colleges and universities eager to have their talent search brokered. Above all, a widespread scholarship program would require all parties to accept the idea that the best minds could be scouted and sorted through a system of standardized tests—that the students who received scholarships had already been guided through their education by a series of standardized assessments and vetted for their award by a series of standardized tests.

The entities that would run postwar standardized test-based scholarship programs thus had to represent the interests of all stakeholders—public and private, corporate and academic—without fully belonging to any. The most notable example of these efforts, the National Merit Scholarship Corporation (NMSC), marked a shift in corporate educational philanthropy. By sponsoring scholarship programs, corporations could indirectly invest in higher education while

¹⁷⁶ Cole, *Encouraging Scientific Talent*, 186.

¹⁷⁷ Cole, *Encouraging Scientific Talent*, 28-31; Dael Wolfle, “Intellectual Resources,” *Scientific American* 185, no. 3 (September, 1951): 42-46. For relatively concurrent research on this topic, see also: Mary Frances Hodgson, “A Study of the Discovery and Development of Science Talent in Secondary School Pupils,” master’s thesis, Kansas State College of Agriculture and Applied Science, 1956.

also promoting business interests. Scholarship competitions' dependence on standardized tests allowed corporate philanthropists to skirt around core economic and racial factors plaguing the cultivation of American brainpower. These programs, in turn, influenced the meaning of democracy in discussions of postwar American higher education. Meritocratic language and logic—in which standardized test scores represented a student's investment risk rather than reflected systemic shortcomings—helped reinforce the idea of educational attainment as an individual effort. Within the logic of corporate scholarship programs, some students may have needed a leg up to attend college, but their test scores proved they already had the brainpower required for success.

Corporate scholarship programs did exist before the Cold War, including contests geared to promote collegiate scientific study among talented youth. These programs typically involved a single corporation having a direct role in the funding and selection of award recipients. One notable example, the Westinghouse Science Talent Search, was a large-scale production involving a giant corporation (Westinghouse Electric), its philanthropic wing (the Westinghouse Educational Foundation), and a nonprofit that carried out testing operations (Science Search). Westinghouse and Science Search launched the talent hunt in 1942, awarding scholarships to 40 high school students who displayed uncommon ingenuity, innovation, and workmanship—characteristics that could be made profitable through advanced scientific training. Before winning these scholarships, however, students first took a science aptitude test, conducted by Science Service and administered to participants through officially sanctioned extracurricular science clubs. This two and one-half hour aptitude test included questions on engineering, anatomy, biology, chemistry (“Which of the following is *not* a use of the mineral quartz?”),

science history, spatial visualization, and reading comprehension.¹⁷⁸ Science Service encouraged students to work through sample test questions printed in its publication *Science News-Letter*, even including an address to request the most recent Science Talent Search exam. Because the aptitude test had been designed to never elicit a perfect score, even by actual scientists, the organization felt no risk letting high school students toying with sample questions. Through the Science Talent Search, which it continued to sponsor throughout and beyond the Cold War, Westinghouse set a template for corporate scholarship funding. By relying on third-party aptitude test designers, companies could promote the shared interests of business and science while keeping the upfront cost of promoting talented youth to a relative minimum.¹⁷⁹ The public relations boon these annual searches generated for Westinghouse (and for advanced scientific training in general) greatly outweighed the cost of the actual scholarship, all while establishing standardized testing as a key method for quickly determining the very best in the talent pool.

Large-scale scholarship programs truly took off, however, when third-party developers promised to coordinate talent hunts on behalf of multiple corporations for the general promotion of higher education. Created in September 1955, the National Merit Scholarship Corporation (NMSC) provided the means for corporations to centralize their educational philanthropy. NMSC emerged from the joint efforts of the Ford Foundation and Carnegie Corporation, which together contributed millions of dollars for scholarship funds and administrative costs. Upon its creation, NMSC immediately became “the largest independent scholarship program in the

¹⁷⁸ Allen Long, “Science Aptitude Test,” *Science News-Letter* 65, no. 5 (January 30, 1954): 74; “Test Your Science Ability With Sample Problems,” *Science News-Letter* 55, no. 5 (January 29, 1949): 72-73, 78; Watson David, “Test Your Science Aptitude,” *Science News-Letter* 63, no. 4 (January 24, 1953): 58-59.

¹⁷⁹ Seven G. Terzian and Leigh Shapiro, “Corporate Science Education: Westinghouse and the Value of Science in Mid-Twentieth Century America,” *Public Understanding of Science* 24 (2): 147-166

history of American education.”¹⁸⁰ NMSC’s initial Board of Directors reflected an aim to portray the interests of higher education as those of big business, with presidents of Dartmouth, NYU, and the University of Colorado sitting alongside those of Goldman Sash, Standard Oil, and Detroit Edison Company. To incentivize businesses participation, NMSC appropriated 8 of its initial \$20.5 million in seed money toward corporate and individual donation-matching funds. Corporations who contributed to NMSC’s funds could retain naming rights for their scholarships while benefitting from an outsourced “professionally conducted national annual talent search.”¹⁸¹ Sears, Roebuck & Company became an early corporate sponsor, offering over a half-million dollars for 100 Sears Foundation Merit Scholarships designed to cover a complete undergraduate education; Time, Inc. and other corporate entities soon followed suit.¹⁸² Some companies, such as the Stewart-Warner Corporation and General Foods Corporation, teamed with National Merit to provide scholarships for specific paths of study—mechanical and electrical engineering in the case of the former and physical sciences in the latter.¹⁸³

The two organizations that initially funded NMSC shared a belief that business could help solve the problems plaguing American education, particularly where state and society fell short. Despite venturing into education-related projects only at the beginning of the decade, the Ford Foundation found its sense of mission in what it perceived as inadequate federal educational support. The Foundation asserted the federal budget—which apportioned just 2.4% of the gross national product to education—was wholly inadequate, especially at a moment when

¹⁸⁰ “News Briefs,” *Science* 122, no. 3168 (September 16, 1955): 508.

¹⁸¹ Carnegie Corporation of New York, *Annual Report for the Fiscal Year Ended September 30, 1954* (New York: Carnegie Corporation of New York, 1955), 34.

¹⁸² Ford Foundation, *Annual Report: October 1, 1954 to September 30, 1955* (New York: Ford Foundation, 1956), 19.

¹⁸³ “Grants, Fellowships, and Awards,” *Science* 122, no. 3181 (December 16, 1955): 1181-1182.

the number of children in American schools continued to grow with no projected end in sight. Even the most conservative demographic estimates indicated twice as many Americans would be enrolled in college between 1966 and 1971 than were enrolled in 1954—leaping from 2.5 to at least 5 million people—yet no long-term nationwide plans existed to expand the pool of American teachers and professors. Those who attended college would likely encounter professors whose wages, when adjusted for inflation, were lower in the mid-1950s than just before World War II.¹⁸⁴ To the Ford Foundation, a decentralized educational system could not operate without some form of guardianship.

To the Carnegie Corporation, far too many discussions about postwar education started from the misguided assumption that anyone who could go to college, should go to college. The organization maintained that this problem stemmed from a widespread superficial understanding of the relationship between intelligence and college readiness. College administrators and educational specialists had debated over the IQ needed for likely college success—108? 110? 120?—without giving much attention to the actual logic of probability. A young woman with an IQ of 115 may have been in the upper fifth of intelligence among the general population, but her heightened capabilities did not necessarily guarantee that she would eagerly complete her studies, or know exactly what field she wanted to pursue, or even want to go to college in the first place. Likewise, a man with an IQ of 80 was far less likely than the average college-goer to complete his studies, but intelligence test scores alone did not account for the multiple traits needed to successfully earn a degree. In either case, focusing too much on the likelihood

¹⁸⁴ The Ford Foundation, *The Ford Foundation Report for 1954: January 1, 1954-September 30, 1954* (New York: Ford Foundation, 1955): 1; Ford Foundation, *Ford Foundation Report: October 1, 1954 to September 30, 1955* (New York: Ford Foundation, 1956): 11-13.

someone with a certain IQ would complete college got in the way of actually “selecting those who can profit most from college.”¹⁸⁵

The Carnegie Corporation maintained that most criticism of contemporary education fell into a logical trap of confusing growing social expectations with genuine social good. Those who espoused equal educational opportunity often carried “the tendency arbitrarily to define college as a valuable experience”; insisting most people could use extended training, particularly vocational instruction, siphoned resources away from bright youth who would benefit from the type of liberal arts education that once served as the foundation of all higher ed.¹⁸⁶ Unless experts and advocates carefully balanced the need to train more young adults with the need to provide a rich and challenging education, other social forces and stakeholders would strip American higher education of its greatest potential good. Without any examination of the issues colleges and educators faced, the United States risked “a lowering of intellectual standards, the smothering of individuality, and the assembly-line production of half-educated men and women.”¹⁸⁷ As the corporation warned, a thoughtless approach to national higher education would create an “enthronement of the ‘average’”—and only encourage democracy’s worst tendencies.¹⁸⁸

NMSC used standardized testing as its basis for scouting, assessing, and rewarding academic talent. To determine who would ultimately receive a National Merit scholarship, the corporation first proportionally allocated the number of scholarships available by state. The

¹⁸⁵ Carnegie Corporation of New York, *Annual Report for the Fiscal Year Ended September 30 1954* (New York: Carnegie Corporation, 1955), 12.

¹⁸⁶ Carnegie Corporation, 1954 *Annual Report*, 13.

¹⁸⁷ Carnegie Corporation of New York, *Reports to the Officers for the Fiscal Year Ended September 30, 1953* (New York: Carnegie Corporation, 1954), 11.

¹⁸⁸ Carnegie Corporation, 1954 *Annual Report*, 16.

corporation directed each participating school—over 10,000 in the competition’s first year—to allow the top 5 percent of its senior class to take a screening test. The corporation then used the results from this screening test to pare down the applicant pool to ten times the number of available scholarships. These remaining students were then invited to take the SAT. Based on semifinalists’ SAT scores, NMSC selected a pool of finalists—and after collecting their biographical information, school grades, and recommendations, the corporation decided who would receive an award.¹⁸⁹ From the 58,000 students who participated in the initial talent hunt, NMSC chose nearly 5,100 semifinalists and granted 555 National Merit Scholarships.¹⁹⁰ NMSC’s approach to awarding scholarships normalized standardized testing by making it a participatory, incentive-driven activity that revealed more to colleges than academic and biographical information could alone.

Although the corporation used test scores to determine who received scholarships, the actual amount of the scholarship depended on the perceived need of the student. Whereas some students received full coverage for four years of tuition and expenses, others received considerably smaller stipends. In either case, the corporation made payments to colleges and universities rather than to the students.¹⁹¹ By making the institution the point at where education was subsidized, the corporation also increased the incentive for universities to get on board with the program. Schools would not have to pay out of pocket to receive the best and the brightest,

¹⁸⁹ “News of Science—National Merit Science Program,” *Science* 122, no. 3168 (September 16, 1955): 508.

¹⁹⁰ “Merit Scholarships,” *Science* 123, no. 3195 (March 23, 1956): 499; “NMSC History and Facts,” National Merit Scholarship Corporation website, <http://www.nationalmerit.org/s/1758/interior.aspx?sid=1758&gid=2&pgid=451> (last accessed August 3, 2016).

¹⁹¹ “Merit Scholarship Corporation,” *Science* 125, no. 3249 (April 5, 1957) 639-640.

thus encouraging them to participate in the program and, in time, actively spend resources on courting national merit finalists.

While some early Cold War scholarship programs seemed to promote vastly different agendas than the National Merit program, many used standardized testing to reward students who, with college training, could boost American prowess while strengthening conservative social relations. Beginning in 1955, General Mills Foods administered the Betty Crocker Homemaker of Tomorrow Scholarship program, which aimed to “to focus national attention on the so-called ‘forgotten career’ of homemaking.”¹⁹² Like the National Merit program, the Betty Crocker Scholarship used a standardized test—in this case, a 50-minute, 150-question multiple-choice exam—to determine winners at statewide and national levels. By providing young women the means to pursue college studies that would promote the science of homemaking, the Betty Crocker Scholarship pursued an objective similar to other corporate scholarship providers: harnessing brainpower to place men and women in positions that would ensure constant American progress with minimal social friction. Corporations used their scholarships to improve existing social structures—never to dismantle them.

Even so, some women awarded Betty Crocker scholarships didn’t conform to the social expectations they believed these prizes symbolized. Many women (including Massachusetts Senator and 1966 Betty Crocker scholarship recipient Elizabeth Warren) recalled the scholarship program as a pragmatic means to get into college when other methods weren’t possible, even if they had no personal interest educational experience in home economics. As with those teens who did well on the SAT, some of the women recalled simply being good at standardized tests.

¹⁹² Grant Moos, “Remembering Betty Crocker Homemakers of Tomorrow,” *General Mills Taste of General Mills Blog*, April 9, 2013 (accessed May 18, 2016), <http://blog.generalmills.com/2013/04/remembering-betty-crocker-homemakers-of-tomorrow/>

Their test savvy proved the key to success, suggesting that standardized scholarship tests not only rewarded those who showed promise in a particular field of knowledge, but also those whose cleverness revealed a different, albeit unintentional, set of abilities. Although some corporate scholarship programs used standardized tests to help reinforce a particular worldview, many young Americans used the tests as an opportunity to springboard to a different, untraditional life path.¹⁹³

Realizing their award recipients could be treated as an unparalleled data set, the National Merit Corporation tracked the life paths of its earliest award recipients in order to see if their scholarships allowed the gifted to reach their greatest potential. In 1968, the organization released a study on the educational trajectories of its first two cohorts of scholarship recipients. Even among these men and women—ostensibly the most gifted young adults the nation had to offer in the mid-1950s—educational disparities persisted. Eight years after receiving National Merit scholarships, 6 percent of men had yet to receive a bachelor’s degree, and another 13 percent had completed yet to complete any graduate work. Although these rates of college completion were drastically higher than the national average, the corporation mined into the backgrounds of award recipients to see what factors prevented these men and women from advanced studies—in other words, what issues were beyond the company’s responsibility.

¹⁹³ Moos, “Remembering Betty Crocker Homemakers of Tomorrow”; Susan Marks, “Betty Crocker Search for the American Housewife of Tomorrow Scholarship,” *Finding Betty Crocker* blog October 26, 2008 (accessed May 18, 2016), <http://findingbettycrocker.blogspot.com/2008/10/betty-crocker-search-for-american.html>; Claire Davidson, “Snapshot: Future Homemakers of Tomorrow,” Betty Crocker website, no date given (accessed May 18, 2016), <http://www.bettycrocker.com/menus-holidays-parties/mhplibrary/parties-and-get-togethers/vintage-betty/snapshot-future-homemakers-of-tomorrow>, Leslie Ann Jones, “Interview with a One-Time ‘Betty Crocker Homemaker of Tomorrow,’” *the hairpin*, June 13, 2013 (accessed May 18, 2016), <http://thehairpin.com/2013/06/interview-with-a-one-time-betty-crocker-homemaker-of-tomorrow/>; and Suzanna Andrews, “Elizabeth Warren: A Warrior for the Middle Class,” *More* May 2014, available online at: <http://www.more.com/news/womens-issues/elizabeth-warren-warrior-middle-class>

The comments for Marks’s blog post reveal, in particular, the way many women (and, after rule changes in the early 1970s, some men) used the scholarship opportunity to gain tuition and board money for schools they would have otherwise never been able to attend.

Among women who received National Merit scholarships in 1956 and 1957, nearly one-third had yet to advance beyond undergraduate work by 1964. Gender cross-cut socioeconomic patterns: women who received scholarships typically came from better-educated families than men in their cohort, but women who continued onto advanced studies received considerably higher stipends than those who didn't—a pattern reversed among men. Whereas women who left their studies to become homemakers overwhelmingly indicated a desire to continue school, the several dozen men who had yet to receive a bachelors degree “appeared to have relatively more than their fair share of emotional problems.”¹⁹⁴ Through this longitudinal data, the Corporation was able to hold 95% as a natural statistical limit for the proportion of National Merit recipients who would complete college, all while posing broader social forces as the quirks of individuals and their families. The corporation remained as concerned with mitigating responsibility for scholarship recipients' failure to achieve full their potential as it was with measuring what, exactly, a student could achieve after receiving the sudden advantage of a National Merit scholarship.

Critics lambasted the National Merit Scholarship competition and similar business-backed talent hunts for their perpetuation of class and racial inequities. Mapping the location of National Merit winners in several states, educationist Horace Mann Bond concluded that award recipients did not reflect the truest peak of young America's talent but simply the brightest among whites of some considerable means. As Bond argued, the scholarship company ignored the sociocultural and economic factors that would allow well-educated middle-class white children to succeed. By allocating scholarships on a state-by-state basis, National Merit did nothing to discourage policies that limited educational equity for blacks and poor whites—and by maintaining the scholarship was awarded on the basis of talent, only reinforced the idea that

¹⁹⁴ Donovan J Watley, “Career Progress of Merit Scholars” (Evanston, IL: National Merit Scholarship Corporation, 1968): 18.

blacks were inherently less intelligent than their white counterparts. This idea would gain even more steam, Bond feared, as federal programs aimed to replicate the National Merit model of nationwide scholarship programs allocated on a state-by-state basis. Bond insisted the same results for a national talent-search could be achieved by using a different “‘diving rod’ for ‘talent’”¹⁹⁵: the average number of working toilets per household. For Bond, the competition “‘might better be called ‘States Rights, Upper Class’ Scholarships’”—a criticism that did not disappear as the corporation attempted to highlight talented African-American youth through the National Achievement Scholarship program.¹⁹⁶

Other critics of the National Merit program believed standardized testing would stifle society by relying on an overly narrow idea of talent. Analyzing the limitations of his own field, psychologist David McClelland maintained existing standardized testing devices were only equipped to measure “‘academic excellence, skill in taking examinations, in following instructions and finding solutions to problems set by others.’”¹⁹⁷ Although this type of ability was needed for success in higher education, McClelland worried the growth of educational testing would only encourage Americans’ fondness for conformity. Widespread conformity would, in turn, discourage other kinds of extracurricular intelligence vital for American vitality, such as curiosity and entrepreneurship. For such critics, standardized testing programs stifled what higher education could provide by reducing talent to a specific quantifiable psychological concept.

Scholarship competitions—and the language of merit—offered a way for corporations to help individuals’ chances of attending college without ever having to disrupt or challenge the

¹⁹⁵ Horace Mann Bond, “Talent—and Toilets,” *Journal of Negro Education* 28 no 1 (Winter, 1959): 7.

¹⁹⁶ Bond, “Talent—and Toilets,” 12; “Deliver Us From the Racial Hypocrisy of the National Merit Scholarship Program,” *Journal of Blacks in Higher Education* 39, (Spring, 2003): 61-63.

¹⁹⁷ David C. McClelland, “Encouraging Excellence,” *Daedalus* 90, no. 4 (Fall, 1961): 713.

forces that kept many groups outside of the winner's circle. More than the rhetoric of the educational elite, the language of merit reflected the way corporations approached postwar educational philanthropy: the best prospects would make themselves apparent, and would succeed with the right investment. These scholarship programs reinforced the idea that standardized tests made excellent sorting devices, easily making clear the vast majority of applicants who merited no genuine attention while also revealing the very few with superlative aptitude and abilities. Although these programs used biographical information and school grades to make their final decision, this data always followed the test scores: one's identity was always a supplement to one's numbers. Test designers and scholarship funders often acknowledged—and in their philanthropic capacity, tried to undo—trenchant socioeconomic disparities, but many scholarship programs often failed to account for those discrepancies when considering award winners. The logic of merit, that the best students would receive the best training, reconciled the gaps between the narrow abilities the tests measured, the brainpower demands corporate philanthropists aimed to satisfy, and cultural clout bestowed upon young adults who received a college scholarship. Merit presented the sense that test-takers arrived at an examination at a level playing field: after all, if tests measured aptitudes, there's a limit to what accounting for external social differences could accomplish.

Sputnik and the Federal Embrace of Testing

It would take a geopolitical space crisis to fully cement standardized testing as a state-sanctioned educational practice for promoting talent. Over the course of several weeks in late 1957, the Soviet Union launched two artificial satellites into the earth's atmosphere. These two artificial moons, *Sputnik I* and *Sputnik II*, sparked widespread concern over how the United States had fallen behind in technological supremacy. Americans used the surprise of the satellites

to probe how their national education system had failed to provide their children adequate training. The federal legislative response to the sputniks—the National Defense Education Act of 1958—set a precedent for the type of intervention the federal government would make in educational agendas. In the process of codifying which testing programs it would support, the federal government also delineated its limited responsibility for the shape of American science education.

The sputniks drove American educationists, corporate leaders, and technical experts into a fit of introspection: How had American science fallen short? Pundits pointed to several systemic causes for American scientific shortcomings. In its comparison of American and Soviet education systems, General Electric blamed a widespread educational “doctrine of permissiveness” within a society that placed more emphasis on consumer spending patterns than educational expenditures.¹⁹⁸ Atomic scientist John Dunning argued in a *New York Times* essay that, although “the Government cannot legislate scientists into being,” it had to move beyond “the conference-and-cliché method of planning policies” if United States were to ever outpace the Soviet Union.¹⁹⁹ As the president of the New Jersey Education Association deduced, the Soviet satellites represented the American educational system’s long-term failure to build a corps of trained experts:

“It is not hard—just a little less comfortable—to say we could rather have two American sputniks in the air than two cars in every garage.’[...]’But two sputniks in the air are the result of 2,000 or 200,000 scientists in their laboratories. And scientists are made in elementary schools, as well as colleges.”²⁰⁰

¹⁹⁸ Charles Graham McClintock, *The Competition in Education: US vs. USSR (Tempo Report RM 58TMP-57)* (Santa Barbara, CA: General Electric Company, 1958), 51.

¹⁹⁹ John R. Dunning, “If We Are to Catch Up in Science,” *New York Times* November 10, 1957: 92, 29.

²⁰⁰ “U.S. Satellite Lag Tied to Education,” *New York Times* November 9, 1957: 2.

To determine how well American education stacked up against that of its main rival, the National Science Foundation (NSF) conducted its own study of the Soviet education system in the mid-1950s. Using available Soviet statistics and refugee reports, the NSF examined the full breadth and evolution of the Soviet schooling from the early Stalinist era to the immediate post-Stalin moment. What set American and Soviet schooling apart, the foundation argued, were fundamentally different ideological beliefs about the relationship between education, society, and the state. Whereas American education prioritized humanities-heavy “training for the realization of the full capabilities of an individual,” the Soviet system valued a centralized “service-to-the-state” model that created scientists and engineers who fulfilled the state’s research agenda.²⁰¹ Soviet scientific work, by extension, involved demonstrating and applying known scientific truths rather than challenging conventional wisdom or posing feasible alternative explanations for observed reality. The Soviet state instilled these core scientific and educational values through its polytechnic model for compulsory schooling, which mirrored Marxism’s opposition to the division of labor by training youth both mentally and physically. By valuing “productive work as an integral part of training,” polytechnic education offered the promise of immersive, integrative science instruction at every level of education.²⁰² Even when Soviet schools fell short of what polytechnic education promised—which happened often, the NSF insisted—they instilled a political imperative for scientific study from an early age.

Yet, the NSF warned against placing too much emphasis on the role of communist ideology in Soviet education. Focusing squarely on socialist indoctrination obscured the Soviet state’s primary aim to use education as a means for rapid industrialization. To assume Soviet

²⁰¹ De Witt, *Soviet Professional Manpower* (Washington, D.C.: National Science Foundation, 1955), 1.

²⁰² De Witt, *Soviet Professional Manpower*, 31.

youth wholeheartedly believed the ideological aspects of their schooling would ignore an obvious reason for pursuing a career in the sciences: social prestige. Scientific careers offered better guarantees of higher wages and social comfort, albeit at the cost of enduring additional propaganda. By centrally managing job placement and benefits, the Soviet Union had inadvertently created an incentive system for occupational prestige divorced from overt Marxist-Leninist rhetoric: the state will reward the brightest with the best jobs, so long as they remained loyal. Whether or all young adults sincerely believed the promises of the socialism, the Soviet state had created an education-occupation network that, above all, valued the production of scientists.

As the NSF determined, both the American and Soviet educational systems possessed particular strengths for producing scientists and engineers, creating a virtual dead heat in brainpower production. Although the Soviet Union had far fewer young adults in its higher education system—the revolution hadn't truly unmoored the peasantry from its rural agricultural base—proportionally far more Soviet students completed university studies in sciences and engineering than their American counterparts. While American students completed two more years of compulsory schooling than Soviet youth, they spent considerably less classroom time learning science and mathematics, a pattern even more pronounced when comparing each country's collegiate programs for scientific studies. As a whole, American youth were more educated while Soviet youth had more proficiency in specific fields. Vastly different approaches to public education left the United States and Soviet Union fairly evenly matched by the middle of the twentieth century, and American educational policy choices would determine “whether within the next decade or so the scales will be tipped off balance.”²⁰³ Hesitate and be left behind.

²⁰³ De Witt, *Soviet Professional Manpower*, 257.

Businessmen also staked in interest on where American brainpower fell short. During the mid-1950s, scientists, technologists, and psychologists from CalTech researched the potential drain humanity would place on resources over the next several generations. The team presented their findings to dozens of corporate executives and publishing their study in 1957. The published report, ominously titled *The Next Hundred Years*, made clear where the United States should focus its efforts for resource preservation: shoring up the massive erosion of potential scientists and engineers caused by premature dropouts. As the CalTech experts argued, American brainpower shortages didn't stem from a low proportion of science and engineering majors, but from steep drop-offs at every stage of education. For every 100 students who entered grade school, the researchers observed, only 59 graduated high school and just 20 entered college. The 13 young adults who completed a college degree—even the 3 who earned science and engineering degrees—did not reflect all the “technical brainpower” the United States could offer.²⁰⁴ “Educational attrition” took away between one-half and two-third of those capable of completing college, all while the Soviet Union produced scientists and engineers at twice the rate of the United States.²⁰⁵

Despite the vast room for improvement, the experts behind *The Next Hundred Years* warned of the political limits to brainpower development. The nation could not follow the Soviet Union's path to scientific prestige. Rigorous compulsory science education and state-managed job markets would boost the United States' pool of trained brains, but would ultimately prevent “a world where people can lead lives of abundance and creativity within the framework of a free

²⁰⁴ Harrison Brown, James Bonner, and John Weir, *The Next Hundred Years* (London: The Scientific Book Club, 1957), 137; “Brain Power Hope of Long Future,” *New York Times* May 27, 1956: 173.

²⁰⁵ Brown, Bonner, and Weir, *The Next Hundred Years*, 127.

society.”²⁰⁶ The authors insisted the United States could undo its impending science deficit simply by plugging in gaps where capable youth left school. By their calculations, American schools and businesses could jointly expand national technical brainpower by a factor of eight before approaching any natural population limits. By identifying how and when students become discouraged from advanced studies, private enterprise could devise ways to encourage structurally disadvantaged youth into participating in the American marketplace of ideas.

To make the case that the United States possessed far more potential technical brainpower than it actually employed, the authors of *The Next Hundred Years* defined intelligence and ability through standardized testing models. Researchers used Army General Classification Test score distributions to illustrate the rate of educational attrition between high school and college. Asserting AGCT scores were roughly equivalent to individuals’ IQ scores, the authors drew several conclusions about untapped American brainpower. Even though AGCT score for college graduates was 120, far less than half of the population with that level of aptitude completed higher education. An even smaller fraction of Americans with the average aptitude found in science and engineering graduates pursued such studies. Such a steep drop-off wasn’t a natural phenomenon, the authors insisted, but the inevitable result of a secondary school system that employed unqualified math and science teachers while neglecting to guide talented youth toward challenging coursework. Addressing the dearth of women scientists and engineers—a gender gap far more pronounced in the United States than in the Soviet Union—could effectively double American technical brainpower. By leaning on standardized test score models, researchers could identify social causes for squandered American brainpower, and in

²⁰⁶ Brown, Bonner, and Weir, *The Next Hundred Years*, 145.

doing so, emphasized the United States was not destined for second place so long as it genuinely managed its human resources.²⁰⁷

The Eisenhower Administration, meanwhile, anticipating the following year's midterm election, leveraged the shock of the sputniks against the ongoing desegregation crisis in Little Rock, Arkansas. Despite initially aiming to keep the integration of Central High School outside of partisan politics—indeed, being sour to the idea of educational desegregation in itself—Eisenhower hoped to use Little Rock as a symbol for the Democratic Party's inability to understand the United States' role in the postwar world. Being caught unawares by the sputniks may have been bad, but Democrats' willingness to flaunt judicial rulings only helped create scenarios that played to Communists' strength. A party that pledged to cut taxes while resisting integration, the administration warned, failed to understand the depths of America's scientific shortcomings as well as the impact images of white resistance would have in newly independent unaligned nonwhite nations. As Vice President Richard Nixon warned, the sputniks were not “a scientific stunt” but a sign that the federal government, under continued Republican leadership, needed to aid the development of American scientific brainpower.²⁰⁸

Proclaiming “a severe blow—some would say a disastrous blow—[had] been struck at America's self-confidence and at her prestige in the world,” The Congressional Committee on Labor and Public Welfare held a series of hearings during the first three months of 1958 to

²⁰⁷ Brown, Bonner, and Weir, *The Next Hundred Years*, 123-136. One recommendation worth noting that the authors didn't make was to, quite simply, pay teachers more. The researchers emphasized that those who chose a life in education over a career in industry must truly have a deep passion for teaching—without giving attention to the idea that incompetence in science and math education continued to fester because there was no economic incentive for talented Americans to pass up private enterprise for public service. To this end, the report reinforced a division of “hearts and minds” in employment.

²⁰⁸ W.H. Lawrence, “G.O.P. to Cite Little Rock Minimize ‘Moon’ in ‘58 Bid,” *New York Times* October 16, 1957: 24; Dana Adams, “U.S. Foreign Policy: Outlook for the Future,” *New York Times* October 13, 1957: 183; Thomas J. Hamilton, “World Seen Reaching a Balance of Terror,” *New York Times* October 13, 1957: 183; William deJong-Lambert, “Rethinking Little Rock: The Cold War Politics of School Integration in the United States,” *European Education* 38, no. 4 (Winter 2006-7): 65-81.

consider how the federal government should revise its approach to scientific education support.²⁰⁹ Throughout the hearings, senators grappled with the boundaries of the federal government's responsibility for education. Whereas South Carolina Senator and virulent segregationist Strom Thurmond routinely used his time to insist the federal government had no authority in the state of American education—the word “education,” after all, appeared nowhere in the Constitution—others on the committee were more interested in finding a line for federal involvement without breaching traditional local authority. The solution, many witnesses testified, would include merit-based scholarships for promoting science education and funds for helping states develop widespread standardized testing and guidance programs. As Health, Education, and Welfare Secretary Marion Folsom argued, providing matching funds for state testing programs would allow communities to better guide youth through school in ways that served local needs and upheld local values. To Folsom, the idea that federal assistance for identifying and supporting talented youth would lead to centrally dictated education was “a groundless fear of an imaginary danger.”²¹⁰ The secretary also assured the committee that a national scholarship program would be a better use of federal funds than an extensive educational loan program; merit-based awards would “put a little more prestige in the intellectual level of achievement” and stimulate college attendance among talented youth more fruitfully than any model for economic redistribution.²¹¹ Through these hearings, Congress aimed to settle on an approach that would

²⁰⁹ Lister Hill, *Hearings before United States Senate Committee on Labor and Public Welfare meeting on Science and Education for National Defense*, 85th Congress, 2nd Session, January 21, 1958 (Washington, D.C.: US Government Printing Office, 1958), 2.

²¹⁰ Marion Folsom, *Hearings before United States Senate Committee on Labor and Public Welfare meeting on Science and Education for National Defense*, 85th Congress, 2nd Session, February 6, 1958 (Washington, D.C.: US Government Printing Office, 1958), 186.

²¹¹ Folsom, *Hearings*, 205.

change the shape of American science education without triggering suspicions of federal intervention or social control.

The resulting National Defense Education Act (NDEA) cemented this middling federal approach, ensuring that the government would not interfere in the content of instruction while also setting an agenda and incentives that would inevitably alter the shape of American education. Within this environment—neither radically transformed nor structurally untouched—standardized testing thrived. Title V of NDEA not only provided explicit provisions for testing, but also promoted a national educational environment that would give greater weight to standardized testing. Title V allocated state educational agencies sixty million dollars over a four-year period “to assist them to establish and maintain programs of testing and guidance and counseling.”²¹² States were eligible to receive funds if they developed testing programs for public school secondary school students that would “identify students with outstanding aptitudes and abilities,” the results from which would then be used by guidance programs to steer exceptional youth into demanding college-preparatory high school coursework.²¹³ The title also provided an additional 28 million dollars for colleges to design guidance and counselor training programs, so that those working in secondary schools could better steer their charges toward nationally meaningful life paths. When combined with NDEA’s other provisions—hundreds of millions for student loans, grants for states for science equipment, and a new National Defense Fellowship graduate study program, among others—Title V cemented a general approach to scouting talent, in which standardized testing programs would help satisfy national needs without breaching traditional educational authority or embedded discriminatory behavior.

²¹² National Defense Education Act of 1958 (Public Law 85-864), 2 September, 1958: V, Sec. 501. For a clear summary of the entire act, see: Barbara Barksdale Clowse, *Brainpower for the Cold War: The Sputnik Crisis and the National Defense Education Act of 1958* (Westport, CT: Greenwood Press, 1981), 162-167.

²¹³ National Defense Education Act of 1958: V, Sec. 503 (a) (1)

Sensing the guidance provisions allowed by Title V would confuse laypeople, the Department of Health, Education, and Welfare (HEW) eventually published a primer for parents, teachers, and administrators explaining the basics of educational testing. The Department stressed that any testing program was “not an end in itself, but rather a valuable tool” guidance counselors and educators could use to adjust youth toward ideal paths of study and lines of work.²¹⁴ The booklet advised those without technical expertise to adopt a moderate opinion of testing: neither the cure for all society’s ills nor a pseudoscientific racket. Instead, the department assured readers that, when designed correctly, standardized objective tests were devices for making reasonable judgments about a person’s behavior or ability to perform. Good standardized tests went through an extensive construction process—with designers drafting several times as many test questions as would appear on the final version, and publishers arranging numerous experimental administrations in order to actually standardize a test. A well-constructed standardized test ideally targeted a very specific, clearly delineated skill set. As the department suggested, trained guidance professional would never use results from one standardized test to make judgments about a student’s other capabilities, nor use a single standardized test score to project a student’s entire educational and employment future. Rather, the department stressed, standardized tests should be seen as an integral part of a student’s cumulative record, which would ideally “begin with information about the pupil before he enters school and continue until he has been out of high school at least 5 years.”²¹⁵ As school districts began to adjust to Title V’s provisions, HEW aimed to convince parents and teachers that the best approach to standardized testing was to use many, many more tests.

²¹⁴ Kenneth F. McLaughlin, ed., *Understanding Testing: Purposes and Interpretations for Pupil Development* (Washington, DC: U.S. Government Printing Office, 1960), 1.

²¹⁵ McLaughlin, *Understanding Testing*, 14.

While the federal government promoted the benefit of comprehensive standardized testing programs to parents and teachers, the College Board sought to deepen its relationship with colleges and students. Ironically, the organization did so by first encouraging college officials to deemphasize the role standardized test scores played in admissions decisions. CEEB's aim was not to lessen the importance of its exams in college admissions, but to heighten their value by stressing their technical limitations. The SAT and College Board achievement exams scores were only truly useful admissions tools when used alongside high school grades, and when treated more like a range than an exact measurement of educational aptitude.²¹⁶ CEEB also decided to grant colleges permission to disclose test scores to students for the first time. The move, effective for all scores after December 1958, seemed to be a way to placate students' curiosity while pressuring schools not to use secret test scores as a convenient way to maintain racially segregated campuses.²¹⁷ By 1959, the organization developed a new standardized test specifically for high school juniors: the Preliminary Scholastic Aptitude Test (PSAT). The College Board maintained the PSAT would "give the junior a forecast of his senior-year performance on the [SAT]," and encourage him to make better-reasoned college plans sooner.²¹⁸ CEEB stressed, however, that the new test would not make high school juniors familiar enough with the format of aptitude tests to game the system, as only senior-level courses provided the mathematical and verbal material needed to noticeably raise scores. The organization instead suggested that college hopefuls may as well familiarize themselves with the SAT with a preliminary version of the exam because, by its own projections, all degree-granting colleges

²¹⁶ "College Boards' Value Appraised by Expert," *New York Times* September 7, 1958: E11.

²¹⁷ "Colleges to Bare Entrance Scores," *New York Times*, October 31, 1957: 33.

²¹⁸ "New College Boards for Juniors," *New York Times* March 8, 1959: E9.

would require some sort of entrance examination before the end of the next decade.²¹⁹ For the College Board, standardized admissions tests were becoming nearly unavoidable—and if higher education was increasingly pegged to the idea of successful citizenship, then preliminary and supplementary standardized tests could only help students adjust their ambitions accordingly.

Well before the Eisenhower administration responded to Sputnik, the federal government, extra-governmental educational organizations, and corporations began building a network dedicated to catching the United States' brightest youth and steering them toward scientific and engineering careers that would secure American geopolitical supremacy. Sputnik did not spark a commitment to science and mathematics so much as solidify the direction multiple stakeholders, both private and public, took in harnessing brainpower for geopolitical ends. Scholarship programs allowed corporations to use standardized tests to cull talent while promoting the merits of American capitalism. Standardized testing also allowed the federal government to dodge accusations of centralization at a moment when some alleged it had superseded traditional local control for education. Rather than restructure U.S. education, standardized testing became the netting that bound all vested parties together. Where the rhetoric of American democratic brainpower fell short, standardized testing provided a convenient solution.

Although standardized test designers noted the limited statistical usefulness of their products, they also developed a rhetoric that stressed standardized testing could create tremendous social good. The more popular standardized testing became, the more opportunities arose for students, parents, educators, administrators and policymakers to draw from this psychometric optimism and make outsized conclusions about what standardized tests meant and measured. These tensions between expert, layperson, and corporate understandings of what

²¹⁹ Leonard Buder, "Testing Increases for College Seen," *New York Times* October 30, 1958: 64.

standardized tests could do created an educational culture in which standardized tests easily became associated with bigger political and cultural forces. Test scores became shorthand for an individual's worth to the state and the quality of American brainpower: a symbol for the fate of the America's future. What became difficult to accept for many, then and now, was the basis by which educational aptitude was defined. If standardized tests measured natural talent, then the United States had failed to steer the majority of its brightest youth into college. But if standardized tests measured the skills that came from years of schooling, then test scores provided clear proof of how deeply racial and class inequities shaped social educational disparities. Despite experts' specific understanding of test validity and reliability, aptitude remained a slippery concept—and one the public would continue to invest with multiple, often-contradictory meanings.

The use of standardized testing as a means for harnessing brainpower persists—but, more frequently, international standardized tests have served as way for Americans to fret over whether their youth will maintain geopolitical preeminence through scientific innovation. The Organization for Economic Co-operation and Development (OECD) issues a multi-subject standardized test to hundreds of thousands of students in several dozen countries every three years. Along with other capabilities, the Program for International Student Assessment (PISA) gauges the complex mathematic and scientific problem-solving skills of teenagers in participating member countries. Although the United States spends disproportionately more money on students than most other PISA participant countries, American students continually lag behind their Asian and European counterparts on the math and science portions of the standardized test. While the geopolitical dimension of PISA scores remains less tightly bound to any particular state—Estonia, Singapore, or Finland seem far less threatening than the Soviet

Union did—the fear remains that American brainpower is withering on the vine.²²⁰ So, too, remains Americans’ faith that standardized tests can help them make sense of their place in the world.

²²⁰ Organization for Economic Cooperation and Development, “About PISA,” OECD website, <http://www.oecd.org/pisa/aboutpisa/>; Emily Richmond, “How Do American Students Compare to Their International Peers?” *The Atlantic*, December 7, 2016, <https://www.theatlantic.com/education/archive/2016/12/how-do-american-students-compare-to-their-international-peers/509834/>; Jill Barshay, “U.S. Now Ranks Near Bottom Among 35 Industrialized Nations in Math,” *Hechinger Report*, December 6, 2016, <http://hechingerreport.org/u-s-now-ranks-near-bottom-among-35-industrialized-nations-math/>; and Amanda Ripley, “What America Can Learn from Smart Schools in Other Countries,” *New York Times*, December 6, 2016, https://www.nytimes.com/2016/12/06/upshot/what-america-can-learn-about-smart-schools-in-other-countries.html?_r=0, (all last accessed June 1, 2017).

CHAPTER THREE: MAPPING MARGINAL MEN: STANDARDIZED TESTING AS A FORM OF RACIAL SOCIAL CONTROL IN THE POSTWAR STATE

Race is a defining feature of American society, but its persistence as a categorical marker for social disparity is too often attributed primarily to the misdeeds of individual actors. Race, by this circular logic, depends upon the racism harbored by racists. The burden of proof for the presence of racism thus becomes whether or not a person openly admits that racial ideology fueled their actions, thereby turning racialized motivation into incomprehensible malice. When unmoored from their socioeconomic bases, race and racism thus seem analytically impenetrable: ineffable mysteries only combatable through the politics of affect. Yet what fuels the usefulness of race as a hierarchal system for social sorting—and, in turn, gives currency to racial ideology and racist actions—are the institutions and organizations that give nation-states a sense of continuity and legitimacy. State-sanctioned structures carry out projects that entrench and justify racial difference in everyday life, and give anchorage to the various ways groups and individuals make their own versions of racial meaning. At each level, racial ideology shapes a reality in which the fact of race is self-evident. The tension that keeps race afloat is thus not between fact and fantasy—ideology creates its own logic—but between the conservative impulse of the nation-state and the profound flexibility and variety in individuals' racial interpretations.²²¹

²²¹ See, for a hodgepodge of arguments: Martin Gilens, "'Race Coding' and White Opposition to Welfare," *American Political Science Review* 90, no. 3 (September, 1996):593-604; Arthur R. Williams and Karl F. Johnson, "Race, Social Welfare, and the Decline of Postwar Liberalism: A New or Old Key?" *Public Administration Review* 60, no. 6 (November-December, 2000): 560-572; Howard Winant, "Race and Race Theory," *Annual Review of Sociology* 26 (2000): 169-185; Daniel Kryder, *Divided Arsenal: Race and the American State During World War II*

The salience of race persists even when the state has backed away from overtly racial policies. This persistence occurs because the state depends upon technologies—such as standardized testing—that are built upon and contribute to racialized frameworks for understanding ability and social utility. Postwar institutional desegregation was one such moment in which the state abandoned racist policies while maintaining the racialized technology of standardized testing. Over a several-year period following the Second World War, the United States Armed Forces integrated its troops, marking one crucial early moment in the dismantling of legal racial segregation within American institutions. Certain branches—namely the Army—underwent more extensive overhauls of their race-based personnel management policies. In addition to gradually integrating its troops, the Army also abandoned restrictions to the amount of black soldiers within its ranks at any given time. After a year of consistent pressure from President Harry Truman’s Committee on Equality of Treatment and Opportunity in the Armed Services, Army leaders agreed to abandon this racial quota. In its place, the Army adopted a quota based on the Armed Forces’ standardized aptitude entrance exam: the General Classification Test (GCT). By replacing the racial quota with a GCT quota, policymakers used standardized testing to forge a postwar social order that, on the surface, prioritized ability over race—but in effect used standardized testing as a means to avoid disrupting too much of the existing white supremacist social order. By allowing some viable channel for deeper black

(Cambridge: Cambridge University Press, 2000); Mary Poole *the Segregated Origins of Social Security: African Americans and the Welfare State* (Chapel Hill: UNC Press, 2008); Ann Morning, “‘Everyone Knows It’s a Social Construct’: Contemporary Science and the Nature of Race,” *Sociological Focus* 40, no. 4 (November, 2007): 436-454; George Wilson and Amie L. Nielson, “‘Color Coding’ and Support for Social Policy spending: Assessing the Parameters among Whites,” *annals of the American Academy of Political and Social Science* 634 (March, 2011): 174-189; Steve Martinot, *The Machinery of Whiteness: Studies in the Structure of Racialization* (Philadelphia: Temple University Press, 2010); Adolph Reed, Jr., “Marx, Race, and Neoliberalism,” *New Labor Forum* 22, no. 1 (Winter, 2013): 48-57; Karen E. Fields and Barbara J. Fields, *Racecraft: The Soul of Inequality in American Life* (London: Verso, 2014); and Michael Omi and Howard Winant, *Racial Formation in the United States*, 3rd edition (New York: Routledge, 2015).

participation in the armed forces, however limited, the Truman Administration could satisfy its aim to integrate and modify the military without creating too much discontent within military leadership at a vital moment in American statecraft. These actions would set a precedent for the postwar state, in which standardized testing would continue to use as a seemingly race-neutral device that nonetheless reinforced the perception that black men were marginally useful to both state and society.

Racial Ideology, the GCT, and Desegregation

The ways in which the American state mobilized for World War Two and sorted men for military service depended upon psychologists' ability to maximize manpower within prevailing racial ideology. Although the Selective Service Act of 1940 established a superficially nondiscriminatory draft policy—establishing quotas only for geographic divisions—the Army carried out white-only drafts in many regions, and continued to segregate its forces even after public pressure from black journalists and political leaders compelled the Roosevelt Administration to make several administrative concessions. Military segregation endured even after Executive Order 8802 prohibited defense industry workplace discrimination, an order made effective when black women pressured federal agencies and labor unions to ensure manufacturing corporations did not weasel out of nondiscriminatory hiring practices. Throughout the war, the American state used the language of manpower to thread the logic in which nondiscrimination policies and segregationist practices could occur simultaneously.²²²

²²² Phillip McGuire, "Desegregation of the Armed Forces: Black Leadership, Protest, and World War II," *Journal of Negro History* 68, no. 2 (Spring, 1983): 147-158; George White, Jr. "'I Am Teaching Some of the Boys': Chaplain Robert Boston Dokes and Army Testing of Black Soldiers in World War II," *Journal of Negro Education* 81, no. 3 (Summer, 2012): 200-217; Megan Taylor Shockley, "Working for Democracy: Working-Class African-American Women, Citizenship, and Civil Rights in Detroit, 1940-1954," *Michigan Historical Review* 29, no. 2 (Fall, 2003): 125-157. This isn't to say that federal policy overrides regional differences; see: Beth Bailey and David Farber, "The 'Double-V' Campaign in World War II Hawaii: African Americans, Racial Ideology, and Federal Power," *Journal of Social History* 26, no. 4 (Summer, 1993): 817-843

During the war, psychologists and Army leaders amassed a tremendous amount of data about men's military usefulness by subjecting millions of incoming recruits to the Army General Classification Test (GCT). For the Army, GCT data became a useful tool for valorizing American democratic principles in contrast to Axis racial ideology. In an omnibus report on black soldiers, the Army Service Forces' Military Training Division asserted that assumptions of black Americans' mental inferiority premised on the Army Alpha and Beta test data were not only scientifically unsound but also reminiscent of the very fascistic rhetoric the United States was currently combatting. In the ensuing quarter-century, the Army asserted, testing devices had improved, and the GCT provided "a quick and reasonable dependable measurement of an enlisted man's working level and *ability to learn*," no more and no less.²²³ By posing the differences in black and white aptitude test scores as a matter of job readiness—without ever suggesting which parties, exactly, were responsible for maintaining racially disparate educational funds and job opportunities—the Army could congratulate itself on not falling prey to the race science of a previous generation, all while avoiding casting blame on any politicians who kept white supremacist structures intact. The Army, as it portrayed itself, was an institution that granted equal opportunity of service: it was not the Army's problem to fix the lack of opportunity elsewhere in American society.

Some of the chief psychologists behind the GCT held less overtly nationalist analyses of the exam—but they nonetheless established a connection between a man's standardized test scores and his worth to society and the state. After the end of conflict, War Department Chief Psychologist Walter Bingham published general findings about the GCT's reliability in predicting a recruit's eventual military successes in *Science* magazine. Bingham insisted the test

²²³ Army Service Forces Military Training Division, draft of *The Negro Soldier*, p. 40; "Instructional Aids for Use in Connection with ASF Manual, M 5, 'Leadership and the Negro Soldier'; Philleo Nash Papers, box 59.

data illustrated a clear general pattern: the lower a man's GCT score, the less likely he would successfully complete a job training course with at least average marks, and the greater likelihood he would "prov[e] to be more of a burden than a help during mobilization."²²⁴ Yet, the sheer number of men who took the GCT—about 10 million—allowed Bingham to make broad generalization about a massive population while also noting sizeable exceptions. While about 5,000 men who lacked grammar school educations still received high enough GCT scores to earn the highest possible Grade classification, this sizeable number of men was still one half of one tenth of one percent of all men examined. This unprecedented amount of score data thus allowed the armed forces—and the American state—to be convenient with how it discussed population statistics, toggling between percentages and numbers depending on which sense of type of figure would better serve state needs.

During World War II, American military leaders assembled to determine how to make the best use of black personnel, setting the foundation for using standardized testing data to uphold white supremacist power dynamics. One military committee, led by Assistant Secretary of War John McCloy, considered how to juggle the dire need for able-bodied men, regardless of race, with the seeming intellectual shortcomings many black recruits displayed in their low GCT scores. Nearly half of all black recruits scored in the lowest GCT score category (Grade V), compared to nine percent of their white counterparts. In an aborted scheme, the McCloy Committee planned to replace low-scoring black personnel with higher-scoring men, relocating Grade V men to labor troops. The committee reasoned this reshuffling would create little friction among low-scoring black soldiers, as "the majority of Negroes falling into Grade V [were] from the South where they have been accustomed to performing common labor under immediate

²²⁴ Walter Bingham, "Inequalities in Adult Capacity—From Military Data," *Science* 104, no. 2694 (August 16, 1946): 150.

white supervision.”²²⁵ Disagreements between the Army’s Ground and Service Forces eventually undid this plan, yet the McCloy Committee’s model reflected a willingness to trade upon existing racial power dynamics—black men in menial tasks overseen by powerful white men—in order to ensure the military ran as efficiently as possible.²²⁶

Although the McCloy Committee’s redistribution plans failed to take shape during the war, some members spent time just after the end of hostilities considering where black men would fit into the postwar defense system. In a late August 1945 memo to his superior, McCloy’s assistant Davidson Sommers recommended the “eventual non-segregation and assignment of Negro troops solely on the basis of ability.”²²⁷ Yet Sommers referenced GCT scores to claim that, overall, black soldiers were poorly equipped for Army demands when compared to white soldiers. This claim, Sommers reassured McCloy, did not need to be premised on “racial theories,” as black soldiers’ relative “lack of educational, occupational, and social opportunities [was] a sufficient explanation.”²²⁸ By asserting that environmental factors created black recruits’ low aptitude test scores, Sommers hinted at the Army’s institutional resistance to abolish its racial policies. If the Army provided one of the few consistent social opportunities for black Americans within a broader racial system designed to subordinate them, then the abolition of racial controls would provide an even more compelling case for black men to serve in the armed forces. The Army would thus make itself responsible for providing what other American

²²⁵ I.H.E., memorandum to the Assistant Secretary of War, re: “Utilization of Grade V Negro Personnel,” February 17, 1943, p. 1; RG 220, Records of the President’s Committee on Equality of Treatment and Opportunity in the Armed Services, Box 14; Truman Library.

²²⁶ E.W. Kenworthy, letter to Charles Fahy, 8 August 1949, RG220, Box 14; Adjutant General’s Office, memo to Assistant Secretary of War John J McCloy, memorandum re: “Advisory Committee on Negro Troop Policies,” August 27, 1942, RG 220, Box 14.

²²⁷ Davidson Sommers, memo to Assistant Secretary of War John J McCloy, August 28, 1945, p. 1, RG 220, Box 14.

²²⁸ Sommers, memo to McCloy, p. 2.

institutions and systems refused to do, at the risk of diluting its overall talent pool. Yet in his proposed solution—using GCT scores as a barrier for entry—Sommers also acknowledged that the opposition would be both internal and external. Although most parties agreed that black men would continue to have a role in securing the postwar American state, any changes that threatened the institutional characteristics of the Army risked upsetting at least some of its constituent groups.

Several months after the dissolution of the McCloy Committee, the War Department assembled a new panel to revise its use of black troops in the uncertain postwar period. The panel, led by Army Chairman Lieutenant General Alvan Gillem, released their recommendations as a circular memorandum in April 1946. Within Circular 124, the Gillem Board established the unenviable position the Army believed itself to be in. Postwar defense required maximum possible manpower, but “the manpower available, of itself, [varied] in quality.”²²⁹ Black soldiers were far more qualified during World War Two than their counterparts in World War One, and the Army prided itself on offering additional training and education opportunities otherwise unavailable for black troops outside of the military, but considerable gaps between the quality of manpower black and white troops brought upon entering service endured. The Army felt an institutional responsibility to continue serving as a conduit for opportunity—but it also aimed to prevent wasting resources on improving the manpower of too many black men. Hence, the Gillem Board announced that, in order to attain “maximum utilization of the authorized manpower of the nation in the event of a national emergency,” the Army would accept black troops at a ten percent quota, approximate to the proportion of black Americans in the general

²²⁹ War Department, “Circular No. 124: Utilization of Negro Manpower in the Postwar Army Policy,” April 27, 1946; p. 3.

population.²³⁰ Black soldiers, however, would remain in segregated units, and black units with a disproportionate amount of men with low GCT scores would receive additional supervisors. The Army positioned itself as an institution for black advancement—but as an institution, it also used numerical and testing quotas to set a firm boundary for how much change it would accept.

The Army sought to normalize Circular 124 among troops while also limiting its own responsibility for sustained institutional segregation. The War Department publication *Army Talk* devoted a three-part essay to the new policies mandated by the circular, designed to be read aloud and discussed in organized group settings. Deeming the issue of black manpower a “special problem,” *Army Talk* detailed “the Army’s concerns with matters of race as they relate to its effectiveness and its status in a democracy.”²³¹ As the newsletter explained in digestible segments, the Army did not hold any policy of racial hierarchy, but simply maintained segregated troops as a matter of organizational tradition. The essay’s writers insisted that the Gillem Board did find certain racialized military structures detrimental to national security. The postwar peacetime army had to be treated as “our war machine in miniature”—and overly large black units overseen by white officers would wreck efficient American combat.²³² While the Army held a position of equal opportunity, it declared itself “not an agency of social reform in matters either of race, racial relationship, or anything else.”²³³ As the Army hoped to instill among its personnel, the military was not the party responsible for society’s larger shortcomings—nor was it responsible for addressing those issues at the risk of efficiency.

²³⁰ Circular No. 124, p. 1.

²³¹ United States War Department, *Army Talk* 170 (April 12, 1947), 2; Philleo Nash Papers, Truman Library, Box 59.

²³² *Army Talk* 170, p. 4.

²³³ *Army Talk* 170, p. 7.

As the Army made the policies within Circular 124 an accepted institutional norm, it also complained to the Defense Department that it endured unfair testing and racial standards compared to the other branches of the military. After hearing various arguments and responses from the Army, Navy, and Air Force, Secretary of Defense James Forrestal issued findings on the how feasible it would be for the Armed Forces to maintain uniform GCT caps. The Army alleged that the 1948 revision of the Selective Service Act forced it to accept an unfair amount of marginal manpower compared to other branches; while the Navy and Air Force could refuse men with GCT scores below 87 and 90, respectively, the Army was bound by law to accept men with at least a score of 70. The Army also demanded that other military branches adopt a uniform quota for black personnel, insinuating that the Navy's resistance to racial quotas conveniently obscured the fact that it employed only 4 black officers within its ranks, compared to 1000 within the Army. Forrestal agreed that the Army bore a disproportionate burden employing marginally useful men, but concluded that the branch "has many vital, but 'dirty,' jobs in wartime that compare unfavorably in creature comforts, danger and interest to jobs in the Navy and Air Force": the Army would invariably attract the least acceptable men because it needed more of them to do the most unsavory work.²³⁴ Further, Forrestal acknowledged that creating a uniform GCT standard would help the Army's manpower, "but at a cost to the Navy of 9%, to the Air Force of 10%, and to the National Military Establishment as a whole of 3%."²³⁵ Because

²³⁴ James Forrestal, "Findings and Decision on Questions of Parity and Mental Standards, Allocation of Inductees to Physical and Mental Capabilities and Allocation of Negroes," October 29, 1948, p. 13; Charles Fahy papers, Box 4. According to notes compiled by Charles Fahy based on correspondence with the Department of Defense, the Army had 257 black units as of November 1, 1948. Within those units were 89 transport units, 25 laundry units, 15 Army Band units, 13 ambulance and evacuation units, 10 infantry battalions, 9 dump truck units, 8 field artillery battalions, and 8 fire-fighting companies. Within black troop ranks were just over 15,700 noncommissioned officers (soldiers who had earned master sergeant, sergeant first class, sergeant, or corporal status) and about 1,100 officers (in this case, men who had earned a rank of master officer, 2nd lieutenant, 1st lieutenant, captain, major, lieutenant colonel, or colonel); n.a., "The following Army units are composed of Negro personnel as of 1 November, 1948," n.d., Charles Fahy Papers, Box 3.

²³⁵ Forrestal, "Findings," p. 18.

black soldiers remained disproportionately clustered in lower GCT score brackets, the peacetime burden for black personnel had to be shouldered largely by the Army. As Forrestal remarked, the existing peacetime military, unprecedentedly massive as it was, still had to represent “the nucleus of a future wartime force,” and the manpower burdens faced by one branch could not be altered to the detriment of the entire armed forces’ manpower output.²³⁶

But broader geopolitical pressures compelled the federal government to push further on nondiscriminatory military policy. For the Truman Administration, systemic racial discrimination and segregation had become Cold War rhetorical weapons—a source of international criticism and scorn by enemies and allies alike.²³⁷ To combat one aspect of this geopolitical quandary, President Truman issued Executive Order 9981 in July 1948. The order mandated “equality of treatment and opportunity for all persons in the armed services without regard to race, color, religion or national origin”—with all necessary changes to take place “as rapidly as possible.”²³⁸ To help all branches of the military reach this goal, the executive order also established the President’s Committee on Equality of Treatment and Opportunity in the Armed Services, a seven-member advisory group authorized with determining and recommending desegregation policy. Truman designated former Solicitor General Charles Fahy to chair the new committee. The president filled out the group with members that straddled educational, corporate and civil rights interests, appointing to the committee the presidents of Oberlin College, General Cable Corporation, and A.J. Donahue Corporation, as well as the

²³⁶ Forrestal, “Findings,” 38.

²³⁷ Mary Dudziak, *Cold War Civil Rights: Race and the Image of American Democracy* (Princeton: Princeton University Press, 2011); Christopher Waldrep, “National Policing, Lynching, and Constitutional Change,” *Journal of Southern History* 74, no. 3 (August, 2008): 589-626.

²³⁸ Harry S Truman, “Executive Order 9981: Establishing the President’s Committee on Equality of Treatment and Opportunity in the Armed Forces,” July 26, 1948; Transcript, “Meeting of the President and the Four Service Secretaries with the Presidential Committee on Equality of Treatment and Opportunity in the Armed Services,” January 12, 1949; RG220, Box 9.

executive secretary of Urban League and the publisher of the *Chicago Defender*.²³⁹ Those on the Fahy Committee recognized the geopolitical esteem the United States risked by maintaining segregated forces. In a letter to Cleo Blackburn, superintendent of Indianapolis's African-American social service center Flanner House, James Evans lamented "that the world denies us recognition which would otherwise justly be ours" were the military desegregated—adding that most Americans remained unaware of "the devastating effects of this deplorable situation on the prestige of America abroad."²⁴⁰ President Truman established his aim for concrete, thorough, expansive policy at a brief introductory meeting with the new committee, Secretary of State James Forrestal, and Defense Department leaders the following January. Truman envisioned that the committee would go beyond examining existing military barriers to carrying out his executive order, but also related federal bureaucratic procedures that could also gum up military desegregation. Through this committee, the Truman Administration could partially address a prevailing domestic concern that had become an international spectacle.

Some in the Fahy Committee quickly realized that the only effective line of reasoning to convince the armed forces to desegregate would be an argument about manpower. In early March 1949, the committee's executive secretary, E.W. Kenworthy, warned Fahy about the potential pitfalls in appealing to broader principles. Whether framing military desegregation from the perspective of constitutional authority or as a matter of practical politics, the president's capacity to act as Commander in Chief was checked by Congress's responsibilities to oversee military policy and appropriations: an executive order and committee report alone would carry a weak mandate. Kenworthy also dismissed using the moral high ground embedded in

²³⁹ Eben A. Ayers, Assistant Press Secretary, press release, September 18, 1948; RG220 Records of [President's Committee], Box 9: Truman Library.

²⁴⁰ James C. Evans, letter to Cleo Blackburn, February 9, 1949; RG 220 Record of [President's Committee], box 4.

foundational American legal texts because, from the military's perspective, the lofty ideals of the Bill of Rights had nothing to do with effectively winning wars. Where the committee could instead force the Army's hand is by confronting them on the inefficient use of black manpower. By deferring to racial models of intelligence and failing to create a cadre of black leadership during peacetime, Kenworthy reasoned, the Army rushed into combat with large, unsupported black units. This created low morale and panic among black troops—and ultimately unfavorable service records when compared to the conduct of black soldiers allowed to serve within white regiments on the Western Front. By appealing to efficiency, the Committee could “find a side entrance” to changing Army policy that relied on the military's own alleged priorities.²⁴¹

In its initial report, issued late March 1949, the President's Committee premised its recommendations on the belief that the armed forces and the postwar state could only succeed through “the maximum efficient use of manpower.”²⁴² If the increasingly technical qualities of war rendered manpower an abstract quality detachable from any particular man, then a soldier's race should have little bearing on their job detail: black or white, men were vessels for manpower and brainpower. Hence, the President's Committee recommended that the Army remove discrepancies between black- and white-unit organization that effectively barred black troops from certain military occupational specialties (MOS). The President's Committee also recommended that the Army no longer use race as a factor when assigning men to Army Schools courses, nor for assigning men to units once they've completed advanced training. Above all, the President's Committee recommended that the Army abolish its 10% quota for black troops. If the

²⁴¹ E.W. Kenworthy, memorandum for Charles Fahy, re: “Approach and Method,” March 10, 1949, p. 3; RG220 Records of [President's Committee], Box 1: Truman Library.

²⁴² President's Committee, “Initial Recommendations by the President's Committee on Equality of treatment and Opportunity in the Armed Services,” March 24, 1949, p. 1; RG220 Records of [President's Committee], Box 6: Truman Library.

Army wanted to maximize manpower, it needed to do so on the basis of aptitude and not racial composition.

The Army, meanwhile, adopted various tactics to make its ranks whiter. Between May 1948 and March 1949—a period without any immediate demands for troops—the number of Army officers and enlisted personnel grew by nearly 95,000, to roughly two-thirds of a million people. In that same ten-month window, however, the number of black officers and enlisted personnel grew by only 3,500 men. While technically still accepting more black men, the Army effectively thinned out the proportion of black men in its ranks to almost exactly ten percent. But this adherence to the racial quota did not extend to upper rank: black officers still comprised less than two percent of the officer corps.²⁴³ By April, the Army used the racial quota to halt black recruitment outright. Assessing that its proportion of black soldiers remained unacceptably high, the Army declined black volunteers, all while lowering its GCT minimum to increase the pool of potential white recruits. The Army also extended these race-based personnel policies to its protocol for reenlistment. Although all soldiers had twenty days to reenlist in their same position, the Army granted white soldiers an additional three-month window to reenlist at the same pay grade; black soldiers who missed the first window were barred from reentry “until the Negro quota [was] open.”²⁴⁴ Yet the Army did not gather any GCT scores for black men they turned away, let alone any sense of whether low-scoring white men were accepted into the army in lieu of higher-scoring black hopefuls. The racial quota had already predetermined these dismissed black men were numerically beyond what was the acceptable burden of the national security state.

²⁴³ “Department of the Army Strength Figures,” n.d., n.a.; RG 220 Record of [President’s Committee], box 6.

²⁴⁴ Joseph H. B. Evans, Memorandum for the record, re: “Conference with Major Bowers of MPPSD,” May 31, 1949,” RG 220, box 8.

The Army resisted abandoning its racial quota by instead demanding that all three branches of the military adopt a uniform cutoff for the General Classification Test. At the time, the Navy used a GCT cutoff score of 90, arguing that the technical demands of modern naval defense meant it could not employ the services of men with below-average scores—nor, by deduction, below-average intelligence. The Army countered by contending that, were it to abandon its racial quota without being allowed to adopt a higher GCT cutoff, a disproportionate amount of intellectually deficient men would flock to Army recruiting stations. The Fahy Committee found this line of reasoning unacceptable, citing the Air Force’s policy of using men with lower scores, and instead recommending low GCT cutoffs across all three branches in order to cull all possible useful manpower.²⁴⁵

To make sense of the Army’s insistence on racial quotas—and how standardized testing fit within the racial logic of its revised personnel management systems—the Fahy committee met with several military officials in late April 1949. As departmental and Selective Service officials explained over the course of several hours, the GCT had become just the first of several standardized exams that determined the line of work an incoming recruit would eventually be assigned. After clearing the Army’s peacetime GCT quota—with lower-scoring hopefuls needing to earn a higher physical rating to pass muster—recruits endured a battery of ten aptitude assessments as part of their induction process. Each test within the battery measured some combination of ten capabilities, including pattern analysis, shop mechanics, and clerical speed. Following this series of exams, recruits then took the Army Activities Preference Blank, another standardized inventory designed to objectively gauge the fields of work a person was most attracted to. Army intake officers compiled the results from a recruit’s Preference Blank as

²⁴⁵ President’s Committee, “Initial Recommendations,” 8-10.

well as the average skill scores from each aptitude test to determine which set of jobs he should be assigned. For the Army, the GCT worked as a manpower-harnessing device when it was used as the first in a series of standardized assessments.²⁴⁶

Despite this new system, the War Department officials interviewed by the Fahy committee continued to assert that race remained a meaningful category for sorting and limiting Army troops. When asked why the Army continued to segregate its GCT figures by race, department officials evaded a concrete answer, only noting that GCT scores figures remained categorized by race because all Army data was divided in that fashion. When members of the committee then suggested a soldier's level of education would be a more meaningful frame for analysis, one military representative, Roy Davenport, rattled off a series of statistical extremes to suggest race remained a more meaningful basis of categorization:

However, you might be interested to know that of those individuals who reported some schooling, for whom the records show some schooling but not completion of grade schools, three tenths of one per cent of the whites scored in Grade 1, 35.5 per cent in Grade 5. For the colored two-tenths of one per cent in Grade 1 and 73.1 per cent in Grade 5. [...] At the other end of the scale, those who had completed college, for the whites 47.5 scored in Grade 1 and no individuals were reported as scoring in Grade 5. For the colored, 6.4 per cent scored in Grade 1 and 1.5 per cent scored in Grade 5.²⁴⁷

The committee's executive secretary, E.W. Kenworthy, pressed on, asking Davenport whether non-segregated units would create fewer occupational pathways for black soldiers—that is, whether the Army felt segregation beneficially limited job competition. Even then, Davenport deferred to test score figures: over a third of all white men scored in the two highest score tiers

²⁴⁶ Transcript of meeting between President's Committee on Equality of Treatment and Opportunity in the Armed Services, Roy Davenport, Frank McKernan, and Major James D. Fowler; April 26 1949, Pentagon, Washington, DC, pp. 25-35; RG 220, Box 11.

²⁴⁷ Roy Davenport, in April 26 President's Committee meeting, transcript pp. 16-17.

on the GCT, while fewer than 4 percent of black soldiers did so.²⁴⁸ Through an avalanche of statistical data absent of any sense of scale or scope—percentages and comparisons without tangible numbers for anchorage—the army used GCT scores as a means to justify its racial quotas without having to provide racially motivated rationales for those restrictions: numbers, after all, don't lie.

The Secretary of the Army separately took great pains to stress that, among the four services in the Armed Forces, his branch made the greatest efforts to include black Americans. In a May 1949 memo to Secretary of Defense Louis Johnson, Acting Secretary of the Army Gordon Gray outlined the sheer numerical advantages black Americans had in the Army when compared to the Navy, Air Force, or Marine Corps. Gray stressed that, numerically speaking, black Army servicemen totaled “more than triple that of the Air Force or Navy and almost double that of all other services combined.”²⁴⁹ Further, black soldiers had far more pathways to attain a noncommissioned officer rank than their counterparts in other branches of the Armed Forces. In Gray's estimation, his branch was not the one to blame for the Armed Forces' underutilization of black Americans.

Gray also used General Classification Test score patterns to defend the relatively small number of black commissioned officers in the Army, but his faulty math exposed the trouble many non-specialists faced when attempting to make sense of testing patterns. Grey conceded that black commissioned officers were relatively rare—numbering about 1,300, or just 1.65

²⁸ Davenport, in April 26 President's Committee meeting, transcript p 18.

²⁴⁹ Gordon Gray, memo for Louis Johnson, “Equality of Treatment and Opportunity in the Armed Forces,” May 26, 1949, p. 3; RG 220 Record of [President's Committee], box 6. Assistant Director to the Selective Service Colonel Campbell C. Johnson, when meeting with the President's Committee in April, 1949, detailed the Army's development of black troops during the war, and its continued use of black troops relative to other military branches after the war. When Selective Service began in 1940, black men comprised fewer than 2 percent of the Army's ranks.

percent of all Army officers. The secretary argued, however, that this relative dearth of black commissioned officers was not a product of discriminatory policy but a mere byproduct of the General Classification Test. At the time, the Army required a minimum GCT score of 110 for any soldier to be considered for a commissioned officer position. For every black soldier who scored at least 110 on the GCT, 6.5 white soldiers met the same qualification. Grey then convoluted these score patterns to justify the Army's use of a racial quota: multiplying the existing percentage of black commissioned officers by the racial ratio of soldiers who scored at least 110 on the GCT would "equal competitively to a 10.7 per cent ratio, or slightly in excess of the total Negro strength ratio."²⁵⁰ Whether in good or bad faith, such gibberish mathematics allowed the Army to use the aura of statistics to stand by its racialized personnel policies.

The Army's use of specious calculations for justifying its low levels of black commissioned officers reflect officials' compulsive fear of the branch being overwhelmed by black men with low intelligence. Gray once again used GCT score patterns to justify to his superior the Army's reluctance to abandon its racial quota. Whereas other branches of the Armed forces were "in the fortunate position of being able to maintain an enlistment standard of 90 GCT and higher," the Army was restricted by the Selective Service Act to use a GCT cutoff of 70 before calling upon the draft.²⁵¹ Black recruits were much more likely than white counterparts to receive low GCT scores, and the group of potential recruits with scores between 70 and 90 was thus included a disproportionate amount of black recruits relative to the general population. For Gray, the Army faced the burden of accepting those men who were not intelligent to get into other military branches—and the relative strength of the army corps was thus compromised by

²⁵⁰ *Ibid.*

²⁵¹ Gordon, memo to Johnson, May 26, 1949, p. 5.

this dilution of manpower. In the absence of a racial quota, Gray forecasted the Army would shoulder the immense burden of accepting black men who other branches deemed unfit to defend the American State, warning that “an Army of even 15 percent Negroes suffers a jeopardizing decreases in efficiency.”²⁵² By expressing GCT scores as a representation of military manpower, efficiency, and cooperativeness, the Army could present its racial quota as the soundest way to open opportunity for black soldiers without risking defensive chaos.

Although the Fahy Committee remained flummoxed by the Army’s resistance to abandon the racial quota, its response to the Army’s third series of proposals revealed the limits to the postwar state’s sense of racial justice. In a July 6 memo to other committee members, Executive Secretary E.W. Kenworthy accused the Army of having “the same disregard for efficiency as for equal opportunity.”²⁵³ Although the Army’s proposals to eliminate racial quotas in training divisions considerably expanded the ability for black soldiers to reach specialized training, its insistence on maintaining racially segregated military units rendered any manpower gains moot. Kenworthy found the Army’s plan to apply its 10 percent racial quota within certain military occupations even more disastrous because it made “numerical representation the criterion of equal opportunity.”²⁵⁴ For the Fahy Committee, postwar defense had to be premised on efficiency. Eliminating the racial quota and desegregating units were not primarily matters of redressing social injustice, but instead, means for ensuring that the military’s internal organization, training procedures, and promotion protocols placed the most capable men in the correct places.

²⁵² Gordon, memo to Johnson, May 26, 1949, p. 6.

²⁵³ E. W. Kenworthy, memo for the Fahy Committee, re: “Outline Plan for Utilization of Negro Manpower to Provide Further Equality of Opportunity”—The Army’s Third Reply, July 6, 1949, p. 2; RG 220 Record of [President’s Committee], box 6.

²⁵⁴ E. W. Kenworthy, memo for the Fahy Committee, July 6, 1949, p. 4.

At the end of his memo, Kenworthy suggested a potential solution for the Army's stubbornness: replace the racial quota with a General Classification Test Quota. If the Army resisted abandoning the racial quota was because it feared its ranks would be overrun with black recruits, Kenworthy reasoned, then a quota based on GCT scores could effectively keep the Army from adopting black men it deemed undesirable without resorting to overtly racial methods. The Army could set a cap on the number of men it would accept with GCT scores between 70 and 90; as monthly recruitment reports from the Army confirmed, a disproportionate amount of black recruits earned scores within that range compared white counterparts.[cite] As Fahy would eventually suggest to the Army, the cap would be based on the "the normal distribution in the Army as revealed by the operation of Selective Service during the war"—in other words, using preexisting test data to set the parameters for what the Army would continue to consider acceptable degrees of intelligence.²⁵⁵ The Army could also use other standardized aptitude tests to solve the problem of reenlisting men with low GCT scores. Men with marginal scores on both the GCT and the Aptitude Area I test could be barred from continuing their service beyond the initial term. Kenworthy posited that this new quota "would not be racial discrimination, but a defensible procedure for keeping the Army from having a heavier percentage of Class IV men than there is in the population."²⁵⁶ By shifting the basis of the quota, the Army could outwardly redefine who was marginally useful to the state—and thus whose presence had to be strictly monitored and regulated.

The Fahy Committee also gathered its own population figures to deduce that the Army's fears of being disproportionately filled with black recruits of low intelligence had no statistical

²⁵⁵ Charles Fahy, memo for Secretary of the Army, re: "Evaluation by the President's Committee of the Army's 'Outline Plan for Utilization of Negro Manpower to Provide Further Equality of Opportunity,'" July 25, 1949, p. 6; RG 220 Record of [President's Committee], box 6.

²⁵⁶ E. W. Kenworthy, memo for the Fahy Committee, July 6, 1949, p. 7.

bearing. Kenworthy used current population data and existing GCT score patterns to see under what circumstances the Army would become at least 30% black. For the Army to reach that proportion, it would have needed just over 200,000 black voluntary enlistees among its troops. Yet, statistically speaking, fewer than 140,000 physically healthy black men in the national manpower pool would have scored over 70 on the GCT. Even if the Army shifted from voluntary to mandatory service, the Army mathematically could not be saturated with the amount of black men that it feared unless it abandoned the GCT or expanded its age range for service.²⁵⁷ In case of national geopolitical emergency, the authorized size for military branches would likely expand, even further reducing the odds black men would represent a disproportionate share of Army strength.

Still, Kenworthy acknowledged that the difficulty for the Army—and the difficulty in convincing the Army to adopt a test-based quota—hinged on how it would manage black soldiers whose test scores prevented easy placement. After drafting the committee’s interim report, Kenworthy detailed to Fahy the infrastructural problems that could result from abandoning the racial quota. The “professional privates” whose scores barely qualified them for service would be easy to place in a new system: in the same types of menial jobs their aptitude scores indicated would be the best use of their manpower.²⁵⁸ So, too, would be black soldiers who scored highly on the GCT, as desegregated training schools and unit placements would also ensure the Army was making the most of their capabilities. But Kenworthy remained flummoxed by the considerable number of black soldiers “who have higher scores, and who are now in

²⁵⁷ E. W. Kenworthy, memorandum for the Fahy Committee, re: “Statement by the Army that Abolition of Quota Would Result in Army Having 30-40% Negroes,” June 27, 1949; RG 220 Record of [President’s Committee], box 6.

²⁵⁸ E. W. Kenworthy, memorandum for Charles Fahy, May 5 1949, p. 1; RG 220, Box 1.

Negro combat units, but who do not have the qualifications for school.”²⁵⁹ Dismantling these units immediately would create an organizational nightmare, by Kenworthy’s estimation.

For several months, the Fahy Committee and the Army lobbed proposals and counterproposals at one another, each side using different statistical interpretations of available American manpower to determine the degree to which a GCT quota could be trusted by itself to control the proportion of black troops. The committee cushioned its insistence on abolishing the racial quota by suggesting that the Army could accept fewer marginal-scoring men if it simply accepted more men with higher scores. The GCT quota would thus not be a rigid quantity to fill but, instead, a ceiling. Army General Counsel Karl Bendetsen argued that department should retain its 10 percent racial quota, but only for the men with the most marginal set of GCT scores. Fahy dismissed this offer as an “absurd” way to curb black recruitment. One Army statistician warned that there were “2,000 to 3,000 Negroes over [GCT score] 80 ready, willing and able to join the Army and Air Forces every month”—which, in the absence of a racial quota, would lead to black troops comprising up to 31% of the Army’s strength.²⁶⁰ Kenworthy took the same raw figures, Army intake caps, and potential GCT quota to determine black soldiers would comprise no more than 11% of all new recruits. (Soon after the Army realized the committee had somehow procured these figures, Kenworthy met a major dismayed that the department did not heed his advice to “never on any account submit figures to the Fahy Committee.”)²⁶¹ In a press release, the Army reiterated that its decision to offer training for all men who qualified for Military Occupational Specialties, regardless of race, would remain the full extent of its revised

²⁵⁹ Kenworthy, May 5 memo for Fahy, p. 2.

²⁶⁰ Austin J. Bonis, memo to Col. [John] McFadden, “Negro Percentages in the Army,” September 23, 1949; RG 220 Record of [President’s Committee], box 8.

²⁶¹ E. W. Kenworthy, memorandum for Chales Fahy, re: “The Status of Discussion with the Personnel and Administration Division on the Quota,” October 25, 1949, p. 1; RG 220, box 2.

policy.²⁶² By mid-November, Bendetsen admitted to Kenworthy that the Army's predictions were "based on the Army's over-all military experience in forecasting enlisting tendencies and trends and not on any specific experience"—but, nonetheless, nothing short of a numerical quota would suffice.²⁶³ Numbers became a rhetorical weapon, not for their specificity but their ability to be malleable enough to support vastly different scenarios.

The quota debate came to a head in late November, when the Army attempted to thwart the Fahy Committee's ambitions by instead proposing amendments to Circular 124 that would eliminate some, but not all, racialized limitations. The Army motioned to remove racial restrictions occupations and training schools, while retaining segregated units and racial quotas. Kenworthy observed that the GCT would make these changes effectively "meaningless": if black soldiers typically received low GCT scores, then black-only military units would retain a relative lack of occupational diversity, undermining the amount of manpower they could provide the Army.²⁶⁴ Unless the Army was willing to place black soldiers into existing white units based on occupational demands as well as ensure the most capable black soldiers entered service by enforcing a GCT quota—both measures the Army still refused to consider—then personnel organization and management would remain effectively unchanged.²⁶⁵

As stubborn as the Army remained in its extended discussions with the Fahy Committee, it soon officially abandoned its race-based personnel policies. At the beginning of 1950, the

²⁶² Department of Defense, memorandum to the press, no. 400-49, November 3, 1949; Charles Fahy, letter to Gordon Gray, November 8, 1949; Gordon Gray, letter to Charles Fahy, November 17, 1949; all RG 220 Box 5.

²⁶³ Karl R. Bendetsen, memorandum for [E. W.] Kenworthy, re: "Negro Manpower Pool," November 17, 1949; RG 220 Record of [President's Committee], box 4.

²⁶⁴ E. W. Kenworthy, memo for Charles Fahy, re: "Revised WD Circular 124," November 28, 1949, p. 2; RG 220, box 2.

²⁶⁵ E. W. Kenworthy, memorandum for the record, re: "Telephone Conversation with Mr. Fahy[,] 27 November," November 28, 1949; E. W. Kenworthy, memorandum for Charles Fahy, November 28, 1949; RG 220, box 2.

Army stipulated that black soldiers would henceforth be placed in units by basis of their qualifications rather than by race, and soldiers serving in mixed-race units would be integrated both on and off the job.²⁶⁶ Even so, Fahy still prepared statistical material to show how a GCT quota would aid the Army's desire "for eliminating the perennial low-scoring men, both white and Negro."²⁶⁷ By early March, Fahy received word that the Army would drop the racial quota—a couple months ahead of the publication of the Committee's final report.²⁶⁸ Although segregationist practices endured for another several years, the Army no longer used race as a precondition for entering service or the types of work one could do.²⁶⁹

For the Fahy Committee, the Army's racial quota and race-based personnel policies were unjust devices that punished individuals' capabilities based on broad statistical data about an entire segment of the population. By denying talented black men entrance into the Army while also retaining "professional privates," the Army squandered the potential depth of its manpower—the wasteful effects of which multiplied by keeping black soldiers out of most job tracks and in occupationally-homogenous units. The Committee advocated that the Army replace its racial quota with a GCT quota to guarantee that the brightest men, regardless of race, would be able to contribute their manpower. Yet this maneuver also established the racial logic of the postwar security state. Race may have been removed as an explicit condition for who could and

²⁶⁶ The President's Committee, *Freedom to Serve: Equality of Treatment and Opportunity in the Armed Forces* (Washington: United States Government Printing Office, 1950), 7; Department of Defense Office of Public Information, memorandum to the press, January 16, 1950, RG 220 box 9; Department of Defense Office of Public Information, "Army Revises Policy Governing Utilization of Negro Manpower, press release, January 16, 1950, RG 220, box 9; Department of the Army, Special Regulations No. 600-629-1, "Personnel: Utilization of Negro Manpower in the Army," January 16, 1950, RG 220, Box 9.

²⁶⁷ Charles Fahy, memorandum for the President's Committee, February 1, 1950, pp 1-2; Charles Fahy Papers, Box 3.

²⁶⁸ Charley Fahy, memorandum for Mr. Niles, March 13, 1950; RG 220, Box 8.

²⁶⁹ Randolph White, "Jim Crow Policy At Fort Dix Backed by Top Brass," *New York Amsterdam News*, October 14, 1950.

could not belong in the armed forces, but it remained foundational to the meaning made out of standardized testing data used to determine aptitude and ability. By shifting entrance caps from race quotas to GCT score minimums, the armed forces redirected how it used race to designate who was marginally useful, a policy shift that would reverberate in the postwar American state and society.

Standardized Testing as a Means for Postwar Racial Social Marginalization

The Army's use of GCT testing as a means to appear nondiscriminatory while maintaining a broader racial social order set a foundation for subsequent educational actors to use standardized testing as the means for designating black brainpower as marginally useful to state and society. This pattern developed over decades from several directions. During the 1950s and 1960s, psychologists began to fret about the potential social harm posed by underachievers, reinforcing in the process whiteness as an invisible norm for psychological types.

Psychometrically speaking, an underachiever was someone who scored high on standardized aptitude exams but whose performance on standardized achievement tests was relatively lackluster. In other words, their measured capabilities noticeably outpaced their measured skills.

Psychologists leaned on standardized aptitude and achievement tests in their assessment of underachievement because, by their lights, teachers frequently incorporated irrelevant and subjective factors when determining the relationship between students' capabilities and effort.

(Some notable researchers, namely educational psychologist Robert Thorndike, did stress that poor test design and statistical knowhow often created an exaggerated sense of underachievement among youth.)²⁷⁰ Although distinct from juvenile delinquents, underachievers and other gifted youth were more prone to deep-set emotional turbulence, and who, without

²⁷⁰ Robert Thorndike, *The Concepts of Over- and Underachievement* (New York: Columbia University, 1963).

proper intervention, would prove to be a societal drain in adulthood. Guidebooks advised counselors and parents alike how to encourage underachieving youth, overwhelmingly boys, into finding meaningful life paths.²⁷¹ Because earlier standardized tests such as the GCT had already placed black test-takers—and by extension black Americans in general—at the threshold of socially useful intelligence, the underachiever was a figure whose whiteness was unspoken but always presumed. If black youth by and large scored relatively low on standardized aptitude tests, particularly aptitude tests that reflected white middle-class educational experiences, then their relative level of achievement became hard to gauge. Whereas black underachievers would remain relatively invisible for another couple decades, white underachievers became a popular midcentury psychological phenomenon.²⁷²

What rendered black underachievers invisible for so long was a deep social-scientific pathologization of black families that portrayed educational disparities as a symptom of cultural and genetic shortcomings. This argument hinged on the insistence that low educational attainment among African-Americans, even if it reflected centuries of structures and systems designed to maintain a black underclass in the United States, primarily illustrated blacks' unwillingness to achieve due to genetically-eroded substandard intellects.²⁷³ This psychological

²⁷¹ See, among others: Merle M. Ohlsen, *The Extent to Which Group Counseling Improves the Academic and Personal Adjustment of Under-Achieving Gifted Adolescents* (Urbana: University of Illinois Press, 1960); Leonard M. Miller, ed., *Guidance for the Underachiever with Superior Ability* (Washington, D.C.: U.S. Government Printing Office, 1961); Maurice F. Freehill, *Gifted Children: Their Psychology and Education* (New York: MacMillan Company, 1961); Betty Jane Bosdell, *Evaluation of Counseling Treatments with Underachieving High School Students* (Grand Forks, ND: Cooperative Research Project, 1962); and Merle M. Ohlsen, *Appraisal of Group Counseling for Underachieving Bright Fifth Graders and Their Parents* (Urbana: University of Illinois Press, 1964).

²⁷² Charles H. Beady, Jr. and Stephen Hansell, "Teacher Race and Expectations for Student Achievement," *American Educational Research Journal* 18, no. 2 (Summer, 1981): 191-206; Carolyn Bennett Murray and James S. Jackson, "The Conditioned Failure Model of Black Educational Underachievement," *Humboldt Journal of Social Relations* 10, no. 1 (Fall/Winter 1982-1983): 276-300.

²⁷³ This argument had some of its foundation set, however unintentionally, by Lyndon Johnson's Assistant Secretary of Labor Daniel Patrick Moynihan's report *The Negro Family: The Case for National Action* (Washington, D.C.: United States Department of Labor, 1965). Moynihan, a sociologist, argued that the problems of

pathologization of black intellect depended upon a resurgent interest in standardized intelligence tests during the late 1960s and 1970s. In particular, certain psychologists argued that IQ testing was a meaningful way to deduce the heritability of intelligence, and by extension, the persistent gaps between white and black Americans' capabilities. For Arthur Jensen and his acolytes, the United States' economy and security increasingly relied on technological knowhow, which likely required IQ levels not typically found among black citizens. Black families transmitted both harmful cultural values and diminished mental capabilities to their children—and no amount of social policy or intervention could meaningfully enrich black youths' mental faculties. Society failed its low-intelligence citizens, these researchers argued, when it attempted to educate them as one would white, middle-class youth. For these psychologists, the humane and socially beneficial solution for low-intelligence youth would be to nurture those mental capabilities they did exhibit. Equal education opportunity, then, was the byproduct of misplaced white liberal guilt—a sentiment that caused more harm than good.²⁷⁴

the black underclass resided in its matriarchal structure; centuries of slavery and discrimination created black family structures at odds with mainstream (white middle-class) society, which only exacerbated social discord and economic precariousness. Impoverished black children who grew up without stable father figures, Moynihan observed, displayed noticeably lower IQs than national norms. As Moynihan fretted, among black eighth-grade students in Harlem, one-third scored “at levels perilously near those of retardation” (36). This lack of intelligence explained why nearly three-fifths of black recruits failed the Armed Forces Qualification Test—a standardized test Moynihan asserted “[measured] ability that ought to be found in an average 7th or 8th grade student” (40). Moynihan deduced that the state had an obligation “to bring the Negro American to full and equal sharing in the responsibilities and rewards of citizenship” (48), but indicated that policies should focus on the restructuring black families. See also: James S. Coleman, et al. *Equality of Educational Opportunity* (Washington, D.C.: U.S. Government Printing Office, 1966), esp. 217-333.

²⁷⁴ For examples of heritability models, and a sense of its longevity, see: Arthur R. Jensen, “How Much Can We Boost IQ and Scholastic Achievement?” speech given before annual meeting of California Advisory Council of Educational Research (San Diego: October, 1967); Sandra Scarr-Salapatek, “Race, Social Class, and IQ,” *Science* 174, no. 4016 (December 24, 1971): 1285-1295; Arthur R. Jensen, “A Reply to Gage: The Causes of Twin Differences in I.Q.,” *Phi Delta Kappan* 53, no. 7 (March, 1972): 419-421; Richard J. Herrnstein and Charles Murray, *The Bell Curve: Intelligence and Class Structure in American Life* (New York: Free Press, 1994); Neven Sesardic, “Philosophy of Science That Ignores Science: Race, IQ and Heritability,” *Philosophy of Science* 67, no. 4 (December, 2000): 580-602; and J. Philippe Rushton and Arthur R. Jensen, “The Totality of Available Evidence Shows the Race IQ Gap Still Remains,” *Psychological Science* 17, no. 1 (October, 2006): 921-922

For critiques of heritability models, see: N. L. Gage, “Heritability, Race Differences, and Educational Research,” *Phi Delta Kappan* 53, no. 5 (January, 1972): 308-312; Peggy R. Sanday, “An Alternative Interpretation

Some black parents did, however, successfully challenge the use of standardized testing as a diagnostic method that overemphasized intellectual disability among black youth. On October 16, 1979, after years of intermittent litigation, the United States District Court for northern California made a decisive ruling against state special education practices in the case *Larry P. v. Riles*, calling the fundamental premises of standardized aptitude testing into question in the process. *Larry P. v. Riles* stemmed from a complaint filed at the beginning of the decade against the San Francisco Unified School District.²⁷⁵ Parents of black students who had been placed in special education tracks for the “educable mentally retarded” (EMR) believed the standardized intelligence tests used to assess their children were fundamentally flawed. The litigants alleged that the cultural and racial biases built into these standardized tests—and which fueled their misinterpretations—led to a disproportionate amount of black students receiving an EMR diagnosis. Whereas African-Americans comprised 10 percent of California’s population in the 1970s, black youth made up a quarter of the students in EMR classrooms.²⁷⁶ Because students placed in EMR classes received severely limited educations compared to their

of the Relationship between Heredity, Race, and IQ,” *Phi Delta Kappan* 54, no. 4 (December, 1972): 250-254; Thomas Sowell, “Arthur Jensen and His Critics: The Great IQ Controversy,” *Change* 5, no. 4 (May, 1973): 33-37; Norman Daniels, “IQ, Heritability, and Human Nature,” *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science* 1974: 143-180; Paul Gombert, “IQ and Race: A Discussion of Some Confusions,” *Ethics* 85, no. 3 (April, 1975): 258-266; Margaret T. Gordon, “A Different View of the IQ-Achievement Gap,” *Sociology of Education* 49, no. 1 (January, 1976): 4-11; James Lawler, “IQ: Biological Fact or Methodological Construct?” *Science & Society* 41, no. 2 (Summer, 1977): 208-218; Roger Milkman, “A Simple Exposition of Jensen’s Error,” *Journal of Educational Statistics* 3, no. 3 (Autumn, 1978): 203-208; and William L. Conwill, “The Inheritance of IQ and Scholastic Achievement: Further Comments on the Jensen Article,” *Journal of Negro Education* 49, no. 1 (Winter, 1980): 97-104.

²⁷⁵ *Larry P. v. Riles* was not the first case to challenge the educationally discriminatory use of standardized tests. *Hobson v. Hanson*, decided in 1967, prohibited the use of existing tracking systems in Washington, D.C.’s school districts. This tracking system relied on standardized testing and caused a disproportionate amount of black students to receive the most basic level of education. However, *Hobson* also addressed a wider array of educational discrimination practices in the D.C. school district, and also considered the distinct role class also played in rendering inaccurate standardized test results. *Larry P.*, by contrast, created a narrower point of judicial examination. See: *Hobson v. Hansen*, 269 F. Supp. 401 (D.D.C. 1967); and Daniel J. Reschly, *Disproportionate Minority Representation in General and Special Education: Patterns, Issues, and Alternatives* (Des Moines: Iowa State Department of Education, 1997).

²⁷⁶ *Larry P. v. Riles*, 495 F. Supp 926, 1979, 931.

mainstream counterparts, the effects of a low standardized intelligence score snowballed throughout a youth's schooling. Black students who already came in with educational disadvantages were thus more likely to be marked as EMR, placed into separate education tracks, and leave school without many of the skills necessary for successful independent living.

In his decision, District Judge Robert Peckham used the history of intelligence testing to determine that standardized aptitude tests were racially and culturally biased instruments, and thus designed for discriminatory educational practices.²⁷⁷ As Peckham deduced, racial ideology drove the meaning made out of test scores. For decades, black children typically scored about fifteen points lower on IQ tests than their white counterparts. Yet, as virtually every expert witness testified, even if researchers came to an agreement over what intelligence was, existing intelligence tests could not measure intelligence directly. Peckham declared the state's method for identifying youth as EMR ultimately violated black children's civil rights as well as denied them equal protection under state and federal law. The judge thus maintained that standardized testing became a method for the state to beg the question of black intelligence: education officials already presumed that black youth displayed disproportionately lower intelligence more frequently than white children, and the evidence provided by biased testing instruments confirmed that presupposition.²⁷⁸

Despite some legal limitations on how governing authorities could use testing to

²⁷⁷ What's worth noting is that Peckham spent relatively little space in his ruling on the actual questions in the tests used to diagnose EMR youth. His decision gave far more contextual historical analysis than textual item-based examination. This approach would be noted and refuted in a subsequent case on the question of racial bias in standardized tests as a diagnostic tool for low-intelligence students in Chicago schools. This historical approach would also be publicly criticized by one of the psychologists who testified on behalf of the San Francisco schools. See: *Parents in Action on Special Education v. Hannon* 506 F.Supp. 831 (1980); and Nadine M. Lambert, "Psychological Evidence in *Larry P. v. Wilson Riles*: Evaluation by a Witness for the Defense," *American Psychologist* 36, no. 9 (September, 1981): 937-952.

²⁷⁸ *Larry P. v. Riles*, 933.

marginalize its black citizens, white Americans used the logic of standardized testing to resist educational admissions policies designed to undo systemic racial disadvantages. For many white Americans, affirmative action educational policies were—and remain—synonymous with reverse discrimination.²⁷⁹ Standardized testing gave the fantasy of reverse discrimination evidentiary heft: How fair was it to deny white students educational opportunities when, according to standardized admissions tests, they had greater likelihood for success than admitted minorities with considerably lower test scores?

The connection between standardized testing scores and the logic of reverse discrimination gained momentum through the Supreme Court case *Regents of University of California v. Bakke*.²⁸⁰ Allen Bakke, a white engineer who sought admission to UC Davis Medical School, filed suit against the California state university system after being denied admission in 1973 and 1974. Although multiple factors minimized Bakke's odds for admission—his relatively late application in one cycle, a low interview score by a member of the

²⁷⁹ For both historic and enduring white opposition to affirmative action, as well as a sense of the malleability of both support and opposition to affirmative action policies among white (both as a methodological concern and a political tactic), see: Nancy Stein, "Affirmative Action and the Persistence of Racism," *Social Justice* 22, no. 3 (Fall, 1995): 28-44; Charlotte Steeh and Maria Krysan, "Trends: Affirmative Action and the Public, 1970-1995," *Public Opinion Quarterly* 60, no. 1 (Spring, 1996): 128-158; Jennifer L. Hochschild, "The Strange Career of Affirmative Action," *Ohio State Law Journal* 59, no. 3 (1998): 997-1037; David R. Williams, *et al.*, "Traditional and Contemporary Prejudice and Urban Whites' Support for Affirmative Action and Government Help," *Social Problems* 46, no. 4 (November, 1999): 503-527; Jeffrey M. Jones, "Race, Ideology, and Support for Affirmative Action," Gallup website, August 23, 2005, <http://www.gallup.com/poll/18091/race-ideology-support-affirmative-action.aspx> (last accessed June 12, 2017); "Public Backs Affirmative Action, But Not Minority Preferences," Pew Research Center website, June 2, 2009, <http://www.pewresearch.org/2009/06/02/public-backs-affirmative-action-but-not-minority-preferences/> (last accessed June 12, 2017).

²⁸⁰ One Supreme Court case that preceded *Bakke*—*DeFunis v. Odegaard*—attempted to challenge similar admission procedures at the University of Washington Law School. Marco DeFunis, a white Jewish applicant, filed suit against the law school after being denied admission. DeFunis's counsel argued that his LSAT scores, used alongside other metrics, predicted their client would have been a more capable law student than admitted Chicano, indigenous, Filipino-American, and African-American students with considerably lower predicted first-year averages. Although the Supreme Court heard oral arguments for *DeFunis* in 1974, they ultimately ruled that the fairness of the law school's admission policies had been rendered moot; the school had not only ultimately admitted DeFunis, but had allowed him to complete his legal studies regardless of the ruling. See: *DeFunis v. Odegaard*, oral Argument, February 26, 1974, transcript at <https://www.oyez.org/cases/1973/73-235> (last accessed June 13, 2017); *DeFunis v. Odegaard*, 416 U.S. 312 (1974).

admission committee, his relatively advanced age (medical schools at the time were loathe to admit students in their thirties)—Bakke and his lawyers decided to focus on his whiteness. Bakke’s counsel argued that his solid MCAT scores, among other metrics, proved he would be a far more capable medical student than any of candidates chosen to fill the sixteen spaces UC Davis reserved for minority applicants. This racial quota, Bakke’s counsel reasoned, unfairly reduced their client’s chances for admission because only eighty-four spaces were available to Bakke and other white applicants. Bakke’s counsel insisted that, by denying white applicants a full opportunity for admission while also allowing minority applicants a chance to fulfill a quota, UC Davis violated Bakke’s rights to equal protection and nondiscrimination established by the 14th Amendment and Civil Rights Act of 1964. Although the Supreme Court’s ruling in *Bakke* was disjointed and convoluted—colleges and universities could consider race as a positive attribute in admissions, but could no longer use quotas and quota-like practices to guarantee minority candidates would have a presence in student bodies—it did reinforce the idea that standardized test scores marked genuine capability. By extension, affirmative action opponents could more easily use the *Bakke* ruling to accuse colleges and universities that actively aimed to undo intergenerational educational discrimination through their admission policies of engaging in tyrannical, anti-empirical social engineering.²⁸¹

Even in subsequent Supreme Court cases that upheld more holistic collegiate affirmative action practices, jurists continued to portray standardized testing as a technology that revealed true capabilities in light of liberal educational activism. In *Grutter v Bollinger*, the court affirmed

²⁸¹ Stephen Gillers, “Affirmative Action: The Dilemma of Making Amends,” *The Nation* December 25, 1976: 691-692; *Regents of the University of California v. Bakke* Oral Argument, October 12, 1977, transcript available at <https://www.oyez.org/cases/1979/76-811> (last accessed June 12, 2017); *Regents of Univ. of California v. Bakke*, 438 U.S. 265 (1978); Nancy E. Dowd, “*Bakke* and *Weber*: The Concept of Societal Discrimination,” *Loyola University Law Journal* 11, no. 2 (Winter, 1980): 297- 326; Howard Bell, *The Bakke Case: Race Education, and Affirmative Action* (Lawrence: University Press of Kansas, 2000); and Goodwin Liu, “The Causation Fallacy: *Bakke* and the Basic Arithmetic of Selective Admissions,” *Michigan Law Review* 100, no. 10 (March, 2002): 1045-1107.

that the University of Michigan's narrow consideration of race as a categorical positive in law school applications was an acceptable practice, neither falling into the quota trap nor violating white applicants' 14th Amendment rights. In his dissenting opinion, however, Justice Clarence Thomas upheld the use of the standardized law school admission test (LSAT) as a worthwhile gauge of students' readiness for advanced studies. If collegiate practices such as legacy admissions already made the concept of meritocracy a farce, Thomas reasoned, then law schools should employ the LSAT as they see fit—so long as those practices did not correct for well-known racial differences in test scores. Even in the majority opinion, Justice Sandra Day O'Connor asserted that LSAT scores could not be eliminated outright for law school admissions because "they are important (if imperfect) predictors of academic success in law school"²⁸² As with *Bakke*, a certain standardized test had already become a stable unit of information used by virtually all accredited degree-granting graduate programs for a specific profession. Thus, even if jurists and educators alike felt that the LSAT (or MCAT or GMAT) had its predictive limits, the basis by which people could be admitted by other categorical or socio-historical factors was limited to the extent that the LSAT (or MCAT or GMAT) served as a de facto unit of objective information. Holistic affirmative action admission policies thus held test scores as clear data, however limited and qualified, against which supplementary applicant information could be weighed. Indeed, the qualified use of standardized test scores strengthened their educational utility. Now universities could portray LSAT, MCAT, GMAT and other post-baccalaureate admission exam scores as a component with checks and balances in a just admissions system.²⁸³

²⁸² *Grutter v. Bollinger*, 539 U.S. 306, at 315.

²⁸³ Robert M. Hendrickson, "Rethinking Affirmative Action: Redefining Compelling State Interest and Merit in Admission," *Peabody Journal of Education* 76, no. 1 (2001): 117-135; William C. Kidder, "Does the LSAT Mirror or Magnify Racial and Ethnic Differences in Educational Attainment? A Study of Equally Achieving 'Elite' College Students," *California Law Review* 89, no. 4 (July, 2001): 1055-1124; *Grutter v. Bollinger* Oral Argument, April 01, 2003, transcript available <https://www.oyez.org/cases/2002/02-241> (Last accessed June 12, 2017); Phoebe

The final action that solidified standardized testing as a method for maintaining black social marginalization was the development of minimum competency testing in the late 1970s and early 1980s. Minimum competency testing emerged mostly in southern states amid the backlash to mandated desegregation, as well as in response to growing concerns that students were leaving high school functionally illiterate and innumerate. State-level school systems, heightening their capacity as an intermediary between the traditional local nexus of control and the growing federal role in education, introduced minimum competency tests as a way to monitor high school students' mastery of fundamental skills. In several states, these standardized basic skills tests became the determining factor for whether a high school senior received a diploma. Such minimum competency exams often focused on fundamental reading skills that state educational officials believed students ought to have developed over the course of their education. Yet, as some pundits at the time worried, such tests may have instead encouraged early cohorts of students to drop out before facing a final competency exam.

More often, as seen through a series of Florida court class action suits filed by black students and their parents, minimum competency testing regimes failed to accommodate for early cohorts of black students who received substandard segregated elementary education and who did not benefit from any early intervention programs or secondary school curricula geared toward ensuring mastery of basic reading and mathematical skills. Although these legal challenges delayed the timetable for when minimum competency testing could be used as a prerequisite for matriculation, the use of such tests nonetheless created, as Scott Baker argues, an

A. Haddon and Deborah W. Post, "Misuse and Abuse of the LSAT: Making the Case for Alternative Evaluative Efforts and a Redefinition of Merit," *St. Johns Law Review* 80 (2006): 41-105; and Kimberly M. Tatum, Joyce Coleman Nichols, and Fernandra Ferguson, "Examining Institutional Barriers that Impede Access to Legal Education" *Negro Educational Review* 59, nos. 1-2 (Spring/Summer, 2008): 79-92.

educational landscape in which the burden of education shifted from the school system to the student. High-stakes standardized accountability testing placed the onus of functional illiteracy and innumeracy on the products of the school system—students—rather than the system itself. This shift allowed state-level politicians to avoid the onus of increasing educational spending in ways that would upset white constituents. The surface-level neutrality of accountability testing programs rendered social marginalization as individual shortcomings. Because the system was no longer overtly discriminatory, test scores could offer the sheen of fair and genuine measurements of students' skills—and whatever shortcoming black students faced on such exams were no longer the problem of school systems.²⁸⁴

Accountability standardized testing regimes persist, more ingrained into curricular design and pedagogical practice than ever. What also continues is the burden placed on black students for their persistent low test scores, without any meaningful intervention into educational funding policies that leave students in majority-black schools comparatively ill-equipped to master material on such standardized tests. Because accountability regimes often peg school funding to standardized test scores, students in chronically underfunded areas—typically majority-black urban schools—are compelled to accept testing regimes in order to prove they merit social

²⁸⁴ *Debra P. v. Turlington*, 474 F. Supp. 244 (M.D. Fla. 1979); Jerrold R. Coombs, "Can Minimum Competency Testing Be Justified?" *High School Journal* 62, no. 4 (January, 1979): 175-180; Jacob G. Beard, "Minimum Competency Testing: A Proponent's View," *Educational Horizons* 58, no. 1 (Fall 1979): 9-13; Lorrie Shepard, "Issues in Minimum Competency Testing," *Review of Research in Education* 8 (1980): 30-82; Daniel P. Resnick, "Minimum Competency Testing Historically Considered," *Review of Research in Education*, 8 (1980): 3-29; W. James Popham, "The Case for Minimum Competency Testing," *Phi Delta Kappan* 63, no. 2 (October, 1981): 89-91; Martha M. McCarthy, "Minimum Competency Testing for Students: Educational and Legal Issues," *Educational Horizons* 61, no. 3 (Spring, 1983): 103-110; *Debra P., a Minor, by Irene P., Her Mother and Next Friend, et al., Plaintiffs- Appellants, v. Ralph D. Turlington, Individually and As Commissioner Of Education, et al., Defendants-appellees*, 730 F.2d 1405 (11th Cir. 1984); Linda B. Gambrell, "Minimum Competency Testing and Programs in Reading: A Survey of the United States," *Journal of Reading* 28, no. 8 (May, 1985): 735-738; Scott Baker, "The Paradoxes of Desegregation: Race, Class, and Education, 1935-1975," *American Journal of Education* 109, no. 3 (May, 2001): 320-343; and Scott Baker, "Desegregation, Minimum Competency Testing, and the Origins of Accountability: North Carolina and the Nation," *History of Education Quarterly* 55, no. 1 (February, 2015): 33-57.

investment, even if the resultant school funding hinges on “teaching to the test” practices detrimental to robust learning and development. The consequences for not accepting these testing regimes—or rather, failing to perform to these standards—often involve public disinvestment from problematic schools, the introduction of profit-driven charter schools, and the collapse of educationally-oriented community investment.²⁸⁵ Through such testing mechanisms students are forced to play along with the charade of equal educational opportunity without ever being provided the economic means that could make equal opportunity possible, let alone desirable. Black students thus often have the most to lose in accountability testing systems.

Meanwhile, researchers continue to examine the phenomenon of the achievement gap between white and black students on standardized tests, without interrogating the racial pathologization that undergirds such an idea. Psychologists and pundits repeatedly question why black students fare poorly on annual accountability exams, often anchoring their explanations in sociocultural rationale rather than economic explanations. By ignoring the economic foundations for standardized testing disparities—and, by extension, the way race is concomitant with economic inequality in American society—educational pundits continue to locate low black test scores as a phenomenon explainable by students’ ineffable blackness. As such, black students who perform outside the expected band of marginal performance continue to evade

²⁸⁵ See, for example, the *Tampa Bay Times* investigative series “Failure Factories,” which examined the Pinellas County School System throughout 2015 and 2016, available at <http://www.tampabay.com/projects/2015/investigations/pinellas-failure-factories/>. See also: Thomas J. Kane and Douglas O. Staiger, “The Promise and Pitfalls of Using Imprecise School Accountability Measures,” *Journal of Economic Perspectives* 16, no. 4 (Autumn, 2002): 91-114; Brian A. Jacob, “Accountability, Incentives, and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools,” *Journal of Public Economics* 89 (2005): 761-796; S. Michael Gaddis and Douglas Lee Lauen, “School Accountability and the Black-White Test Score Gap,” *Social Science Research* 44 (2014): 15-31; and J.T. Richardson, “Accountability Incentives and Academic Achievement: Distributional Impacts of Accountability When Standards are Set Low,” *Economics of Education Review* 44 (February, 2015): 1-16.

psychological diagnosis: black students are still markedly less likely to be diagnosed as gifted.²⁸⁶ Black nerd subcultures, meanwhile, are examined for the phenomenon of “acting white,” as though black Americans could not conceive of fandom and passionate study otherwise.²⁸⁷ The concept of the achievement gap has, meanwhile, endured its ability to dislocate black students from the structural bases for their disadvantageous standardized testing scores. The logic for the achievement gap emerged in the early postwar period, when the American state began to shift from overtly racist institutional policies toward technologies that maintained, to a considerable degree, the country’s foundational racial order. Standardized testing allowed a psychometric sheen for racial disparities in the military and, eventually, higher education—all while absolving the state from culpability in the persistence of a black underclass. Race-patterned score gaps, whatever the psychometric nomenclature, continue to be made the burden of black students

²⁸⁶ Laura Isensee, “In Houston’s Gifted Program, Critics Say Blacks and Latinos Are Overlooked,” *Morning Edition*, National Public Radio, September 30, 2015, audio at <http://www.npr.org/sections/ed/2015/09/30/441409122/in-houstons-gifted-program-blacks-and-latinos-are-underrepresented>; Anya Kamenetz, “To Be Young, ‘Gifted’ and Black, It Helps to Have a Black Teacher,” nprED, National Public Radio, January 20, 2016, <http://www.npr.org/sections/ed/2016/01/20/463190789/to-be-young-gifted-and-black-it-helps-to-have-a-black-teacher>; and Jason A. Grissom and Christopher Redding, “Discretion and Disproportionality: Explaining the Underrepresentation of High-Achieving Students of Color in Gifted Programs,” *AERA Open* 2, no. 1 (January-March 2016): 1-25; ; and Ibram X. Kendi, “Why the Academic Achievement Gap is a Racist Idea,” *Black Perspectives*, October 20, 2016, <http://www.aaihs.org/why-the-academic-achievement-gap-is-a-racist-idea/>.

²⁸⁷ Signithia Fordham, “Black Students’ School Success: Coping with the ‘Burden of ‘Acting White,’” paper presented at the Annual Meeting of the American Anthropological Association (Washington, D.C., December 3-7, 1985), copy available at <http://files.eric.ed.gov/fulltext/ED281948.pdf>; Mary Bucholtz, “The Whiteness of Nerds: Superstandard English and Racial Markedness,” *Journal of Linguistic Anthropology* 11, no. 1 (2001): 84-100; Carolyn Tyson, William Darity, Jr., and Domini R. Castellino, “It’s Not ‘A Black Thing’: Understanding the Burden of Acting White and Other Dilemmas of High Achievement,” *American Sociological Review* 70, no. 4 (August, 2005): 582-605; Cheryl Corley, host, “Does Race Play a Factor in ‘Nerdiness?,” *Tell Me More*, National Public Radio, August 1, 2007, transcript available at <http://www.npr.org/templates/story/story.php?storyId=12419664>; “New ACT Report Shows Many African-American Students Inadequately Prepared for Postsecondary Education,” ACT Media Relations, March 26, 2014, <http://www.act.org/content/act/en/newsroom/new-act-report-shows-many-african-american-students-inadequately-prepared-for-postsecondary-education.html>; Nia-Malika Henderson, “What President Obama Gets Wrong about ‘Acting White,’” *Washington Post*, July 24, 2014, https://www.washingtonpost.com/blogs/she-the-people/wp/2014/07/24/what-president-obama-gets-wrong-about-acting-white/?tid=a_inl&utm_term=.118ba420ce7e; and Andre M. Perry, “Stop Blaming Black Parents for Underachieving Kids,” *Washington Post*, July 30, 2014, https://www.washingtonpost.com/posteverything/wp/2014/07/30/stop-blaming-black-parents-for-underachieving-kids/?utm_term=.4e153b9de3da.

rather than the product of standardized testing regimes that naturalize white middle-class educational opportunities as an ideal norm.

CHAPTER FOUR: BAD AT MATH: STANDARDIZED TESTING, MATHEMATICS ACHIEVEMENT, AND GENDER POLITICS

Over the past several decades, the relationship between girls and mathematical achievement has gone from being seen as a largely unremarkable fact of life to a chronic social problem. Girls who are good at math pursue degrees and careers in mathematical fields less frequently than boys with similar talents. In the early postwar period, functionalist theory chalked this phenomenon up to sex roles: society ran more smoothly when men and women pursued life paths that fit commonly-accepted sex roles, and those who bucked those trends must have some abnormal qualities to their personality. Along with second-wave feminism came a thorough examination of why women remained in the margins of math and science—and many argued that the actual problem was the widespread social belief in mathematical talent as a distinctively masculine trait. Yet the belief persists among many researchers that something innate, some small bit of genetics, holds girls back from mathematical excellence. As these arguments stormed within academic and mainstream circles, the number of women pursuing high education vastly expanded. Women now earn the clear majority of postsecondary degrees, yet remain a minority in math and related subjects—with no clear consensus on why this phenomenon continues.

All along the way, standardized testing provided the fuel for researchers' arguments about the true reason girls had considerably different experiences with math than boys. For some

researchers, standardized test data provided a way to show that social attitudes toward proper feminine behavior made otherwise talented girls disinvest in mathematics during adolescence. For others, standardized testing data illustrated that girls were rewarded for behavior and obedience in the classroom rather than talent, and that boys naturally had a numerical edge among the most mathematically gifted youth. Some held an ambivalent position between these two extremes—but argued that the blunt-force misuse of standardized test scores by college admissions officer and scholarship organizations essentially reinforced whatever social problems already existed by denying mathematically talented young women equal education opportunities. None of these positions, however, seriously questioned the use of standardized testing in itself. Critics of the findings gathered through one standardized math test would instead use other standardized testing data to stake a competing claim. Standardized testing data thus became an increasingly valuable currency in the ongoing argument about what, exactly, limited girls' mathematical achievement—an argument unlikely to go away anytime soon as educational policymakers and pundits consider how the United States can best develop experts in STEM fields.

Gender Ideology, Postwar Science, and Functionalist Theory

Gender is a foundational power relation in American society. Gender not only includes the social meaning of perceived biological differences—those which transform *males* into *men* and *females* into *women*—but also the laws, policies, practices, and symbolic representations that enforce a power dynamic in which men possess control over women. This system is typically rooted in the idea that society should reflect a presumed natural order in which heterosexual couples reproduce (both themselves and their values), and in which women's primary responsibilities revolve around children and the home. Gender thus not only involves the power

dynamics within interpersonal relations, but also the mechanisms used to give societies order and structure, as well as the cultural productions borne from social relations. As an ideology, gender is pervasive but incomplete—filled with inconsistencies, contradictions, and variations at all levels. Gender persists because it is replicated, however messily, at every social level, through institutions and with other forms of power, such as race or class. Gender also persists because of the repercussions for deviating too far from social norms: homosexuals, transgender and genderqueer people, masculine women, and feminine men have historically faced legal, occupational, and physical punishment for their actions and identities. Hence gender persists as a power dynamic because its malleability and incompleteness also makes it resilient.²⁸⁸

The modern American state has served an important but peculiar role in altering gender norms. On the one hand, the post-1945 American state has been a source of enumerated rights based on sex and gender, whether by federal legislation (the 1964 Civil Rights Act, Title IX of the Education Amendments of 1972), or judicial decision (*Roe v. Wade*, *Griswold v. Connecticut*, *Obergefell v. Hodges*). On the other hand, specific constitutional protections of rights on the basis of sex remain absent, despite a nearly successful push for the Equal Rights Amendment in the 1970s. The federal protections that do extend to equitable private sector treatment—namely, the Lily Ledbetter Act of 2009—only provide extended timelines for the right to file suit against discriminatory companies rather than mandate equal pay for equal work.

²⁸⁸ Candace West and Don Zimmerman, “Doing Gender,” *Gender and Society* 1, no. 2 (June 1987): 125-151; Dorothy E. Smith, *The Everyday World As Problematic: A Feminist Sociology* (Boston: Northeastern University Press, 1987); Joan Wallach Scott, *Gender and the Politics of History* (New York: Columbia University Press, 1988); Michael S. Kimmel, “Masculinity as Homophobia: Fear, Shame, and Silence in the Construction of Gender Identity,” in *Theorizing Masculinities*, eds. Harry Brod and Michael Kaufman (New York: Sage, 1991): 119-141; Judith Lorber, *Paradoxes of Gender* (New Haven: Yale University Press, 1995); Gail Bederman, *Manliness and Civilization: A Cultural History of Gender and Race in the United States* (Chicago: University of Chicago Press, 1996); Karin A. Martin, “Becoming a Gendered Body: Practices of Preschools,” *American Sociological Review* 63 no. 4 (August, 1998): 494-511; and Judith Butler, *Bodies That Matter: On the Discursive Limits of Sex* (New York: Routledge, 2011).

The federated structure of the United States, meanwhile, creates different gender-based rights and regulations based on where they live. The ability to push back against gender-based discrimination is complicated when an aggrieved class inhabits multiple marginalized social categories (e.g., black women) because American law—and by extension, social institutions related to the law—often cannot account for the unique state of discrimination borne from intersecting identities. The state, as a political entity, thus plays a role in the politics of gender, but in a haphazard and variable fashion reflective of the idiosyncratic features of American federalism.²⁸⁹

Experimental science has likewise been a historically powerful tool for maintaining and reinforcing gendered social relations. As a historical phenomenon, science remains attached to its social contexts, as does its practitioners. The scientific appeal to objectivity does not erase the political contexts for certain lines of inquiry, the basis of credibility, or burdens of proof. Scientific inquiry often naturalizes social gender disparities in several ways. Scientists often use male bodies (and, more specifically, white male bodies) as the default in their research—inadvertently relegating female bodies as defective variables or ignoring them altogether. Scientific research has also been used as the means to justify preexisting gender stereotypes and subordination; such research begs the question, but typically shields itself with a scientific appeal to objectivity in order to dodge claims of logical fallacy. These practices are reinforced by journalistic and layperson treatments of scientific studies that either exaggerate the results of a study or insert a gendered interpretation where none had explicitly existed before. The enduring

²⁸⁹ Kimberlé Crenshaw, “Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color,” *Stanford Law Review* 43 (July 1991): 1241-1299; Glenda Gilmore, *Gender and Jim Crow*: (Chapel Hill: University of North Carolina Press, 1996); Leslie Reagan, *When Abortion Was a Crime: Women, Medicine, and the Law in the United States, 1867-1973* (Berkeley: University of California Press, 1997); Joan Acker, “Inequality Regimes: Gender, Class, and Race in Organization,” *Gender and Society* 20, no. 4 (August 2006): 441-464.

result is that scientists persistently seek to explain gendered differences through biological rather than social causes. These explanations often also ignore the relationship between society and biology—namely, how long-term conditions for a certain social group may create observable biological peculiarities. Instead, non-genital organs are often assigned a sex and treated as though they exhibit discernably male or female tendencies. Above all, the human brain becomes gendered—“the male brain,” “the female brain”—in a way that often presumes that a brain’s capacity is predetermined by prenatal sexual development. The gap between scientific research and its nonscientific discussion thus frequently turns the brain into a sexed organ.²⁹⁰

Sitting at the intersecting gender politics of the state and scientific discourse, postwar American state-sponsored science maintained a complicated relationship with women in the field. Women did enter higher education in larger numbers after World War II, but their participation in scientific fields of study was complicated by social policies that privileged educating male veterans, as well as persistent social beliefs that limited women’s advancement in the workplace. Those women who did secure roles in the emerging postwar state-science complex—by and large women from the white middle class—often had to assure colleagues and broader society that their participation in basic and applied research would not come at the cost of their responsibilities as wives and mothers. These women scientists and engineers, working largely between first- and second-wave feminism, instead emphasized the importance of their trained brains to the state. Deemed “technocratic feminists” by the historian Laura Micheletti Puaca, many early postwar women engineers and scientists emphasized that their brainpower

²⁹⁰ Ruth Bleier, *Science and Gender: A Critique of Biology and its Theories on Women* (New York: Pergamon Press, 1984); Anne Fausto-Sterline, *Myths of Gender: Biological Theories about Women and Men* (New York: Basic Books, 1985); Londa Schiebinger, “The History and Philosophy of Women in Science: A Review Essay,” *Signs* 12, no. 2 (Winter, 1987): 305-332; Margaret W. Rossiter, “The Matthew Matilda Effect in Science,” *Social Studies of Science* 23, no. 2 (May, 1993): 325-341; Cordelia Fine, *Delusions of Gender: The Real Science Behind Sex Differences* (London: Icon, 2010).

was as useful to national security as that of their male counterparts. Yet, by insisting a “‘Woman’s Place Is In the Lab, Too,’” these technocratic feminists could also subtly assure colleagues and the general public that they were not opposed to traditional feminine spaces.²⁹¹ While these women made tremendous headway gaining acceptance in their field, they also operated largely without workplace protections or positions of power within their disciplines—ultimately limiting the way individual achievements benefited women within their fields, or were even recounted within the historical record.

Postwar social science, meanwhile, remained wedded to functionalist theory, which influenced the types of questions researchers asked about their human subjects. Within functionalist logic, social differences between men and women were not a problem but, rather, a core feature of society itself. Functionalist theories upheld “the status quo in sex role allocation, differentiation and reward”—that is, the maintenance of distinct sex-based social roles.²⁹² Boys and girls, then, needed to be socialized appropriately in order to fulfill social expectations—expectations rooted in sustained, naturalized, sexualized differences between men and women. As seen in one 1955 U.S. Department of Health, Education, and Welfare (HEW) publication on raising adolescents, the family unit was at once an incubator of appropriate roles and an arsenal against subversive ideas. HEW encouraged parents to “forestall criticism of a youth by helping him build up interests that will make him acceptable to his peers,” lest they risk their son to “be

²⁹¹ Laura Micheletti Puaca, *Searching for Scientific Womanpower: Technocratic Feminism and the Politics of National Security, 1940-1980* (Chapel Hill: University of North Carolina Press, 2014): 95; Margaret W. Rossiter, *Women Scientists in America, Volume 1: Struggles and Strategies to 1940* (Baltimore: Johns Hopkins University Press, 1982); Margaret W. Rossiter, *Women Scientists in America, Volume 2: Before Affirmative Action, 1940-1972* (Baltimore: Johns Hopkins University Press, 1995); Jennifer S. Light, “When Computers Were Women” *Technology and Culture* 40, no. 3 (July, 1999): 455-483; Amy Foster, *Integrating Women into the Astronaut Corps: Politics and Logistics at NASA, 1972-2004* (Baltimore: Johns Hopkins University Press, 2011); Londa Schiebinger, “Has Feminism Changed Science?” *Signs* 25, no. 4 (Summer, 2000): 1171-1175

²⁹² Jean Lipman-Blumen and Ann R. Tickamyer, “Sex Roles in Transition: A Ten-Year Perspective,” *Annual Review of Sociology* 1 (1975): 300.

considered a ‘sissy’—something any normal boy shies away from.”²⁹³ Tomboys who refused “to see satisfactions in the feminine role they [were] expected to play,” had been, like their sissy counterparts, failed by their parents and risked dismantling social order.²⁹⁴ Functionalist theory helped blur the line between what aspects of sex roles were natural and which were social constructs by suggesting that, in the end, parsing the difference was not productive.

Midcentury social scientists carried these ideas into their research on mathematical achievement, often naturalizing the difference between girls’ and boys’ math skills in the process. The very idea of achievement was frequently discussed in gendered terms: the social problem of the underachiever was often framed as an issue largely faced by boys and men. (One trade paperback went so far as to assert “80 percent of children with educational problems are boys.”²⁹⁵) Although researchers typically found that girls fared better in many verbal standardized tests, boys scored better in certain mathematical standardized tests, particularly those involving spatial reasoning and arithmetic reasoning. Some psychologists who did not find sex differences in achievement test scores openly questioned whether their own findings were baseless or if the test-makers had already reformatted their test without anyone’s knowledge.²⁹⁶ Psychologists also observed marked sex differences in school achievement—while girls outperformed boys in grade school, men fared better than women in college—but also suggested that the “cultural artifacts” which created these trends were not at odds with a well-functioning

²⁹³ Marion L. Faegre, “What roles are boys and girls expected to play?” in *The Adolescent in Your Family* (Washington, D.C.: U.S. Department of Health Education, and Welfare—Children’s Bureau, 1955): 60, 58

²⁹⁴ Faegre, “What roles are boys and girls expected to play?” 58.

²⁹⁵ Barry Bricklin and Patricia M. Bricklin, *Bright Child—Poor Grades: The Psychology of Underachievement* (New York: Dell, 1967), 3.

²⁹⁶ Marvin Powell, Henry A. O’Connor, and Murray Deutsch, “Are There Really Sex Differences in Achievement?” *Journal of Educational Research* 57, no. 4 (December, 1963): 210-212.

society.²⁹⁷ Girls performed better in math class because teachers rewarded cooperation and obedience, traits that boys had difficulty expressing at earlier ages. The solution for these problems in math achievement, some psychologists reasoned, was not to create sex-neutral pedagogy, but simply to have boys start school a year later: educational institutions would thus allow the underlying natural order of things to run more smoothly.

Some early postwar psychologists tried to make sense of women who bucked trends by faring better on math standardized tests than verbal ones. This sometimes involved the use of personality inventories to determine whether girls who did well on standardized math tests were also abnormally masculine.²⁹⁸ In one early postwar study, women at Santa Barbara College who scored disproportionately high on the quantitative components of the American Council on Education Psychological Examination were then administered the Minnesota Multiphasic Personality Inventory (MMPI).²⁹⁹ The researcher conducting the study, William Altus, categorized questions on the MMPI by their potential for revealing personality oddities, and compared the answers of mathematically minded women to their verbal-minded schoolmates. Altus deduced from the responses that a young woman skilled at math “tends statistically to be a somewhat more anxious, straightlaced, conventional, dysphoric person who dislikes to read”—a

²⁹⁷ Marvin Powell, *The Psychology of Adolescence* (Indianapolis: Bobbs-Merrill Company, 1963), 92.

²⁹⁸ George D. Yonge, “The Use of Masculinity-Femininity Measures to Account for Sex Differences in Problem Solving,” *California Journal of Educational Research* 12, no. 5 (November, 1961): 208-212, 220.

²⁹⁹ William D. Altus, “Personality correlates of Q-L variability on the ACE,” *Journal of Consulting Psychology* 16, no. 4 (August, 1952): 284-291. Not all researchers from this era who tested mathematically-inclined women (as demonstrated through both class performance and standardized test scores) with personality inventories found a preponderance of masculine traits. Some instead found such women to display more feminine personality traits than peers who did not pursue degrees in mathematics. Others indicated that the very premise was suspect, considering that girls outperformed boys on standardized mathematic achievement tests, and that the overall gap between low- and average-performing girls in mathematics remained narrower than the equivalent gap measure in boys: mathematics achievement could not be a masculine trait if girls empirically fared better on standardized tests. See: Philip Lambert, “Mathematical Ability and Masculinity,” *The Arithmetic Teacher* 7, no. 1 (January, 1960) 19-21; Marian Wozencraft, “Are boys better than girls in arithmetic?” *The Arithmetic Teacher* 10, no. 8 (December, 1963): 486-490.

lavender-tinted obsessive-compulsive type who, “beneath their Pollyanna-ish, conventional façade, [has] considerable anxiety and some prickly-pear attitudes.”³⁰⁰ In this line of thinking, women who learned differently than typical women—who learned like men—had to have been socially and psychologically abnormal.

Mathematical Feminism and Biological Essentialism in the Second Wave

During the late 1960s and early 1970s, feminist and liberationist politics began to change the kinds of questions many researchers asked in their work—and, in turn, mainstream publications influenced the way their audiences perceived academic examinations of socially radical questions. This often led readers to encounter feminist ideas through a filter of detached journalistic skepticism. Consumers of *Psychology Today* could, for example, read an article on observable differences between baby boys and girls—one peppered with observations that the nature-versus-nurture debate “now sets off furious blasts among interest groups from women’s liberation to gay liberation.”³⁰¹ Readers could also find this swirl of journalistic coverage and rhetorical dismissiveness in the pages of *Time*, which devoted many column inches during the time to actions of prominent feminists. Kate Millett could, at once, be on the cover of *Time* and regarded as “the Mao Tse-tung of Women’s Liberation” within the feature story.³⁰² Writers for the magazine portrayed Shulameth Firestone and Elizabeth Janeway as women with good premises but extremist conclusions—and observed that “[m]any of the new feminists are surprisingly violent in mood, and seem to be trying, in fact, to repel other women than attract

³⁰⁰ Altus, “Personality Correlates of Q-L Variability,” 290, 289.

³⁰¹ Michael Lewis, “Culture and Gender Roles: There’s no Unisex in the Nursery,” *Psychology Today* 5, no. 12 (May, 1972): 54.

³⁰² “Who’s Come a Long Way, Baby?” *Time* 96, no. 9 (August 31, 1970).

them.”³⁰³ This journalistic approach to women’s liberation—as Gloria Steinem would characterize it, an effort to “[report] it as a small, privileged, rather lunatic event instead of the major revolution in consciousness”—helped cultivate a public that became more aware of gender politics while calling into question the legitimacy of its tenets.³⁰⁴

Federal actions also changed the gender dynamics of educational policy and organizational culture during the 1970s. The broadest, most substantive changes came through Title IX of the Education Amendments of 1972, which was eventually signed into law by President Gerald Ford in May 1975.³⁰⁵ The title pegged equal educational opportunity to federal financial funding: schools at every level of education had to commit to nondiscriminatory practices if they wanted to keep receiving money.³⁰⁶ This requirement extended not only to curricular and extracurricular offerings, but also admissions, housing provisions, and retirement annuities. Title IX also created a new level of educational bureaucracy—the Title IX Coordinator—and mandated that such officials be easy to identify and access, whether at an individual university or in a state education agency. Yet the vast majority of state educational agencies failed to create Title IX bureaucratic structures within the one-year timeline mandated

³⁰³ Ruth Brine, “Women’s Lib; Beyond Sexual Politics,” *Time* 98, no. 4 (July 26, 1971); “The New Feminists; Revolt Against ‘Sexism,’” *Time* 94, no. 21 (November 21, 1969).

³⁰⁴ Gloria Steinem, “‘Women’s Liberation’ Aims to Free Men, Too,” *Washington Post* June 7, 1970: 192.

³⁰⁵ To see some sense of how the Nixon and Ford Administrations considered issues related to Title IX, as well as the shifting role of women in higher education, see: Department of Health, Education, and Welfare, “Charter of Secretary’s Advisory Committee on the Rights and Responsibilities of Women,” May 2, 1974; and Commissioner’s Task Force on the Impact of Office of Education Programs on Women, “A Look at Women in Education: Issues And Answers for HEW,” U.S. Office of Education, November 1972, Patricia Lindh and Jeanne Holm Files, 1974-77, Box 41, Gerald R. Ford Library, Ann Arbor, Michigan.

³⁰⁶ Per the exact language of the title, “No person in the United States shall, on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving Federal financial assistance.” See: U.S. Department of Education, “Title IX and Sex Discrimination,” Office of Civil Rights webpage, https://www2.ed.gov/about/offices/list/ocr/docs/tix_dis.html (last accessed February 23, 2017).

by the law.³⁰⁷ Although the federal government had created policy protecting the education of women, not all components of higher educational bureaucracy were quick to comply.

These broader social, political, and educational changes made it possible for certain researchers to challenge prevailing assumptions about why girls failed to find equal footing in mathematics. Some—particularly the University of Madison–Wisconsin mathematics educationist Elizabeth Fennema—used standardized testing data as the means to question the actual roots of sex differences in girls’ mathematical achievement. Before releasing her own research, Fennema had already tapped into the broader emergent feminist critique of schools as sites that reinforced sexist social hierarchies. Fennema engaged in conversations sparked by other educational writers who argued that, contrary to the prevailing notion that schools were spaces for girls to shine, existing educational curricula had been “denying girls and women the right to full development of their cognitive and emotional abilities.”³⁰⁸ Fennema believed that, through “cold empirical data,” feminist researchers could clearly illustrate that every aspect of American education—from elementary school through college—contained material, attitudes, and practices that kept women from seeing themselves as mathematically-minded beings.

In 1974, Fennema published a meta-analysis designed to challenge the prevailing assumption within both academia and the greater public that girls simply were not as capable at mathematics as boys. In a meta-analysis, researchers comb through previous studies on a topic,

³⁰⁷ Emily Taylor and Donna Shavlik, “Institutional Self-Evaluation: The Title IX Requirement” (working paper), Washington, DC: American Council on Education, October 1975, Patricia Lindh and Jeanne Holm Files, Box 31; Project on the Status and Education of Women, “Partial List of Actions Institutions Must Take Under Title IX,” Washington, DC: Association of American Colleges, November 1975, Patricia Lindh and Jeanne Holm Files, Box 32; “Forty States Face Title IX Charges,” *Peer Perspective* 2, no. 2, July 1976:1, and The Project on Equal Educational Rights, “Title IX and the States: A Report on State Agency Compliance,” Washington, DC: National Organization for Women, May 1976, Patricia Lindh and Jeanne Holm Files, Box 37.

³⁰⁸ Elizabeth Fennema, “Overwhelming Evidence of School Sex Discrimination,” review of *Sexism in School and Society* by Nancy Frazier and Myra Sadker, *Phi Delta Kappan* 55, no. 2 (October, 1973): 147.

often conducted over a number of years and with many particular focuses, and use statistical techniques to synthesize findings in a way that makes a more potent statement about what research suggests in general. Looking through fifteen years' worth of educational research, Fennema wanted to address two core questions: What, if anything, did studies indicate were actual differences between girls and boys in mathematics achievement? And what did the research suggest were the root causes of any differences? Fennema did not want her meta-analysis to be misinterpreted as some hunt for the exact age at which girls became naturally less mathematically capable than boys. Instead, Fennema used the method of meta-analysis to make a feminist critique of the American educational system: girls' diminished math skills were a symptom of learning environments that held little value in the idea that women could participate as equals. If the aggregate research indicated that boys had better math skills than girls, Fennema argued, then American educators had "a moral responsibility as well as an economic necessity" to thoroughly redesign American mathematics education to eliminate sex bias.³⁰⁹ Only then could girls live up to actual natural potential.

Fennema's meta-analysis, printed in the *Journal for Research in Mathematics Education*, revealed no deep-seeded differences in girls' and boys' mathematical skills. The existing research on children up to the third grade showed "no consistent significant differences" in sex-based mathematical capabilities, and studies on students between fourth and ninth grade were too erratic to make any meaningful overarching conclusions.³¹⁰ Among later grades, the only clear

³⁰⁹ Elizabeth Fennema, "Sex Differences in Mathematics Learning: Why?" *The Elementary School Journal* 75, no. 3 (December, 1974): 183.

³¹⁰ Elizabeth Fennema, "Mathematics Learning and the Sexes: A review," *Journal for Research in Mathematics Education* 5, no. 3 (December, 1974): 128. This is not to say that none of the studies Fennema used in her meta-analysis suggested no sex-based differences in mathematical skills. One study, conducted by Richard Stafford in 1972, suggested that qualitative reasoning had a sex-linked genetic component that, although greatly affected by environment, was at least observable when comparing standardized math test performances of different

point at which boys outperformed girls in mathematics was in standardized achievement tests administered to high-school seniors—and, even then, only one study had produced such results. In that study, conducted by Margaret Backman in 1972, sex marked a greater distinction between achievement test scores than ethnicity or socioeconomic status. Backman nonetheless cautioned against assuming these differences had any biological component, given that “the social implications of the physical differences [between males and females] can play an important role in the development of characteristic abilities.”³¹¹ Even the scant pieces of research Fennema could find suggesting boys out-tested girls in math already came with the disclaimer that nothing about the results was strictly a matter of biology.

Fennema deduced that the only reasonably clear trend was that high school boys had greater tested strength in “higher-level cognitive tasks”—leading her to conclude that this late-developing difference must have stemmed from a series of social causes.³¹² As Fennema noted, girls were less likely to take as many math courses as boys, while boys were more likely to drop out of high school; these concurrent phenomena created high schools in which boys who remained often had more extensive mathematical instruction and thus tested better on standardized tests and in studies. For Fennema, the social basis for these differences meant that sexism pervaded both American education and mathematics-education research. The presumption that girls would eventually fare worse than boys in mathematics created lazy research methods and destructive classroom practices that held teenage girls back from their full potential. At the very least, teachers and researchers who understood that girls may approach

sets of male and female twins; see: Richard E. Stafford, “Hereditary and Environmental Components of Quantitative Reasoning,” *Review of Educational Research* 42, no. 2 (Spring, 1972): 183-201.

³¹¹ Margaret E. Backman, “Patterns of Mental Abilities: Ethnic, Socioeconomic, and Sex Differences,” *American Educational Research Journal* 9, no. 1 (Winter, 1972): 11.

³¹² Fennema, “Mathematics Learning and the Sexes,” 137.

math differently than boys yet found no incentive to develop new learning methods only reinforced the ideas that girls' approach to the subject was abnormal, unorthodox, and imprecise.

Fennema soon teamed up with University of Wisconsin–Madison psychologist Julia Sherman to develop an instrument for measuring girls' comfort with mathematics. Using a grant from the National Science Foundation, the researchers developed the Fennema-Sherman Scales in 1976 to empirically gauge girls' perceptions of parent and teacher interest in their math studies, their levels of confidence and anxiety toward math, and “the degree to which [they] see mathematics as a male, neutral, or female domain.”³¹³ Fennema and Sherman sought to measure the factors that kept girls out of advanced math classes in high school despite displaying similar aptitude to boys when given standardized math tests at earlier ages. The pair built their namesake scales partly in reaction to published findings from the recently developed National Assessment of Educational Progress (NAEP), which indicated that marginal differences in math skills between 9- and 13-year-old boys and girls ballooned into gaps that could “only be described as overwhelming” by young adulthood.³¹⁴ Even though NAEP asserted that the score gap was “the result of different but systematic cultural reinforcements for the sexes,” Fennema and Sherman argued that the study suffered from poor sampling techniques and, worse yet, the results had

³¹³ Elizabeth Fennema and Julia A. Sherman, “Fennema-Sherman Mathematics Scale: Instruments Designed to Measure Attitudes toward the Learning of Mathematics by Females and Males,” *Journal for Research in Mathematics Education* 7, no. 5 (November, 1976): 325.

³¹⁴ Ina V. S. Mullins, *Educational Achievement and Sex Discrimination* (Denver: National Assessment of Educational Progress, 1975), 7, <http://files.eric.ed.gov/fulltext/ED115701.pdf> (last accessed February 19, 2017). Girls actually scored slightly higher overall than boys in the 9- and 13-year-old cohorts in the initial 1973 assessment. Within the 17-year-old cohort, however, boys outperformed girls in each of six tested areas—including concepts such as numeration and variables in which younger girls had fared better. These differences were even more severe among the group of adults tested (9). This pattern held among minors in the series of NAEP mathematics tests in 1978; see: National Center for Education Statistics, “Table 222.85: Average [NAEP] Mathematics Scale Score, by Age and Selected Student Characteristics: Selected Years, 1973 through 2012” (.xls file), *Digest of Education Statistics*, https://nces.ed.gov/programs/digest/d15/tables/dt15_222.85.asp (last accessed February 19, 2017).

been misreported in broader circles as a sign of boys' natural superior mathematical minds.³¹⁵

The Fennema-Sherman Scales, by contrast, could provide empirical evidence on mathematical attitudes that would be less prone to interpretive mangling among nonexperts.

Using their scales, Fennema and Sherman made the case that attitude was everything, at least when it came to achievement in mathematics. Fennema and Sherman argued that girls pulled back from taking additional math classes in later grades “because mathematics is perceived as a male domain” in broader society.³¹⁶ Girls instead gravitated toward fields of study that reflected socially sanctioned sex roles. This conditioned avoidance of mathematics led girls to pursue fewer extracurricular activities that strengthened their spatial reasoning skills as well as underestimate their problem-solving skills—social phenomena misread as a natural phenomena when comparing girls' and boys' performance on relevant standardized test questions. Conversely, when Fennema and Sherman compared standardized math test performances for boys and girls who took a similar number of math courses, significant differences between the two groups evaporated. Although other factors, particularly socioeconomic status, altered the attitudes girls and boys held toward mathematics, Fennema and Sherman maintained that girls' measurably lower confidence in learning math and stronger belief in mathematics as a men's field provided a clearer basis for test score gaps than any alleged natural differences.³¹⁷

³¹⁵ Mullins, *Educational Achievement and Sex Discrimination*, 15; Elizabeth Fennema and Julia A. Sherman, “Sexual stereotyping and Mathematics Learning,” *The Arithmetic Teacher* 24, no. 5 (May, 1977): 369-372.

³¹⁶ Elizabeth Fennema and Julia Sherman, “Sexual Stereotyping and Mathematics Learning,” *The Arithmetic Teacher* 24, no. 5 (May, 1977): 370.

³¹⁷ Fennema and Sherman discovered that among lower socioeconomic groups, success in school was seen as more of a feminine trait—and among higher economic groups, the same trait was viewed as a manlier role. For the remaining extent of Fennema and Sherman's research using their affect scales, see: Elizabeth Fennema and Julia Sherman, “Sex-Related Differences in Mathematics Achievement, Spatial Visualization and Affective Factors,” *American Educational Research Journal* 14, no. 1 (Winter, 1977): 51-71; Julia Sherman and Elizabeth Fennema,

As Fennema and Sherman argued for the social basis of girls' mathematical demoralization, other women in mathematics created work for youth, adults, and colleagues alike that showcased women's longstanding contributions to the discipline. These efforts emerged as women mathematicians formally organized during the early 1970s. Realizing that American academic mathematical associations paid little attention to the small amount of women within their ranks, American University professor Mary Gray and several other women mathematicians created the Association for Women in Mathematics (AWM) in 1971. The AWM spent part of its early years not only holding well-attended panels on women's historical contributions to mathematics, but also printing biographical and thematic essays on women mathematicians in its newsletter, such as "Lady Lovelace and the Analytical Engine," "Emmy Noether: Twentieth Century Mathematician and Women," and "Women in Combinatorics."³¹⁸

Other researchers eventually developed biographical anthologies for broader audiences. Some anthologies, particularly those created by mathematics educator Teri Perl, integrated biographical sketches of past and present women mathematicians with short activities related to their innovations. In Perl's workbooks, readers examined parabola diagrams after learning about Hypatia of Alexandria, folded various three-dimensional models following an essay on Grace Chisholm Young's contributions to plane geometry, and translated decimals into binary numbers after reading the stories of the women mathematicians of the Masdison Academic Computing

"The Study of Mathematics by High School Girls and Boys: Related Variables," *American Educational Research Journal* 14, no. 2 (Spring, 1977): 159-168; Elizabeth H. Fennema and Julia A. Sherman, "Sex Differences in Mathematics Achievement and Related Factors: A Further Study," *Journal for Research in Mathematics Education* 9, no. 3 (May, 1978): 189-203;

³¹⁸ Lenore Blum, "A Brief History of the Association for Women in Mathematics: the Presidents' Perspectives," Association for Women in Mathematics website, <http://www.awm-math.org/articles/notices/199107/blum/> (last accessed February 3, 2017). For examples of biographical and thematic AWM essays from the 1970s, as well as AWM panels, see the digitized archive of *AWM Newsletter* at <https://www.drivehq.com/folder/p8755087.aspx> (last accessed February 3, 2017).

Center. While other authors created less interactive biographical anthologies, they shared a similar belief: the accomplishments of the women highlighted in these collections should not be taken as a sign that anything resembling gender parity in the field of mathematics had been achieved. Rather, these women became preeminent mathematicians amid a modern Western “feminine mathique” that placed mathematical proficiency “at variance with one’s womanhood.”³¹⁹ Their deep influence on mathematics, these authors emphasized, should be viewed as a reminder of the countless women discouraged from the field. How many Sophie Germain or Sonya Kovalevskys *could* have existed in the absence of patriarchal institutions and social mores?

Not all researchers ascribed to a social model for girls’ mathematical shortcomings, however. Fennema and Sherman’s main critics, the educational psychologists Camilla Persson Benbow and Julian Stanley, offered a competing hypothesis based on their observations of academically gifted youth during the early 1980s. Although Benbow and Stanley accepted Fennema’s claim that gender-related mathematical differences could be observed in early adolescent students, the pair argued that such differences stemmed from girls’ “less well-developed mathematical reasoning ability.”³²⁰ The pair argued that girls did not do worse than boys in mathematics because they took fewer math courses—but, instead, that girls took fewer math courses because they had inferior math skills, a fact reflected by their test scores. For Benbow and Stanley, “superior male mathematical ability” was most likely “an expression of a

³¹⁹ Lynn M. Osen, *Women in Mathematics* (Cambridge: MIT Press, 1974), 165; Teri Perl, *Math Equals: Biographies of Women Mathematicians + Related Activities* (Menlo Park, CA: Addison-Wesley Publishing, 1978); Teri Perl and Joan M. Manning, *Women, Numbers and Dreams: Biographical Sketches and Math Activities* (Santa Rosa, CA: National Women’s History Project, 1982); and Louise S. Grinstein and Paul J Campbell, eds., *Women of Mathematics: A Bibliographic Sourcebook* (Westport, Ct: Greenwood Press, 1987).

³²⁰ Camilla Persson-Benbow and Julian C. Stanley, “Sex Differences in Mathematical Ability: Fact or Artifact?” *Science* 210, no. 4475 (December 12, 1980): 1262.

combination of both exogenous and endogenous variables,” but differences in the math courses teenagers took carried no weight.³²¹

The pair also used standardized testing data to build their competing claim. The researchers analyzed the SAT performances of roughly 10,000 junior high-aged girls and boys, all of whom had taken exam as part of the long-term Study of Mathematically Precocious Youth.³²² Benbow and Stanley maintained that measuring younger adolescents’ mathematical capabilities through the SAT would eliminate the need to control for differential coursework; gifted junior high girls and boys were far more likely to have identical math classes than their senior high counterparts. Although gifted boys and girls performed equally well on the verbal portion of the SAT, boys consistently outperformed girls on the math portion of the standardized test. Further, boys in the study were far more likely than girls to receive high math scores on the SAT (at least 600 on the classic 200-800 scale)—and in any given cohort, the highest-ranking boy earned anywhere from 30 to 190 points more on the SAT math section than the highest-scoring girl.³²³ When students from earlier cohorts later took the SAT for college admissions purposes, the same general pattern held: boys continued to outperform girls on the math section. If the nation’s most academically gifted youth continuously produced gendered score gaps on the

³²¹ Benbow and Stanley, “Sex Differences,” 1264.

³²² Not every cohort used in the study had identical age/grade parameters, but overall, participants were usually either in the 7th grade or were “accelerated students of 7th grade age” (1262). In either case, the students were typically 4 or 5 years younger than the average high school juniors and seniors who took the SAT for its primary intended purpose.

³²³ Although Benbow and Stanley made note of these gaps, these score differences in themselves carry no inherent meaning. The statistical significance of a 30-point gap on a standardized test depends on where, exactly, those scores fall. Standardized scores are designed to cluster around a designated median; the further scores fall from the median, the less likely a person is to make that score and, in turn, the less change it would take to alter a test-taker’s performance to elicit a different scaled score. This means score gaps observed closer to the median are likely to be the result of a greater number of differently-answered questions similarly sized gaps found at a score scale’s poles. Whereas the 32-point difference Benbow and Person calculated among the average boy and girl test-taker in the January, 1979 cohort, for example, would carry more statistical meaningfulness than the nearly identical gap between the highest-scoring boy and girl in that cohort (790 and 760, respectively).

SAT math section before any of the students had a chance to differ in their math class choices, Stanley and Benbow concluded, then Fennema and Sherman's hypothesis did not hold up. Whether or not such boys eventually took calculus more frequently than girls could not change the fact that a foundational, partially innate, difference was already in place.

Benbow and Stanley's study, published in *Science* at the end of 1980, soon received criticism from various scientists, psychologists, and educational researchers. *Science* published an assortment of these critiques several months later. Some writers, such as the physicist Carl Tomizuka and consultant Sheila Tobias, argued that boys and girls' willingness to participate in the study was itself the byproduct of gendered social attitudes toward mathematics. If girls were typically more reluctant to join math programs for gifted youth, and often discouraged from participating in extracurricular activities that sharpened math skills, the pool of students available for the study had already been markedly socially influenced before early adolescence. Other critics, particularly the scientists Elizabeth Stage and Robert Karplus, found fault with the assumption that students who took similar courses had similar experiences in the classroom: girls and boys had been shown to receive considerably different treatment in math instruction even in elementary school. The mathematician Edith Luchins and psychologist Abraham Luchins questioned why, knowing how difficult it would be to control for pervasive cultural factors, Benbow and Stanley did not focus instead on designing question types that would eliminate sex differences in math test performance. One Department of Education employee wondered what, if anything, the SAT genuinely revealed about mathematical ability. The math section of the SAT consisted largely of word problems—a question type prone to content bias. The employee lamented that, more than anything, Benbow and Stanley's convenient depiction of the SAT as a

test of mathematical reasoning only contributed to the rhetorical laziness with which researchers generalized their findings:

“Should we ever discover a genetic and organic basis for mathematical ability, we can be certain at a minimum that some girls will have more ability than almost all boys—a subtlety that keeps getting lost in our ‘Boys are more or less X than girls’ language.”³²⁴

Benbow and Stanley chastised their critics for failing to understand the nuance within their findings. As the two researchers insisted, “superior male mathematical ability” and “greater male ability in spatial tasks” should not have been misinterpreted to mean boys were “inherently, intrinsically, or genetically abler” than girls—just that endogenous factors could be one of several reasons boys outperform girls on the SAT math section and spatial reasoning tests.³²⁵ To Stanley and Benbow, questions about earlier forces that discouraged girls from participating in the study evaded the numerical reality that fewer girls qualified—and those girls who did qualify for the study earned lower average scores and lower peak scores than the boys in their cohorts. Benbow was far more blunt in her statements to popular media about the meaning of her and Stanley’s research. Insisting that the observed test score data was accurate—that the numbers did not lie—Benbow stated that the backlash to her research stemmed from a discomfort with the truth: women could not “bring themselves to accept sexual difference in aptitude,” much less help girls “accept it and go from there.”³²⁶ For Benbow and Stanley, *Science* had not provided a

³²⁴ Susan Chipman, in “Letters: Mathematical Ability: Is Sex a Factor?” *Science* 212, no. 4491 (April 10, 1981): 116. Other letters referenced in this paragraph can be found in the same letters section.

³²⁵ Camilla Persson Benbow and Julian Stanley, response in “Letters: Mathematical Ability: Is Sex A Factor?” 118.

³²⁶ Camilla Persson Benbow, quoted in “The Gender Factor in Math: A New Study Says Males may be Naturally Abler than Females,” *Time* 116, no. 24 (December 15, 1980). See also: Dennis A. Williams and Patricia King, “Do Males Have a Math Gene?” *Newsweek*, December 15, 1980: 73.

space for meaningful criticism but “a forum for subjective judgments and anecdotal evidence.”³²⁷

By their logic, the math portion of the SAT had captured irrefutable data—something calculable and therefore true—that feminist politics willfully ignored.

Benbow and Stanley’s line of reasoning held value at a moment when sociobiology revived the idea of primordial bases for human social patterns. Stemming largely from the work of insect biologist E.O. Wilson, sociobiology argued that the most useful starting point for understanding human behavior is to observe similar behaviors found in other animals, and to consider the evolutionary and preservationist purposes for certain social phenomena. As Wilson suggested in the infamous final chapter of his 1975 work *Sociobiology: The New Synthesis*, “the time [had] come for ethics to be removed temporarily from the hands of the philosophers and biologized.”³²⁸ Altruism, territoriality, and homosexuality made far more sense, Wilson noted, if one first accounted for the behavior of chimpanzees and other mammals. Wilson’s arguments, along with those made by contemporaries such as David Barash and Mary Midgley, sparked a wave of contentious debates within scientific circles—soon morphing into a verbal slugfest in which sociobiology’s critics accused its supporters of being quasi-fascist eugenicists, and its defenders dismissed detractors as groupthink leftist conformists. Although sociobiology did not necessarily advance wildly novel ideas, it built upon the previous generation of sociological functionalist theory to offer an even more essentialist explanation for social sex roles: men and women were, in many ways, no different than the males and females of other species, whose social patterns also follow from the reality of sexual difference.³²⁹ More generally, it added

³²⁷ Benbow and Stanley, “Letters,” 121.

³²⁸ Edward O. Wilson, *Sociobiology: The New Synthesis* (Cambridge: Belknap Press of Harvard, 1975), 562.

³²⁹ Sociobiology Study Group of Science for the People, “Sociobiology: Another Biological Determinism,” *Bioscience* 26, no. 3 (March, 1976): 182, 184-6; Edward O. Wilson, “Academic Vigilatism and the Political

weight to criticisms of feminism. If women's liberation had the right ideas, critics mused, why hadn't these changes already come to fruition?³³⁰ While sociobiology did not directly influence Benbow or Stanley's research, the theory gave currency to academic arguments at the time that suggested deep-wired biology played some unavoidable role in the differences between men and women—that social forms followed biological function.

Benbow and Stanley followed up their initial report with additional studies that held the SAT as the most accurate gauge of sex-based differences in mathematical reasoning ability. The researchers repeated their analysis of precocious adolescent performance on the SAT with a second, larger group of boys and girls. The same patterns Benbow and Stanley observed with their earlier subjects reappeared. Far more boys than girls earned high scores on the math section of the SAT, and the ratio of boys to girls grew more severe as the level of performance increased.³³¹ Benbow and Stanley also analyzed the high-school math coursework, later SAT performances, and college majors of first-year college students who had participated in their earlier Study of Mathematically Precocious Youth. Overall, these students' average SAT scores were now roughly 200 points higher than the national average for college-bound high school seniors—but the gap between men and women's SAT math scores had inched even wider than what had been measured in their early adolescence. Although the young men and women had, on

Significance of Sociobiology," *Bioscience* 26, no. 3 (March, 1976): 183, 187-190; Nicholas Wade, "Sociobiology: Troubled Birth for New Discipline," *Science* 191, no. 4234 (March 19, 1976): 1151-1155; David Smillie, "The Challenge of Sociobiology," *American Psychologist* February 1979: 187-188; and W.R. Albury, "Politics and Rhetoric in the Sociobiology Debate," *Social Studies of Science* 10, no. 4 (November, 1980): 519-536

³³⁰ For one contemporary example of a writer musing over feminism's inability to overhaul certain career patterns among women and assuming biology must play some role in that lack of change, see: Samuel C. Florman, "Engineering and the Female Mind," *Harper's* February 1, 1978: 57-63.

³³¹ Camilla Persson Benbow and Julian Stanley, "Sex Differences in Mathematical Reasoning Ability: More Facts," *Science* 222, no. 4627 (December 2, 1983): 1029-1031. Among nearly-identical pools of seventh-grade boys and girls—roughly 20,000 in each case—the ratio of boys to girls who cleared certain level on the SAT math test grew from 1.5:1 for scores above 420 to 4.1:1 for scores above 600.

average, taken over four years' worth of math courses in high school, men typically began taking advanced math classes earlier than women and took more extensive coursework, particularly calculus. Despite this, gifted girls typically received slightly better class grades than boys, and were more likely in their first year of college to declare a major in mathematical sciences. Benbow and Stanley nonetheless concluded that, "[i]f one is interested in the question of why women do not pursue careers in mathematics and science as frequently as men do," one had to make use of data which provided a logical explanation for such phenomenon.³³² For Stanley and Benbow, only the gap in math SAT scores and what it suggested about innate differences offered that.

The two also dismissed the claims of researchers who challenged the usefulness of the SAT over other standardized math tests. What ensued was an argument over what, really, different standardized math tests measured and whether an individual's performance on one type of standardized math test gave any meaningful data about the skills measured by a second standardized math test. Educational sociologists Aaron Pallas and Karl Alexander reconsidered the differential coursework hypothesis, arguing that sex-related differences in SAT math scores more likely reflected the factors that kept girls out of advanced math classes rather than whatever skill differences already existed by the time girls and boys entered high school. Pallas and Alexander tested their hypothesis by calculating what factors caused a test score gap to emerge between when ninth-grade boys and girls took School and College Ability Test (SCAT) to when they took the SAT as high school seniors. The pair determines that test score gaps developed

³³² Camilla Persson Benbow and Julian Stanley, "Consequences in High School and College of Sex Differences in Mathematical Ability: A Longitudinal Perspective," *American Educational Research Journal* 19, no. 4 (Winter 1982): 619.

largely in high school and were largely “due to the sparse quantitative programs of study typically pursued by girls.”³³³

Stanley and Benbow countered Pallas and Alexander by pointing out, exasperatedly, the incompatibility of the standardized tests used in the study. Of course Pallas and Alexander found no sex-related test score gap in the SCAT only to see one emerge in the SAT: the tests measured drastically different aspects of mathematical ability. Performances on the SCAT, which was thick with word problems and focused on computations and arithmetic concepts, could not be used to gauge how test-takers fared on an algebra-dense test such as the SAT.³³⁴ When challenged more bluntly by other researchers to defend the interpretive significance of high SAT scores, Benbow asserted that superior math scores on the SAT “[related] closely to success in high-level, fast-paced mathematics and science courses, to educational acceleration, and (especially) to the choosing of careers requiring excellent quantitative ability.”³³⁵ The SAT remained useful for Benbow and Stanley because it was the only standardized test to give a reasonably clear idea of gifted adolescents’ likely college and career paths.

As Benbow and Stanley published and defended their findings, the demographics of American higher education and labor were in the middle of several major shifts. In 1970, U.S. colleges and universities enrolled about 50% more men than women: 4.4 million versus 3

³³³ Aaron M. Pallas and Karl L. Alexander, “Sex Differences in Quantitative SAT Performance: New Evidence on the Differential Course Hypothesis,” *American Educational Research Journal* 20, no. 2 (Summer, 1983): 179.

³³⁴ Camilla Persson Benbow and Julian Stanley, “Differential Course-Taking Hypothesis Revisited,” *American Educational Research Journal* 20, no. 4 (Winter, 1983): 469-473.

³³⁵ Camilla Persson Benbow, reply to Jonathan Beckwith and Michael Woodruff, in “Advancement in Mathematics,” *Science* 223, no. 4642 (March 23, 1984): 1248. Benbow soon after used the SAT scores of subjects in the Study of Mathematically Precocious Youth to determine the extent to which sex-based differences in SAT math scores revealed later differences in science achievement. See: Camilla Persson Benbow and Lola Minor, “Mathematically Talented Males and Females and Achievement in the High School Sciences,” *American Educational Research Journal* 23, no. 3 (Autumn 1986): 425-436.

million. While men's enrollment numbers rose throughout the 1970s, women's skyrocketed, nearly doubling by the end of the decade. During the 1980s, women's enrollment rates continued to grow as men's stagnated. By the end of the 1980s, about 1.25 million more women than men were enrolled in higher education. This broad shift, however, varied depending on students' ages. Men outnumbered women in U.S. colleges at every measured age level in 1970—but by 1980 men only outnumbered women among college students in their twenties. The only consistent point at which men outnumbered women at the end of the 1980s were among students aged 22 to 24. Despite the change in the composition of college students, the number of men who completed at least four years of higher education remained substantially higher than that of women during this period, even though more women had completed at least some college. Within a generation, women became the larger group on American campuses by a sizeable margin—but nonetheless remained less likely to have a four-year degree in hand by the time they reached twenty-five.³³⁶

The gender dynamics of the U.S. workforce also changed noticeably during this period. Between 1970 and 1990, the percentage of women over the age of sixteen in the U.S. labor force grew from roughly 43% to about 58%, while the percentage of men coasted downward from nearly 80% to 76%. During this same time period, the number of civilian men and women in the U.S. grew by nearly about 52 million—but as millions of women entered the workforce, millions of men left.³³⁷ When viewed alongside the rising age of first marriages during the same time

³³⁶ United States Census Bureau, "Table A-1 Years of School Completed by People 25 Years and Over, by Age and Sex: Selected Years 1940 to 2015" and "Table A-6: Age Distribution of College Students 14 Years Old and Over, by Sex: October 1947 to 2012" (XLS tables), CPS Historical Time Series Tables, U.S. Census Bureau website, <https://www.census.gov/hhes/socdemo/education/data/cps/historical/> and <https://www.census.gov/hhes/school/data/cps/historical/TableA-6.xls> (last accessed February 11, 2017).

³³⁷ United States Bureau of Labor Statistics, "Table 2: Employment Status of the Civilian Noninstitutional Population, 16 Years and Older, By Gender, 1948-2014 Annual Averages," *BLS Reports* 1059 (December, 2015): 11-16, <https://www.bls.gov/opub/reports/womens-databook/archive/women-in-the-labor-force-a-databook-2015.pdf>

period, these labor and education figures suggest a partial—but very incomplete—change for the terms of successful young adulthood for many American women.³³⁸ Attending college and entering the workforce took more and more priority for many women over getting married.³³⁹

Federal Fascination with the Gender Gap

By the late 1980s, the academic debate about “sex differences” in mathematical test performance had expanded into a broader discussion about the “gender gap.” The argument had expanded beyond the causes of the differences between men and women’s mathematical and scientific capabilities, growing to include debates about why the differences endured despite the persistent attention paid to it. Standardized testing was central to this shift. As more students took more standardized tests, particularly in mathematics, researchers had more data to test hypotheses and offer explanations.³⁴⁰ Depending on the study, the gender gap in math

(last accessed February 11, 2017). The number of women outside of the U.S. labor force remained relatively stable between 1970 and 1990, fluctuating between 41.2 and 43.4 million, even as the number of women over the age of 16 grew overall from 72.8 to 98.8 million in the same time period. Although the population of civilian noninstitutional men grew by nearly the same number—also increasing by about 26 million, from 64.3 million to 90.4 million—the number of men not in the labor force grew by nearly 9 million, to about 21.4 million.

³³⁸ Overall, the average age for first marriages rose by nearly four years for both American men and women between 1970 and 1990 (23.5 to 27.6 for men, and 21.5 to 25.4 for women). The proportion of American women who married between the ages of 20 and 24 correspondingly fell during this time period, from 3 in 5 to less than 1 in 3. See Michael R. Haines, “Long Term Marriage Patterns in the United States from Colonial Times to the Present,” National Bureau of Economic Research Historical Paper 80 (March 1996): “Table 2: Nuptuality Measures: United States, 1880-19[9]0,” <http://www.nber.org/papers/h0080.pdf> (last accessed February 11, 2017).

³³⁹ This statement must come with the caveat that the late-twentieth-century trend back toward later marriage and fuller participation in the workforce is a phenomenon largely for white women. The relationship between labor, education, and marriage has historically been—and remains—different for black women, Latinas, indigenous women, Asian-American women, Arab-American women, and foreign-born women. Put another way: the broader social forces that enabled white women to renegotiate their relationship to the workplace and the academy did not produce identical opportunities for racial minorities or other multiply-marginalized women. Nor were the feminist approaches adopted by many heterosexual white women identical to the theories and activist methods employed by women of color or lesbians. Such is the danger in observing broad, long-term national trends for a large demographic group—here, women—without paying attention to the other social identities that change experiences and political (re)action.

³⁴⁰ This was not only the result of more students going to college, but also the rapid growth of state-level high-stakes minimum competency testing during the 1970s and 1980s. In minimum competency testing, students pass by answering more than a predetermined percentage of questions correctly; in many state- and district-level uses, passing this type of test was used as the final requirement for a high school diploma. During this same period, many states turned to standardized tests as a means to hold school districts accountable. See: Robert L. Linn,

standardized test scores was either closing, holding steady, or resurfacing; girls were either continually underrated by tests or overrated in the classroom; comparing high school students' standardized test performances provided a useful gauge for assessing social pressures on adolescents or ignored more meaningful differences established early in grade school; and women's continued underrepresentation in math and science jobs was a profoundly social issue or a bit of feminist mythmaking. Because various researchers used different standardized tests to reach contradictory conclusions, testing data held currency for nearly all participants in the debate. Whether researchers and pundits believed standardized tests were biased tools that perpetuated social disparities, or that they were refined instruments that simply reflected social and educational problems, standardized testing data became a valuable commodity.³⁴¹

During the late 1980s, the federal government also took an interest in considering the role standardized testing played in perpetuating the gender gap, as well as the general extent of sex bias in standardized tests themselves. In a three-panel hearing before its Subcommittee on Civil and Constitutional Rights, the House Judiciary Committee examined the role standardized testing played in the math and science gender gap. But the testimony did less to point to one clear

"Assessments and Accountability," *Educational Researcher* 29, no. 2 (March, 2000): 4-16, <http://www.education.umd.edu/EDMS/MARCES/mdarch/pdf/1000052.pdf> (last accessed February 13, 2017); Edward B. Fiske, "Standardized Test Scores: Voodoo Statistics?" *New York Times* February 17, 1988, <http://www.nytimes.com/1988/02/17/us/standardized-test-scores-voodoo-statistics.html> (last accessed February 13, 2017).

³⁴¹ For a tiny sample of these various opinions about the gender gap during this period, as well as the differing ways various tests are wielded as rhetorical tools, see: Gina Kolata, "Gender Gap in Aptitude Tests in Narrowing, Experts Find," *New York Times* July 1, 1989: 1, 8; Lynn Friedman, "Mathematics and the Gender Gap: A Meta-Analysis of Recent Studies on Sex Differences in Mathematical Tasks," *Review of Educational Research* 59, no. 2 (Summer, 1989): 185-213; Joseph M. Horn, "Truth, Gender, and the SAT," *Academic Questions* 3 (1989-90): 35-39; Penelope V. Flores, "How Dick and Jane Perform Differently in Geometry: Test Results on Reasoning, Visualization, Transformation, Applications, and Coordinates," paper presented at Annual Meeting of the American Educational Research Association (Boston: April 16-20, 1990), <http://files.eric.ed.gov/fulltext/ED320915.pdf> (last accessed February 24, 2017); David C. Beidleman and Christine L. Cole, "Scholastic Aptitude Test Gender Gap," *American Secondary Education* 19, no. 2 (1991): 2-5; Constance Holden, "Is 'Gender Gap' Narrowing?" *Science* 253, no. 5023 (August 30, 1991): 959-960; Kate R. Sheehan and Mary W. Gray, "Sex Bias in the SAT and the DTMA," *Journal of General Psychology* 119, no. 1 (January, 1992): 5-14; and "Gender Gap Remains a Puzzle," *USA Today* 121, no. 2571 (December 1992): 10.

answer so much as indicate the sheer inextricability of standardized testing from the organization and economics of higher education. Among the first panel was *Ms.* editor Phyllis Rosser, who had written about standardized testing bias for both popular magazines and the National Center for Fair and Open Testing throughout the 1980s. Rosser asserted that the sheer volume of standardized tests used for admissions and scholarships created systemic sex bias in higher education. Over three million high-school students took one of three college-readiness standardized test (the SAT, ACT, and PSAT), and young women consistently earned higher marks in college courses than predicted by their test scores—calling into question the entire legitimacy of these exams. If the SAT were an accurate predictor of college-bound women’s first-year grades, Rosser argued, “girls would score 20 points higher than boys, rather than 61 points lower.”³⁴² This systemic under-prediction shut girls out of higher-ranking schools and scholarship competitions, eventually causing “a real dollar loss for females in later life, as they get less prestigious jobs, earn less money, and have fewer leadership opportunities.”³⁴³

Not all members of the subcommittee, however, found Rosser’s arguments equally compelling. Colorado Representative Patricia Schroeder—the only woman representative in the entire House Judiciary Committee—observed that her sixteen-year old daughter already had fallen prone to the disorienting effects of surprisingly low test scores: she and her friends began

³⁴² Phyllis Rosser, testimony for *Sex and Race Differences on Standardized Tests: Oversight Hearings before the Subcommittee on Civil and Constitutional Rights of the Committee on the Judiciary*, 100th House, 1st Session, April 23, 1987 (Washington, U.S. Government Printing Office, 1989), 3. As Rosser noted, the predictive capability for black women was even weaker than for women as a whole; black women’s first-year grades were not necessarily consistent with average SAT scores “43 points lower than black men, and 264 points lower than white men” (3). This double penalty, on basis of race and sex, cut across all non-white women who took the SAT (13).

³⁴³ Rosser, subcommittee testimony, 4. During the mid-1980s, the ratio of boys to girls who received National Merit Scholarships grew more severe, approaching 2:1 by 1986. Although the ratio has become slightly more equitable in the ensuing three decades—particularly after a complaint to the US Department of Education’s Office of Civil Rights by FairTest in the 1990s led to the College Board adding a writing section to the PSAT in an effort to produce more female National Merit Finalists—considerably fewer young women receive National Merit Scholarships than young men to this day; see: Bryan Nankervis, “Gender Inequity in the National Merit Scholarship Program” *Journal of College Admission* 220 (Spring 2013): 20-25

“to wonder if their high school performance, or if they had charmed their high school teachers, if maybe suddenly they’re not as good as they used to be.”³⁴⁴ Schroeder found it unconscionable that guidance counselors would likely discourage her daughter’s friends and many other girls from applying to certain schools strictly because of these test scores, despite many having studied math well beyond what was tested on the PSAT, ACT, and SAT. Alan Slobodin, the subcommittee’s Republican minority counsel, remained dubious. Slobodin questioned Rosser to define what, exactly, bias was—“I mean, if there’s a one point difference, would you consider that bias? How about five points?”—and whether the overall impact of test score differences was negative given how some fields, particularly law, had recently experienced a massive surge of women.³⁴⁵ How, Slobodin wondered, could these standardized tests be biased if two-thirds of their developers are women? And what if the classrooms themselves produced biased results? Although the subcommittee officially deemed the issue of sex bias a topic worth further examination, some members remained unconvinced the issue was as drastic as some witnesses professed.

Rosser submitted with her testimony a reading comprehension passage to illustrate the depths to which bias against women pervaded standardized testing content. The short passage, taken from the June 1984 version of the SAT, focused on the anthropologist Margaret Mead’s depiction of sex roles in food production. Although test-takers were asked to consider the short passage’s intent, form, and tone, the passage itself was riddled with several logical inconsistencies. The passage’s author chides Mead for making “such a dubious generalization”

³⁴⁴ Patricia Schroeder, subcommittee questioning, 61. It’s worth noting that Catherine A LeRoy served as Counsel for the Subcommittee on Civil and Constitutional Rights—contributing questions to this testimony as well—and that M. Elaine Mielke served as General Counsel for the larger House Judiciary Committee, but Schroeder was the only woman in either instance serving as an elected official.

³⁴⁵ Alan Slobodin, subcommittee questioning, 67.

about commonly-observed gendered divisions in food preparation—but only after directly quoting Mead’s insistence that such divisions “‘can be modified, and the modifications provide proof that the pattern itself is not something deeply biological.’”³⁴⁶ The author then proposes that two patterns actually exist: the basic pattern Mead described, and “one in which food is produced by the men with relatively little help from women,” which the author deems “the female and male systems of farming.”³⁴⁷ This argument not only was logically needless, as it did not depart from Mead’s quoted statement or the idea of sex roles in itself, but also muddled the lines between biological and social functions by deeming the variations “male” and “female.” Faulty logic aside, Rosser’s bigger concern remained the multiple effects such a passage would have on girls who took the SAT. In one of the few instances girls encountered a women researcher on a standardized test, they also saw her ideas discredited by an anonymous uncredentialed commentator. This experience, by Rosser’s estimation, would not only jar young women in the moment by possibly slowing them down and harming their test scores. Such an encounter could also have had an discouraging and enduring subconscious effect on their future paths of study.

The College Board and Educational Testing Service officials interviewed in the subsequent panel did not, unsurprisingly, share the same disdain for standardized testing as Rosser did. Their defense was that, if anything, standardized testing had opened more pathways for woman into colleges and universities than any other method. The gradual decline in women’s SAT scores was not a reflection of test bias, but of the expanding number of women who took the test. Between the mid-1960s and the early 1980s, women went from being the minority to the

³⁴⁶ Rosser, document submitted to subcommittee, 17.

³⁴⁷ *Ibid.*

majority of SAT candidates—and by 1987, the year of the subcommittee hearing, about 40,000 more women than men took the exam. This broad trend, these officials stressed, “correspond[ed] exactly to the times periods in which the scores of women declined.”³⁴⁸ Because the test-taking population was self-selecting (students voluntarily took the exam, regardless of their skill level and without any prerequisites), the larger number of women taking the SAT also included more racial minorities and women from lower socioeconomic households. These newer test-takers’ lower scores were not a reflection of some inferior capacity, officials stressed, but of their lower readiness for college based on poorer educational opportunities. If anything, the officials noted, the SAT and other standardized admission exams predicted first-year success for women accurately while over-predicting the academic success for men. For ETS and the College Board, the gender score gap simply reflected a broad influx of new college-goers, using the SAT and other standardized tests to access a world previous closed off to them.

To counter allegations of test bias, the testing officials submitted several documents that detailed the processes Educational Testing Service and the College Board had developed during the 1980s to reduce gender stereotypes and analyze peculiar testing results. ETS formalized its Sensitivity Review Process several years earlier to ensure tests “contained questions recognizing the varied contribution that minority members made to our society” and that “there was no inappropriate or offensive material in the tests.”³⁴⁹ Frequently, ETS changed the descriptive material of a question for the purpose of diversification without altering the tested concept; revising a grammar question to include “a representational women’s reference” to the poet Gwendolyn Brooks did not affect the question’s core purpose of seeing whether students could

³⁴⁸ Gretchen W. Rigol, subcommittee testimony, 71.

³⁴⁹ Educational Testing Service, *ETS Sensitivity Review Process: An Overview* (Princeton: Educational Testing Service, 1987), 4; in congressional testimony at 169.

identify a comma splice. ETS also removed sexist stereotypes and language from its test content—depictions of mechanical incompetence, “aviatrix” and other feminized diminutives, use of “generic he,” among many others—unless contextually necessary.³⁵⁰ Although the College Board and ETS insisted that the SAT and other standardized tests had useful predictive capabilities regardless of the demographic group in question, they did acknowledge that some individual questions may have elicited wildly different results between certain groups of test-takers and thus warranted “differential difficulty analysis.”³⁵¹ For each of the 60 math questions on the SAT, ETS and the College Board conducted a differential difficulty analysis between white men and white women, comparing the percentage of test-takers in each group who answered correctly. The test-makers calculated that the correlation between what these two groups found difficult were extremely close, even if the overall gender gap in test performance remained persistent. For ETS and the College Board, the continued differences between men and women on standardized tests, through real, were not the result of either organization failing to scrutinize and improve their measurement tools.

The question of who, exactly, was responsible for the persistent gender gap became no clearer when the congressional subcommittee took statements from an MIT admissions director and a testing reform executive. As the admissions director admitted, women’s lower scores in math and science standardized admissions exams led to a consistent under-prediction of their final grade point average. The SAT had failed to determine that women at MIT had the same

³⁵⁰ Educational Testing Service: *The ETS Sensitivity Review Process: Guidelines and Procedures* (Princeton, Educational Testing Service, 1986), 8, 18, 21-23; in congressional testimony at 223, 233, 236-238.

³⁵¹ Thomas F. Donlon “The SAT in a Diverse Society: Fairness and Sensitivity,” *College Board Review* 122 (Winter 1981-1982): 6; in congressional testimony at 143.

GPA as men, but lower dropout rates.³⁵² Even so, the admissions officer maintained that standardized test scores ideally played a supporting role in determining which applicants would be accepted; although nonselective schools would have little reason to make test scores the key barrier to entry, selective schools typically used them as part of a more complex formula. Yet Denise Carty-Bennia, the executive chair of FairTest, stressed in her testimony that over one-third of colleges and universities had a minimum SAT or ACT score required for admission. Schools kept these hard cutoffs, Carty-Bennia noted, despite the College Board's insistence that a student could potentially retake the SAT and earn a score 130 points higher or lower than their original marks without any genuine statistical difference. (Even discounting this extreme example, most SAT test-takers had a normal variation of about 30 points in either direction for each half of the exam.) When compounded with the number of scholarships premised on standardized test scores, a sizeable chunk of higher education admissions—not only who got in but also who was rewarded for their potential—had become based on an exacting, highly imprecise use of standardized tests that only reinforced existing social biases.³⁵³ This misuse of tests in admissions and scholarships passed the burden of back onto test-takers: the system may be unfair, but you also already knew the hurdle you needed to clear.

Although Congress's investigation of the gender gap stopped short of intervention, some state-level governments amended their educational policies during the late 1980s to reduce the

³⁵² Michael C. Behnke, statement to subcommittee, 284. The lower dropout rate is significant because the larger pool of women had the opposite effect than it did with SAT scores nationwide, in which a larger number of women led to lower average scores. It is also worth noting that the admission director, Michael Behnke, was also a Chairman of the New England Regional Council of the College Board and a member of the scholarship selection committee for the National Merit Corporation; see congressional testimony at 290. This is not noteworthy for being unusual so much as in the difficulty in completely separating testing bureaucracy from admissions bureaucracy.

³⁵³ Denise Carty-Bennia, statement to subcommittee, 297-301. As Carty-Bennia noted, the College Board advised admissions officials that “score differences of less than 66 points on the SAT verbal and SAT mathematical, respectively, [had] little significance” (300), a statistical grace window at odds with the admission procedures at many schools.

effects standardized tests had on girls' and boys' life paths. New York in particular restructured the way it used the SAT for its two largest college scholarship funds. As a cost-saving measure, New York State Education Department began using the SAT as the sole determinant for its Regents Scholarship and Empire Scholarship recipients during the late 1970s. After numerous complaints that this method discriminated against girls—who, despite being the majority of applicants, received well less than half of all state scholarships—the state temporarily switched to a formula that considered both SAT scores and high school grades. The state's legislature allowed this method to lapse, leading to class action suit filed on behalf of girls throughout the state. The plaintiffs in *Sharif v. New York State Education Department* alleged that the use of SAT scores as the sole determinant for New York's merit scholarships violated girls' constitutional right to equal protection as well as Title IX provisions. The defendants maintained that grade point averages were an unreliable way to gauge which students merited a scholarship—particularly after many elite schools fudged records to ensure more of their students received Empire and Regents Scholarships, thus maintaining the prestige of their programs.

The presiding judge for the U.S. District Court of southern New York ruled in favor of the plaintiffs. Judge John M. Walker reasoned that New York's use of SAT scores for scholarship funding was so far from the original intended use of standardized tests—and, even then, the SAT chronically underpredicted girls' college achievement—that its sole use in determining Regents and Empire scholarship recipients violated girls' equal protection and Title IX rights. Walker asserted that these violations occurred even if no discriminatory intent existed: good intentions were no defense for educational sex discrimination. The judge also noted, however, that New York could still use SAT scores to decide Regents and Empire awards if the

state also used other metrics, or that it could instead develop a standardized achievement test specifically suited for determining which students merited state funding. Even when states directly intervened in misuses of standardized testing that contributed to lifelong gender gaps in mathematical achievement, they never abandoned the idea that a fair standardized test could be developed to measure girls' genuine math skills. Specific tests and the way they were mishandled may have perpetuated social gender disparities in math and science, but standardized math testing itself was too integral to the business of higher education to remove completely.³⁵⁴

The Enduring Mystery of the Gender Gap

Decades after the debate about girls' mathematical skills began, several key features persist. Most notably, the gendered gap between men and women's relationship to math and science remains, despite continued shifts in the demographics of American higher education and the U.S. work force. Although women now outpace men in the number of postsecondary degrees they attain—accounting for over half of all doctorates, nearly 60% of all masters degrees, and about 57% of all bachelors degrees conferred in the 2013-14 academic year—they remain relatively underrepresented in certain scientific and mathematic fields. Whereas women receive more degrees in biomedical and psychological disciplines, men significantly outnumber women

³⁵⁴ *Sharif by Salahuddin v. New York State Educational Department*, 709 F. Supp. 345 (U.S.D.C. Southern District of New York), February 7, 1989; Kary L. Moss, "Standardized Tests as a Tool of Exclusion: Improper Use of the SAT in New York," *Berkeley Journal of Gender, Law & Justice* 4, no. 2 (1989): 230-244, <http://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=1027&context=bglj> (last accessed February 25, 2017); James Barron, "New Scholarship Test Sought By Albany Education Official," *New York Times* May, 26, 1988: B6; Mark A. Uhlig, "New York State Returns to National Tests for Scholarships," *New York Times* October 6, 1988: B1, B13; Nadine Brozan, "Suit Asserts Bias Against Women in State Scholarship Awards," *New York Times* November 30, 1988: B6; William Glaberson, "U.S. Court Says Awards Based on S.A.T.'s Are Unfair to Girls," *New York Times* February 4, 1989: 1, 50; and Philip S. Gutis, "New York Acts to Alter Awards of Scholarships," *New York Times* February 7, 1989: B2.

The following year, girls received more Regents Scholarships than boys—which, despite being a rather slim majority, was the first time it had happened in the 77-year history of the award. The New York Department of Education had calculated that using only the (had it been allowed) would have given boys a slight, but not overwhelming, majority—with the inverse holding true if the state only took school grades into account. See: Sam Howe Verhovek, "Girls Take 51.3% in Regents Awards," *New York Times* March 4, 1990: 28.

in computer science, engineering, mathematics, and physical science degrees; taken altogether, the gap within STEM fields has remained stagnant for the last several years.³⁵⁵ While women make up nearly half of the American workforce, as well as half of the college-educated active workforce, they only make up one quarter of workers in STEM occupations. Despite STEM occupations providing considerably higher wages and relatively smaller gender wage gaps than the average job, women with degrees in mathematics, engineering, technology or physical sciences remain “less likely than their male counterparts to work in a STEM occupation.”³⁵⁶ As STEM has grown into a prized, sometimes fetishized field of study, women remain a marginal

³⁵⁵ U.S. Department of Education, National Center for Education Statistics, Fast Facts: Degrees conferred by race and sex” <https://nces.ed.gov/fastfacts/display.asp?id=72> (last accessed January 28, 2017); “Table 318.30: Bachelor’s, master’s, and doctor’s degrees conferred by postsecondary institutions, by sex of student and discipline division: 2013-14” https://nces.ed.gov/programs/digest/d15/tables/dt15_318.30.asp?current=yes (last accessed January 28, 2017); and “Table 318.45: Number and percentage distribution of science, technology, engineering, and mathematics (STEM) degrees/certificates conferred by postsecondary institutions, by race/ethnicity, level of degree/certificate, and sex of student: 2008-09 through 2013-14” https://nces.ed.gov/programs/digest/d15/tables/dt15_318.45.asp?current=yes (last accessed January 28, 2017); all from Integrated Postsecondary Education Data System (IPEDS), Fall 2014, Completions component.

Women now outearn men in college degrees regardless of race or ethnicity; the bachelor degree gap for the 2012-2013 academic year ranged from 54% women among students of Asian descent to 65% women among African-American students (NCES, “Fast Facts”). Within biomedical and biological disciplines, women made up 58% of all bachelors recipients, 57% of all masters recipients, and 53% of all doctorate recipients during the 2013-24 academic year; in psychological disciplines, the respective figures were 77%, 80%, and 75%. In contrast, the same figures for computer science degrees were: 18%, 29%, and 21%; for engineering degrees: 18%, 24%, and 23%; for mathematics degrees: 43%, 41%, and 29%; and for physical sciences: 39%, 38%, and 33% (NCES, Table 318.30). Although the number of American students receiving degrees in STEM disciplines has grown substantially over the past several years—from about 472,000 during the 2008-2009 academic year to almost 604,000 in the 2013-2014 academic year—the ratio of women to all STEM degree recipients has remained stagnant at and beyond the bachelors level, hovering between 31 and 35% (NCES, Table 318.45).

³⁵⁶ David Beede, *et al.*, Office of the Chief Economist, U.S. Department of Commerce Economics and Statistics Administration, “Women in STEM: A Gender Gap to Innovation,” ESA Issue Brief #04-11, August 2011, available <http://www.esa.doc.gov/sites/default/files/womeninstemagaptoinnovation8311.pdf>; p.1. At the time of this report’s publication, women earned on average 79 cents to every dollar that men in did in non-STEM occupations; in STEM occupations, women earned on average 86 cents to every dollar a man did. Women in STEM occupations also earned, on average, more per hour than their non-STEM counterparts: \$31.11 per hour versus \$19.26 per hour (Beede *et al.*, p. 7). For a sense of what the federal government considers a STEM occupation or STEM undergraduate degree, see: David Langdon, *et al.*, Office of the Chief Economist, U.S. Department of Commerce Economics and Statistics Administration, “STEM: Good Jobs Now and for the Future,” ESA Issue Brief #03-11, July 2011, available http://www.esa.doc.gov/sites/default/files/stemfinaljuly14_1.pdf, pp. 8-9. For a sense of how women and men trained in STEM fields face different initial career pathways, see: Catherine Buffington *et al.*, “STEM Training and Early Career Outcomes of Female and Male Graduate Students: Evidence from UMETRICS Data Linked to the 2010 Census,” *American Economic Review: Papers & Proceedings* 106, no. 5 (2016): 333-338.

presence—exacerbating persistent wage gaps attributable to gender and race, particularly as the narrowing of those gaps has stagnated over the past decade.³⁵⁷ If standardized tests play some role discouraging or preventing women access into highly esteemed fields of study, then they ultimately hold some deleterious effect on women’s economic power.

As women continue to lag behind men in their pursuit of STEM degrees and occupations, researchers keep mining standardized testing data as a way to determine the onset and severity of the math and science gender gap. While many educational pundits continue to investigate the enduring SAT-M score gap, others have asserted that the SAT is too limited a test, administered at too late a stage in secondary education, to deliver a full set of observable differences. While an earlier generation of researchers used standardized testing data to fuel a debate about the core causes of the math and science gender gap, contemporary researchers are now more hesitant to

³⁵⁷ The degree to which wage differences can be solely attributed to gender—that is, clearly sexist pay differentials—remains a little hazy, but economists have done considerable work quantifying the gap while accounting for other variables. Economists and social scientists have also done considerable research accounting for the role race and ethnicity complicates the gender wage gap: minority women continue to make less than men of the same race or ethnicity, and markedly less than white men. The gap is also complicated by age: women in older age cohorts face a larger wage gap than their younger counterparts. The same goes for motherhood: mothers pay a price for having children, where men do not. How much of this is a generational issue (that is, women from older generations who made less than male peers double suffered from workplace environments without strong collective bargaining powers that could redress mid-career and late-career gaps, or which effectively penalized pauses in labor for motherhood) and how much of this is a built-in feature of workplace sexism (that is, women are systemically undervalued at each stage of their professional lives, thereby limiting the maximum pay raise they can achieve at any given stage, creating a snowballing wage gap as women age), remains to be fully calculated. This conundrum bears out for the figures from 2015, in which the median weekly income for men aged 45 to 54 was nearly \$250 higher than women in the same age cohort (\$1040 vs. \$799)—but men aged 16 to 24 already had a median weekly income \$60 higher than similarly aged women (\$510 vs. \$450). Even in the most conservative estimates, a wage gap between men and women unattributable to any factor besides gender itself persists. See: American Association for American Women, “The Simple Truth About the Gender Pay Gap: Spring 2017 Edition,” January 2017, <http://www.aauw.org/resource/the-simple-truth-about-the-gender-pay-gap/>; U.S. Bureau of Labor Statistics, “Highlights of Women’s Earnings in 2015,” Report no. 1064, November 2016, <https://www.bls.gov/opub/reports/womens-earnings/2015/pdf/home.pdf>; Francine D. Blau and Lawrence M. Kahn, “The Gender Wage Gap: Extent, Trends, and Explanations,” Discussion Paper no. 9656,” Institute for the Study of Labor, January 2016, <http://ftp.iza.org/dp9656.pdf>; Rebecca Glauber, “Marriage and the Motherhood Wage Penalty among African Americans, Hispanics, and Whites,” *Journal of Marriage and Family* 69, no. 4 (November, 2007): 951-961; Joan R. Kahn, Javier García-Manglano and Suzanne M. Bianchi, “The Motherhood Penalty at Midlife: Long-Term Effects of Children on Women’s Careers,” *Journal of Marriage and Family* 76 (February, 2014): 56-72; and William M. Rodgers and Leslie S. Stratton, “The Male Marital Wage Differential: Race, Training, and Fixed Effects,” Discussion Paper no. 1745, Institute for the Study of Labor, September 2005, <http://repec.iza.org/dp1745.pdf> (accessed February 1, 2017).

point to a singular cause.³⁵⁸ Whether inspired by sociobiological models of development, convinced that capitalism fuels disparities more than culture or biology could—or simply unwilling to retread old arguments—researchers seem less certain to ascribe a singular cause to the gender gap. More potently, whatever the argument, standardized testing data remains a highly valued way to consider women’s worth to society. Test scores continue to serve as both a psychological and an economic form of data, serving as a monitor for how well girls with mathematical skills become women who fulfill the state’s STEM needs.

More than anything, the persistent gender gap in standardized test scores has made it easy for numerous Americans to continue suggesting that there must be something essential—something innate and pure and true—about quantitative data that shows differences in men and women’s mathematical skills. The examples are both relentless and repetitive. In fact, what makes the contemporary period remarkable is how redundant the story has become. Consider, for example, Lawrence Summers’s remarks at a National Bureau of Economic Research conference in early 2005, in which the former Harvard president insisted “issues of intrinsic aptitude” contributed more to women’s underrepresentation in tenured engineering and science professorships than “socialization and patterns of discrimination.”³⁵⁹ For Summers, the

³⁵⁸ Muriel Niederle and Lise Vesterlund, “Explaining the Gender Gap in Math Test Scores: The Role of Competition,” *Journal of Economic Perspectives* 24, no. 2 (Spring, 2010): 129-144; Erin Cech and Mary Blair-Low, “Perceiving Glass Ceilings? Meritocratic versus Structural Explanations of Gender Inequality among Women in Science and Technology,” *Social Problems* 57, no. 3 (August, 2010): 371-397; Katherine M. Finnigan and Katherine S. Corker, “Do Performance Avoidance Goals Moderate the Effect of Different Types of Stereotype Threat on Women’s Math Performance?” *Journal of Research in Personality* 63 (2016): 36-43; Denise Schwery, David Hulac, and Amy Schweinle, “Understanding the Gender Gap in Mathematics Achievement: The Role of Self-Efficacy and Stereotype Threat,” *School Psychology Forum* 10, no. 4 (Winter 2016): 386-396; and Jamin D. Speer, “Pre-Market Skills, Occupational Choice, and Career Progression,” *Journal of Human Resources* 52, no. 1 (January, 2017): 187-246.

³⁵⁹ Lawrence H. Summers, “Remarks at NBER Conference on Diversifying the Science & Engineering Workforce” (Cambridge, January 14, 2005), Harvard University’s Office of the President website, http://www.harvard.edu/president/speeches/summers_2005/nber.php (last accessed April 7, 2014); Scott Jaschik, “What Larry Summers Said,” *Inside Higher Ed* 18 February, 2005, http://www.insidehighered.com/news/2005/02/18/summers2_18 (last accessed April 7, 2014).

“unfortunate truth” was simply that fewer women had intellects “three and a half, four standard deviations above the mean.”³⁶⁰ Or consider, similarly, the numerous studies over the past decade that have sought to make meaning out of test-based sex differences among the most extreme cases of mathematical talent. Even though many of these researchers ultimately discredit—or, at least, greatly minimize—the role biology plays in mathematical talent, they still feel obliged to frame their studies as investigations of the validity of biological essentialism. This framing suggests that, despite mounting evidence of the attitudinal and socioeconomic factors that prevent all but a few girls at select institutions from competing on a level mathematical playing field with boys, the idea that standardized testing data reveals genuine inborn aptitude still carries tremendous social cachet.³⁶¹ By grafting essentialist gender politics onto standardized test data, and by treating numerical information as something untainted by social influence,

³⁶⁰ Summers, “Remarks.”

³⁶¹ See, among others: Janet Shibley Hyde, “Biological, Social, and Organizational Components of Success for Women in Academic Science and Engineering,” paper presented at the National Academies Convocation on Maximizing the Success of Women in Science and Engineering (Washington, D.C.: December 9, 2005), text available at <https://www.ncbi.nlm.nih.gov/books/NBK23762/>; Diane F. Halpern *et al.*, “The Science of Sex Differences in Science and Mathematics,” *Psychological Science in the Public Interest* 8, no. 1 (August, 2007): 1-51; Luigi Guiso *et al.*, “Culture, Gender, and Math,” *Science: Education Forum* 320 (30 May, 2008): 1164-1165; Stephen Machin and Tuomas Pekkarinen, “Global Sex Differences in Test Score Variability,” *Science* 322, no. 5906 (November 28, 2008): 1331-1332; Allison Schrager, “Men are Both Dumber and Smarter than Women,” *Quartz* July 09, 2015, <https://qz.com/441905/men-are-both-dumber-and-smarter-than-women/>; Glenn Ellison and Ashley Swanson, “The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions,” July 2009, <http://economics.mit.edu/files/4298>; Johnathan M. Kane and Janet E. Mertz, “Debunking Myths About Gender and Math Performance,” *Notices of the AMS* 59, no. 1 (January, 2012): 10-21; Rachel Feltman, “Brains aren’t Actually ‘Male’ or ‘Female,’ New Study Suggests,” *Washington Post* December 1, 2015, https://www.washingtonpost.com/news/speaking-of-science/wp/2015/12/01/brains-arent-actually-male-or-female-new-study-suggests/?utm_term=.d37364060b39; Gary Stix, “Is the Brain Gendered? A Q&A with Harvard’s Catherine Dulac,” *Scientific American*, December 28, 2015, <https://blogs.scientificamerican.com/talking-back/is-the-brain-gendered-a-q-a-with-harvard-s-catherine-dulac/>; Anh Tan, *et al.*, “The Human Hippocampus is Not Sexually-Dimorphic: Meta-Analysis of Structural MRI Volumes,” *NeuroImage* 124, part A (January, 2016): 350-366; and Jesse Singal, “Here’s the Biggest Study Yet on the Differences Between Male and Female Brains,” *Science of Us* April 6, 2017, <http://nymag.com/scienceofus/2017/04/heres-the-biggest-study-yet-on-sex-based-brain-differences.html>; Lara Perez-Felkner, Samantha Nix, and Kirby Thomas, “Gendered Pathways: How Mathematics Ability Beliefs Shape Secondary and Postsecondary Course and Degree Choices,” *Frontiers in Psychology*, April 06, 2017, <https://doi.org/10.3389/fpsyg.2017.00386>; and Michael Price, “Study Finds Some Significant Differences in Brains of Men and Women,” *Science* April 11, 2017, <http://www.sciencemag.org/news/2017/04/study-finds-significant-differences-brains-men-and-women>.

Americans still entertain the idea that biology is the clearest arbiter of talent—and that only the most extreme cases of measurable talent can be used to frame an argument about gender and society. In this way, standardized testing has helped ideology shape destiny.

CHAPTER FIVE: THE BIRTH OF THE STUDENT-CONSUMER: STANDARDIZED TESTING, CONSUMER ADVOCACY, AND THE EDUCATIONAL MARKETPLACE

Standardized testing has become an evergreen item of scorn. Websites, think tanks, and periodicals devote considerable time and money producing commentary on the dangers of American education's obsession with, and dependence on, standardized tests. Whether speaking of high-stakes statewide testing schemes or college admissions exams, standardized tests also seem increasingly unpopular among American parents and students, particularly those in the white middle class, among whom opposition to nationwide standardized tests often serve as a proxy for political beliefs. Yet, while pundits and the general public often share a disdain for testing itself, standardized test preparation remains as popular—and profitable—as ever. The market remains flush with guidebooks, practice tests, tutoring programs and even coaches, all for the purpose of improving standardized test scores. The current state of educational testing is anchored in irony: few seem to embrace the degree or types of standardized testing students endure throughout their education, but many seem willing to participate in the test-preparation marketplace in order to score as well as possible. The tests may attract considerable disdain, but the socioeconomic stakes for poor performance are too high for many to avoid participating in the test preparation market.³⁶²

³⁶² This chapter builds substantively off the research I conducted for my master's thesis; see: Keegan Shepherd, "Selling 'Dream Insurance': The Standardized Test-Preparation Industry's Search for Legitimacy, 1946-1989," master's thesis, University of Central Florida, 2011.

This chapter examines how this knotty relationship developed between the late 1960s and the mid-1980s. During this period, middle-class white Americans expressed their dissatisfaction with standardized testing in two key ways: challenging the legitimacy of testing practices through consumer advocacy networks, and purchasing the products and services of test-preparation (“test-prep”) companies. Through both practices, standardized testing came to be publicly discussed as a consumable good. Testing companies, in turn, were posed as corporations subject to certain responsibilities, and test-takers (or, by proxy, their parents) were framed as consumers entitled to certain rights. Two groups of consumers ushered in these changes to standardized testing during this decade. On one front, young consumer advocates nurtured by Ralph Nader conducted a several-year investigation of Educational Testing Service (ETS)—by far the United States’ largest creator of standardized admissions exams. These consumer advocates, many of whom started as student activists, published their searing findings in 1980 as the report *The Reign of ETS: The Corporation That Makes Up Minds*. These consumer advocates also helped usher in New York’s truth-in-testing law, which mandated that test-makers divulge the correct answers and explanations for standardized exams. On the other front, standardized test-preparation (“test-prep”) entrepreneurs developed and expanded standardized test-preparation operations, both in terms of geography products offered. The mixed findings of Federal Trade Commission (FTC) investigation into the ethicality of test-prep services offered the industry enough legitimacy to continue diversifying into the 1980s.

The functional relationship between standardized tests, test-makers, and test-takers changed considerably by the mid-1980s. Whereas testing companies once held colleges and universities as the key consumers of standardized testing—and students as the testing candidates whose results created the data needed for test scores to have any meaning—by the mid-1980s,

students held a more dominant role. Those who took standardized admissions tests were no longer students but student-consumers, who had the right to question the legitimacy of educational gatekeepers by invoking their right to quality products. But transparency laws and an expanding test-prep industry helped transform student-consumers' satisfaction into a commodity in itself, creating an environment in which ETS and other test-makers bundled old tests for commercial publication, and test-prep companies increasingly staked their marketing on persona rather than content. This commoditized satisfaction did more to reinforce standardized testing as an educational gatekeeping device—and, in turn, cemented standardized testing as a key gatekeeping mechanism for the middle class. Test-prep devices and transparency measures made middle-class Americans feel a greater sense of ownership and participation in the test-taking process and thus solidified a pattern in which the clearest way to maintain intergenerational middle-class status was to score well enough on standardized tests to earn entry into a college or university socially designated as befitting one's class status.

Focusing on this consumerist turn—found in both the rhetoric of the publication *The Reign of ETS*, the FTC investigation of test-prep companies, and the actions of both test-prep companies and test manufacturers following transparency laws—illustrates several curious features about the American middle class during the late-twentieth century. In particular, the emergence of the student-consumer reveals the slipperiness with which consumer advocates, test manufacturers, and test-prep companies alike have equated test-takers with the future middle class and consumption with higher education. Indeed, then and now, these three entities have accused each other of defending the interests of a narrow band of Americans—the white-collar middle-class—while hiding behind the idea of “public good.” To this end, consumer advocates' push for transparency in standardized testing operations illustrates a rhetorical slight-of-hand

between “the middle class” and “public interest,” inasmuch as the problems of the former have been used to invoke the latter. This slippery language is much more than just word games, however. It reverberates throughout higher education policy debates, when the structural needs and privileges of the solidly middle class are used to dictate the priorities and ideals for the future U.S. higher education.

Consumerism as a Framework for Understanding Standardized Testing

What typically complicates the historiography on standardized testing is that researchers are often woven into in the history of standardized testing itself. Nicholas Lemann, whose work *The Big Test* remains the most notable critical inquiry on the SAT’s impact on American social structure, was eventually appointed on a 2008 National Association for College Admission Counseling commission that higher education institutions scale back their reliance on standardized admission test scores when selecting students.³⁶³ After writing *None of the Above*, a searing indictment of the logical legerdemain Educational Testing Service used to validate its tests and operations, David Owen composed the forewords for two publications by the Princeton Review—a test prep company whose marketing angle was premised on dismissing the SAT as pseudo-scientific hogwash.³⁶⁴ Conversely, the College Board and Educational Testing Service routinely smudge the author-subject relationship with their numerous institutional histories and defenses of standardized testing methods.³⁶⁵

³⁶³ National Association for College Admission Counseling, “Report of the Commission on the Use of Standardized Tests in Undergraduate Admission” (Arlington, VA: NACAC, 2008); Nicholas Lemann, *The Big Test* (New York: Farrar, Straus and Giroux, 2000).

³⁶⁴ David Owen, *None of the Myth: Behind the Myth of Scholastic Aptitude* (New York: Houghton Mifflin, 1985); Adam Robinson, John Katzman, and David Owen (foreword), *Cracking the System: The SAT* (New York: Villard Books, 1986); Adam Robinson and John Katzman, eds., *College Admissions: Getting into the College of Your Choice* (New York: Princeton Review, 1993).

³⁶⁵ ³⁶⁵ Some institutional histories from standardized test manufacturers include: Michael C. Johanek, ed., *A Faithful Mirror: Reflections on the College Board and Education in America* (New York: College Board, 2001); Manuel Maldonado-Rivera, *The History of the Puerto Rico and Latin American Office*, edited by Michael Farah,

A more stable historiographical foundation for this topic can be built on the history of consumerism and consumer advocacy—particularly postwar U.S. consumer action. Although Lizabeth Cohen and Lawrence Glickman disagree about the relationship between consumers, collective action, and public space in post-1945 America, both authors’ respective work show how the role of consumer has often also been wielded as a political position, to varying success.³⁶⁶ As Matthew Hilton has argued, some hesitance may exist in classifying consumers as political actors because postwar consumer activism—domestic, international, or transnational—has not wedded itself cleanly to left/right political dynamics; consumers haven’t articulated a politics as definably partisan, or at least as easily outlined, as those espoused by postwar activists and social movement actors that historians and social scientists tend to cover.³⁶⁷ Postwar consumers continually used their position to exert political influence (albeit with periods of abeyance), and in the process defined public space and political action through the ability to consume or not.

During the 1960s and 1970s, a growing number of Americans began to express their social and political discontent from the position of wronged consumers. Consumption and

<http://about.collegeboard.org/region-offices/puerto-rico>. See also: John A. Valentine, *the College Board and the School Curriculum: A History of the College Board’s Influence on the Substance and Standards of American Education, 1900-1980* (New York: College Entrance Examination Board, 1987); and Frank Bowles, *The Refounding of the College Board, 1948-1963: An Informal Commentary and Selected Papers* (New York: College Entrance Examination Board, 1967).

Rebecca Zwick’s research epitomizes test makers’ artful dodge in considering their role in fueling social disparities; see: Rebecca Zwick, “Is The SAT a ‘Wealth Test’?” *Phi Delta Kappan* 84, no. 4 (December, 2002): 310. For a deeper look at Zwick’s underlying point here—as well as the reasoning behind her objection to doing away with standardized testing—see: Rebecca Zwick, “Eliminating Standardized Tests in College Admissions,” *Phi Delta Kappan* 81, no. 4 (December, 1999): 320-324. Zwick is a former researcher at Educational Testing Service.

³⁶⁶ Lizabeth Cohen, *A Consumers’ Republic: The Politics of Mass Consumption in Postwar America* (New York: Alfred A. Knopf, 2003); Lawrence Glickman, *Buying Power: A History of Consumer Activism in America* (Chicago: University of Chicago Press, 2009).

³⁶⁷ Matthew Hilton, “Social Activism in an Age of Consumption: The Organized Consumer Movement,” *Social History* 32, no. 2 (May, 2007): 121-143.

consumerism had long served as bases for Americans to frame their identity and air their grievances—from the rejection of chain stores during the Great Depression to the tactical boycotts of the Civil Rights Movement—but the consumer activism of the late 1960s and 1970s reflected a growing dissatisfaction with individually-felt repercussions of postwar Keynesian economics.³⁶⁸ The booming postwar economy may have created unprecedented wealth and marketplace options for the white middle class, but it also agitated many who felt they had been conned into wasting hard-earned money on expensive defective goods. Stuck participating in an economy premised on consistent consumption, in which their social status hinged on conspicuously embracing trends, but lacking any safeguards against bad-faith business practices, many consumers developed a sense of powerlessness and cultural estrangement. Such consumers, reasoned business researchers at the time, also believed the marketplace lacked norms or genuine meaning.³⁶⁹ Airing their grievances as consumers allowed many in the white liberal middle class to press for changes that would materially benefit their own status without either upending the market system or seeming like an imposition of values: What American doesn't want quality goods? Although this wave of consumerism had one foot in the politics of environmentalism and social justice—most notably, the sustained boycott of table grapes from

³⁶⁸ American consumer activism goes even further back, and as T.H. Breen argued, was central to the political identity formation and protests that fueled the American Revolution. Yet, as Lawrence Glickman notes, both American politics and business generationally forgot the very idea of consumers, treating each surge in consumer activism as a phenomenon without precedent. See: T.H. Breen, “‘Baubles of Britain’: The American and Consumer Revolutions of the Eighteenth Century,” *Past & Present* 119 (May, 1988): 73-104; Lawrence B. Glickman, “The Strike in the Temple of Consumption: Consumer Activism and Twentieth-Century American Political Culture,” *Journal of American History* 88, no. 1 (June, 2001): 99-128. See also: Robert E. Weems, Jr., *Desegregating the Dollar: African American Consumerism in the Twentieth Century* (New York: NYU Press, 1998); Daniel Scroop, “The Anti-Chain Store Movement and the Politics of Consumption,” *American Quarterly* 60, no. 4 (December 2008): 925-949; Meg Jacobs, “State of the Field: the Politics of Consumption,” *Reviews in American History* 39, no. 3 (September, 2011): 561-573.

³⁶⁹ Zarrel V. Lambert and Fred W. Kniffin, “Consumer Discontent: A Social Perspective,” *California Management Review* 18, no. 1 (Fall 1975): 36-44. Lambert and Kniffin built off of the concept of alienation theorized by sociologist Melvin Seeman, who posited that the phenomenon could be categorized by group feelings of powerlessness, meaninglessness, normlessness, isolation, and self-estrangement; see: Melvin Seeman, “On the Meaning of Alienation,” *American Sociological Review* 24, no. 6 (December, 1959): 783-791.

the Coachella Valley—such consumer activism was, in many ways, a way to signify social status and class values.³⁷⁰ Forgoing Di Giorgio grapes and avoiding appliances with low *Consumer Reports* ratings allowed in the white middle class to externally portray their self-perception as savvy, rational actors who simply wanted fair conditions in the marketplace, without tearing down the market altogether.

No figure epitomized this era of consumer activism—or had as much influence on how college students would articulate their grievances from the position of citizen-consumers—as much as the consumer advocate Ralph Nader. Nader relied heavily on the efforts of college students for his investigation of the Federal Trade Commission in the late 1960s. Nader believed college students could harness their advanced training to improve their conditions as both citizens and consumers through direct investigation, advocacy, action, and lobbying. This conviction inspired Nader to develop the initial framework for Public Interest Research Groups (PIRGs) in the early 1970s. PIRGs researched, lobbied, and litigated issues on behalf of students, using funds culled from tuition fees. PIRGs also took advantage of American education’s decentralized form, operating as state-level corporations that held offices at various college and university campuses. Although PIRGs ostensibly represented the various political interests of

³⁷⁰ Robert N. Mayer, “The Socially Conscious Consumer—Another Look at the Data,” *Journal of Consumer Research* 3, no. 2 (September, 1976): 113-115; David Vogel and Mark Nadel, “Who is a Consumer? An Analysis of the Politics of Consumer Conflict,” *American Politics Quarterly* 5, no. 1 (January, 1977): 27-56; Matthew Hilton, “The Death of a Consumer Society,” *Transactions of the Royal Historical Society* ser. 6, vol. 18 (2008): 211-236.

The politics of the California table grape boycott—in which farm workers appealed to consumer tactics not only to gain leverage against exploitative growers as well as Teamsters, all while dealing with internal tensions with representative leadership—is largely beyond this discussion except to note that consumers’ participation in a boycott hinged on a complex set of issue-framing related to marketplace practices, social justice, and self-perception in relation to commonly-ascribed liberal middle-class values; see: Sankar Sen, Zeynep Gürhan-Canli and Vicki Morwitz, “Withholding Consumption: A Social Dilemma Perspective on Consumer Boycotts,” *Journal of Consumer Research* 28, no. 3 (December, 2001): 399-417; Matthew Garcia, “Labor, Migration, and Social Justice in the Age of the Grape Boycott,” *Gastronomica* 7, no. 3 (Summer, 2007): 68-74; Elizabeth Lamoree, “Gambling on Grapes: Management, Marketing, and Labor in California Agribusiness,” *Agricultural History* 86, no. 3 (Summer, 2012): 104-127.

college students, a few critics believed the fundraising schemes that kept most PIRGs afloat were themselves acts of consumer fraud. While some states let college students voluntarily pay PIRG-earmarked tuition fees, others mandated that students must explicitly request their college stop assessing PIRG fees with their tuition, and others still gave students to ability to choose one way or the other. (As Nader’s friend-turned-enemy David Sanford noted, while the opt-out model of collecting fees made it possible for certain PIRGs to generate much larger revenues than counterparts with opt-in models, this form of “negative option” selling was more disingenuous than the book-of-the-month clubs Nader had already lambasted for using similar approaches.)³⁷¹ This varied state-by-state fee system left some PIRGs flush with funds, full-time attorneys, and staff, and others with less than \$50,000 in annual funds by the mid-1970s.³⁷² Whatever their funds, however, all PIRGs reported that most of their staff were volunteers—and whether or not their efforts reflected the politics of every student who contributed funds, PIRGs offered a venue for students to invoke their democratic rights and responsibilities as active, dissatisfied consumers.³⁷³ Nader’s PIRGs created a concentration of money, voluntary and paid labor, political agitation, and shared democratic-capitalist vision that enabled New York’s PIRG to take on the standardized testing industry in the latter half of the 1970s.

³⁷¹ David Sanford, “Funding PIRGs: Nader’s Vested Interest,” *The New Republic* June 12, 1976: 11-12. Sanford also questioned how much independence PIRGs actually had from Nader’s agenda, particularly as “he did invent them, he goes to campuses to help organize them, he attacks administrators who resist them, and when otherwise thwarted he goes to Congress for help in putting down the opposition” (11). See also: “Nibbling at the Nader Myth,” *Time*, September 6, 1976: 24.

³⁷² Ronald P. Matross, “A Profile of Public Interest Research Groups,” *University of Minnesota Office of Student Affairs Research Bulletin* 17, no. 9 (February 10, 1977): 10-13; Ralph Nader, “In the Public Interest: Student Activists,” *the New Republic*, February 19, 1972: 10- 11.

³⁷³ Richard J. Leighton, “Consumer Protection Agency Proposals: The Origin of the Species,” *Administrative Law Review* 25, no. 3 (Summer 1973) 269-312; Ralph Nader, “Student Power 101,” *Change* 11, no. 8 (November-December 1979): 47-50; Patricia Powers, “Social Change: Nader Style,” *Journal of Education and Social Work* 13, no. 3 (Fall 1977): 63-69; Al Senia, “The New Student Activists,” *Change* 6, no 8 (October, 1974): 29-33; Arthur Levine and Keith R. Wilson, “Student Activism in the 1970s: Transformation Not Decline,” *Higher Education* 8, no. 6 (November, 1979): 627-640; and Carol Foreman, “Protecting the Consumer,” *Challenge* 20, no. 4 (September/October 1977): 24-28.

Central as Nader's PIRGs would eventually be to standardized-testing reforms, students developed other approaches during this period for challenging educational practices as wronged consumers. Notably, some students sought redress for their consumer dissatisfaction through the courts. These students who filed suit against educational institutions and school districts claimed they were victims of educational malpractice. These attempts to claim educational malpractice were buoyed by changes to the American legal landscape—most notably, the end of *in loco parentis* and the 26th Amendment—that made it much clearer for young adults to make legal claims on their own behalf.³⁷⁴ In such cases—namely, *Peter W. v. San Francisco Unified School District* and *Donahue v. Copiague Union Free School District*—students sued school districts after completing high school, claiming profound educational deficiencies that rendered them unemployable. Although both *Peter W.* and *Donahue* ultimately proved unsuccessful for the plaintiffs, they did establish a precedent that students and parents could at least imagine education as a consumable good whose practitioners, as in medicine, remained liable in instances of willful neglect.³⁷⁵ By extension, these lawsuits posed the student litigants as defective goods. An education was thus not necessarily the product in these lawsuits, but the process by which the students became valuable commodities in the marketplace. Although the Naderist approach would win out in changing educational practices (at least, in the short term), the elision between students as consumers of education and students as commodities produced by educational institutions endured.

³⁷⁴ Robert A Laudiciana and Joseph L Tramultola, *A Legal Overview of the New Student: As Educational Consumer Citizen, and Bargainer* (Springfield, IL: Charles C. Thomas, 1976)

³⁷⁵ Ethan Hutt and Aaron Tang, "The New Education Malpractice Litigation," *Virginia Law Review* 99, no. 3 (May, 2013): 428-440; *Peter W. v. S.F. Unified Sch. Dist.*, 131 Cal. Rptr. (Ct. App. 1976); with *Donohue v. Copiague Union Free Sch. Dist.*, 47 N.Y.2d 440 (1979). Hutt and Tang indicate the idea of "educational malpractice" had swirled in legal thinkpieces since the beginning of the 1970s; see: Gary Saretsky and James Mechlenburger, "See You in Court?" *Saturday Review* October 14, 1972: 50, 55; and Stephen D. Sugarman, "Accountability through the Courts," *The School Review* 82, no. 2 (February, 1974): 233-259.

Educational malpractice also has a root in a series of court cases between Marjorie Webster Junior College and the Middle States Association of Colleges and Secondary Schools (MSACSS) during the end of the 1960s and early 1970s. MSACSS refused to grant accreditation to Marjorie Webster, which was one of about two-dozen degree-granting for-profit higher education institutions in the United States at the time. MSACSS claimed that for-profit institutions fell outside of the purview of accreditation agencies. Marjorie Webster filed suit, claiming that MSACSS's refusal to consider accreditation detrimentally harmed their ability to attract new students, despite being in operation for nearly fifty years. The District Court of Washington, DC sided with Marjorie Webster in 1969, declaring that MSCASS's refusal to consider accrediting the school solely on the basis of its proprietary status was an unreasonable restriction of trade. The district court judge ruled that MSCASS, through its accreditation process, "engaged in a quasi-governmental function," thus "subjecting it to the restraints of the Constitution."³⁷⁶ Even if education was neither an explicit constitutional concern nor primarily a commercial realm, the restriction of accreditation amounted to a violation of Fifth Amendment due process rights in addition to Sherman Antitrust principles. This decision was overturned the following year. The DC Circuit Court of Appeals ruled that Sherman Antitrust regulations did not apply to the nation's six regional accreditation corporations, no matter how vital the quasi-governmental function they served, because protections against "incidental restraint of trade, absent an intent or purpose to affect the commercial aspects of the profession" was not transferrable to "the noncommercial aspects of the liberal arts and the learned professions."³⁷⁷

That is, the primary business of education was education and not business. The Supreme Court

³⁷⁶ *Marjorie Webster Jr. Col., Inc. v. Middle States Association of Colleges and Secondary Schools, Inc.*, 302 F. Supp. 459 (D.D.C. 1969), at 478.

³⁷⁷ *Marjorie Webster Junior College v. Middle States Association of Colleges and Secondary Schools* 432 F.2d 650 (D.C. Cir. 1970).

refused to grant certiorari, thus allowing the appellate court's decision stay. While the courts were thus not opposed to proprietary schools per se, they ultimately decided that the corporations that performed quasi-state functions regulating national education had a primary responsibility in forging standards and not maintaining schools' fiscal solvency. This primary responsibility of quasi-governmental educational corporations would later prove useful—but not entirely successful—for Educational Testing Service in its defense of certain testing practices.³⁷⁸

As the courts refereed the boundaries between education and business, the legislative and executive branches became embroiled in a several-year squabble over whether consumers required separate federal-level representation. Although Congress had created consumer protection measures since the 1920s, it pushed for more substantive bureaucracy during the late 1960s and early 1970s. What undermined this congressional effort, however, was a switch from legislation proposing a cabinet-level department for consumers to bills proposing an independent Consumer Protection Agency. This switch—largely sparked by Ralph Nader's testimony advocating for a more flexible and supervisory body than a department could provide—also inadvertently opened up competing consumer protection agency legislation as well as emboldened a minority of congressional members to filibuster to lengths that weakened support for popular bills.³⁷⁹ The executive branch also resisted approving such measures for a separate

³⁷⁸ 400 U.S. 965 (1971); S.H. , "the Case of Marjorie Webster," *Change in Higher Education* 1, no. 6 (November-December, 1969): 15-16; Stuart A. Sandow, "Emerging Education Policy Issues in Law: Fraud" (Syracuse: Educational Policy Research Center, 1970), <http://files.eric.ed.gov/fulltext/ED062910.pdf> (last accessed March 10, 2017); "Antitrust Laws: College Accrediting Association's Refusal to Evaluate Proprietary College Not a Sherman Act Violation," *Harvard Law Review* 84, no. 8 (June, 1971): 1912-1920; and Mike Cardozo, "P.S. on Marjorie Webster," *American Bar Association Journal* 57 (October, 1971): 1051-1052.

³⁷⁹ Richard J. Leighton, "Consumer Protection Agency Proposals: The Origin of the Species," *Administrative Law Review* 25, no. 3 (Summer 1973): 269-312; Richard J. Leighton, "The Consumer Advocacy Agency Proposal... Again," *Administrative Law Review* 27, no. 2 (Spring 1975): 149-164. For an example of one such debate (among several) about the creation of a Consumer Protection Agency, see: U.S. Senate, 92nd Congress, 1st Session, *Hearings Before the Subcommittee on Executive Reorganization and Government Research of the Committee on Government Operations* (Washington, DC: US Government Printing Office, 1972) .

Consumer Protection Agency, categorizing the effort as an exercise in government overreach. A government that too much influence on ensuring consumer satisfaction, Richard Nixon and Gerald Ford maintained, would be an overreach of government into business affairs. Making consumer advocacy a specific government responsibility would place government in direct opposition to the interests of business—and a marketplace told what to do by the state could not naturally develop the best possible solutions for its own consumer problems. Regulations and their resultant bureaucracy unduly burdened business leaders from being nimble enough to fix new problems as they emerged. A centralized consumer agency, by extension, would rob individual departments of the ability to manage relations with citizen-consumers in the scenarios most frequently encountered.³⁸⁰ The Ford administration in particular maintained that the best solution was requiring all cabinet-level departments and agencies to develop consumer representation plans so they could “determine where additional consumer input might be desirable.”³⁸¹ From a conservative perspective, “consumer rights” could easily be used to cudgel business and government without ever representing what actual consumer interests.

Although the Ford Administration’s bureaucratically decentralized approach to consumer affairs won out, this effort also inadvertently synchronized various federal approaches toward

³⁸⁰ Ford’s objections to a Consumer Protection Agency can be gleaned from the internal memos circulated in response to the Consumer Protection Act of 1975 (S. 200); see: Memo from Jim Cannon to Pam Needham, memo re: “Updated Analysis of S. 200,” May 5, 1975, Dawn Bennett-Alexander Papers, Box 2, Gerald Ford Library, which indicated an CPA would “short- circuit the normal budget review process” and “be detrimental to the free flow of dialogue and issue development within the executive branch” (1), as well as obscure what consumers, exactly, the agency advocated for.

³⁸¹ Office of the White House Press Secretary, message to Congress re: rejection of S. 200 (or H.R. 7575, November 1975; memo from Virginia Knauer to William Baroody and James Lynn, re: “Ford’s Administration Consumer Representation Plan,” May 2, 1975; *Federal Register* 40 no. 229 (November 26, 1975): 55092-55093; all in Dawn Bennett-Alexander Papers, Box 3. Files in this box also detail the laborious process getting all departments and agencies to actually turn in proposals in a timely manner, let alone create proposals that passed muster. More bluntly, Ford remarked in a speech to the National Chamber of Commerce critical of consumer protection proposals that “a Government big enough to give us everything we want is a Government big enough to take from us everything we have”; see: Office of the White House Press Secretary, “Remarks of the President to the 63rd Annual Meeting of the Chamber of Commerce of the United States,” April 28, 1975: 9; Dawn Bennett-Alexander Papers, Box 6.

consumers of education. Because no single governing body had sole control over American education, any effort to maintain a distilled approach to educational oversight had to, ironically, involve considerable interagency coordination. Beyond the Office of Education, the federal bodies alone responsible for funding and maintaining postsecondary schooling included the Social Security Administration, Department of Defense, Department of Housing and Urban Development, Bureau of Indian Affairs, Department of Labor Immigration and Naturalization Service, Federal Aviation Administration, and Veterans Administration. The Federal Interagency Committee on Education (FICE) attempted to forge a standardized federal approach to educational consumers, publishing its findings and recommendations in 1975. FICE aimed to create consensus regarding who was a consumer of education, which rights they held distinct from other types of consumers, and what role government could play in addressing grievances. The subcommittee maintained that students held special consumer rights due to “the expensive and intangible nature of the services [they are] purchasing, and in light of the potential for consumer abuse in ‘future service contracts’ used by most schools.”³⁸² As such, students needed specific consumer protection mechanisms—such as tuition insurance in instances of school bankruptcies, or loan forgiveness for fraudulent training programs—particularly when the federal government had already funded both students and the postsecondary institutions they attended. In the realm of education, a “buyer beware” approach would not only waste government funds but also diminish students’ eventual usefulness to the state. The subcommittee nonetheless insisted that protecting consumers of education had to involve public-private partnerships and all levels of governance. The federal government could not overtake the accrediting and licensing

³⁸² Federal Interagency Committee on Education, *Toward a Federal Strategy for Protection of the Consumer of Education* (Washington, D.C.: Department of Health, Education, and Welfare, 1975): 3-4; Dawn Bennett-Alexander Papers, Box 6.

responsibilities of existing private organizations—but this effort to share responsibility for educational consumers using existing systems still left students without a clear or direct path to file complaints against institutional misdeeds.

By the mid-1970s, then, the federal government was more primed than ever to treat students as a special class of consumer, but still lacked clear avenues for those complaints to be addressed. Students had attempted to build legal precedent for educational malpractice, but courts ultimately decided not to extent antitrust legislation or malpractice concepts to cover educational consumption. Consumers of education were instead treated as a special class of consumer, who were the shared responsibility of public institutions, state and federal governments, and private accreditation organizations. Where students interests became more directly represented were through Public Interest Research Groups—but these Naderite organizations often relied on convoluted tuition fees to fund actions on citizen-consumers' behalf.

Ralph Nader and Public Distrust of Educational Testing Service

As the broader legal and federal system reconsidered its relationship toward education and consumption, the American public grew increasingly critical of standardized tests—particularly the SAT. In particular, many parties obsessed over the long, slow decline of average SAT scores. Between the mid-1960s and the end of the 1970s, the mean SAT score for American high school students dropped by 90 points, sparking a angst-riddled hunt for the cause of this declining national benchmark. The College Board launched its own series of investigations, and concluded that the seventeen-year score decline reflected two consecutive and related phenomena. During the 1960s, as high schools absorbed the first waves of the Baby Boom, an unprecedented number of young adults completed their high school education rather than drop

out beforehand. The organization argued that the sheer number of college hopefuls, particularly those from educationally underserved socioeconomic communities, gradually dragged down the overall average performance among high school seniors year after year. By the 1970s, as Baby Boomers continued to stretch the limits of available educational resources, many school districts reduced the rigor of their high school programs in order to manage large numbers of students without risking high rates of attrition. Declining SAT scores, argued the College Board, reflected what inflated grades, dumbed-down textbooks, and softer course requirements tried to hide: a growing number of American students who wanted to attend college were poorly prepared for its academic challenges. This argument followed from the College Board's position on test coaching. Because the College Board insisted time and again that its tests were not coachable—and that SAT scores were the reflection of the skills and abilities developed throughout a student's years in school—any massive downward trend in test scores had to be reflective of shortcomings in American education at large. The test was “an unchanging measurement” in the College Board's eyes, and any public dissatisfaction with test due to low scores was misplaced frustration toward broader social problems.³⁸³

³⁸³ College Entrance Examination Board, *On Further Examination: Report of the Advisory Panel on the Scholastic Aptitude Test Score Decline* (New York: College Entrance Examination Board, 1977): 8; David G. Savage, “The Long Decline in SAT Scores,” *Educational Leadership* January 1978: 290-293, http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_197801_savage.pdf; William W. Turnbull, “Student Change, Program Change: Why the SAT Scores Kept Falling,” College Board Report No. 85-2 (New York: College Entrance Examination Board, 1985); Shepherd, 98-102, 133.

For the College Board's position on test-coaching, see: Samuel Messick et al, *The Effectiveness of Coaching for the SAT: Review and Reanalysis of Research from the Fifties to the FTC* (Princeton: Educational Testing Service, 1980). <https://www.ets.org/Media/Research/pdf/RR-80-08.pdf>; and Paul I. Jacobs, “Effects of Coaching on the College Board English Composition Test,” ETS Research Bulletin RB-64-24 (Princeton: Educational Testing Service, 1964). For a sense of how midcentury psychologists conceived of testwiseness, or the potentially-trainable skill of using the structure of a test to do well even when one might know the actual content being examined, see: Jason Millman, Carol H. Bishop, and Robert Ebel, “An Analysis of Test-Wisness,” *Educational and Psychological Measurement* 25, no. 3 (1965): 707-726,

Most midcentury criticisms of standardized testing were not rooted in a consumer mindset. Critics such as Banesh Hoffman instead claimed exams like the SAT were fraudulent because standardized testing was a pseudoscientific enterprise. Hoffman, a physicist who once collaborated with Albert Einstein, turned his attention to the flimsy logic of standardized tests in the 1960s. For Hoffman, the very structure of multiple choice testing—the format upon which objective standardized testing had been built decades earlier—was a phony attempt to seem scientific. By asking test-takers to select the best among presented answers, test-writers could dodge complaints that none of the options were exactly correct. Rather than identify the most capable students, Hoffman maintained, standardized tests punished mentally nimble youth. Clever students with enough awareness to sense what test-writers wanted still wasted time and energy discerning the designated “best” answer from other worthwhile choices. Straddling logical precision and compulsive pedantry, Hoffman dissected standardized testing questions used in several high-profile exams in *The Tyranny of Testing*, accusing test-writers of knowing nowhere near as much as they or the general public presumed they did. Hoffman reserved the most scorn for Educational Testing Service. If ETS was the best standardized testing had to offer yet it still produced a considerable amount of deeply flawed questions, Hoffman argued, then less capable testing companies were truly dangerous. The “testolatry” which gripped American society would only punish the best and brightest through false expertise.³⁸⁴

By the mid-1970s, however, a consumerist criticism of standardized testing had developed, fueled mainly the efforts of student volunteers at the largest of Nader’s Public Interest Research Groups, the New York PIRG. Nader regarded Educational Testing Service’s

³⁸⁴ Banesh Hoffman, *The Tyranny of Testing* (New York: Crowell-Collier Press, 1962), 217; Banesh Hoffman, “Psychometric Scientism,” *Phi Delta Kappan* 48, no. 8 (April, 1987): 381-386. For a contemporary criticism of Hoffman and his ilk, see: Kaoru Yamamoto, “Psychological Testing: Invasion of Privacy?” *Educational Leadership* February 1966: 363-368, http://ascd.com/ASCD/pdf/journals/ed_lead/el_196602_yamamoto.pdf

operation as secretive and unethical, and teamed with Princeton undergrad Allan Nairn to lead a several-year analysis of ETS's business practices. Nairn, Nader, and the New York PIRG publishing their findings in 1980 as *The Reign of ETS: The Corporation that Makes Up Minds*.³⁸⁵ Central to *The Reign of ETS* was how Nairn, Nader, and PIRG research assistants positioned themselves as dissatisfied consumers who aimed to make Educational Testing Service play by the established rules for consumer-corporation relationships. Nader set this tone in *The Reign of ETS's* preface, persuading readers to view the consumer-corporation bond between American students and Educational Testing Service as thoroughly rotten. Above all, Nader primed the reader to see ETS as a corporation that thrived on “its long-standing policy to remain hidden from public view.”³⁸⁶ Nader deemed this policy of corporate secrecy as “inimical to the opportunities—both analytic and normative—which a democracy should provide all its citizens from all classes, races, age groups, and genders,” thereby not only accusing ETS of being anti-democratic, but also staking as capacious a claim for citizenship as possible.³⁸⁷ Citizens were thus doubly burdened by ETS, subject to corporate services that heightened social inequities as well as anti-democratic practices that undermined the premises of their citizenship. Nader asserted that educational contexts alone could not fully explain why Educational Testing Service was so deeply entrenched in American society. Rather, “the consumer perspective [was] needed to examine the assumptions and consequences of contemporary standardized testing.”³⁸⁸ To this

³⁸⁵ Lemann, *The Big Test*, 221- 227.

³⁸⁶ Ralph Nader, Preface, in Allan Nairn and associates, *The Reign of ETS: The Corporation that Makes Up Minds* (New York: Ralph Nader, 1980), xiv.

³⁸⁷ *Ibid.*

³⁸⁸ Nader, *Reign of ETS*, xvi.

end, Nader linked opportunity and democracy to the frame of consumerism, and established consumers as the most reliable actors for social analysis and institutional critique.

To built upon Nader’s introductory claims, Nairn and his associates posed a market-based relationship between ETS and the public that differed from the nonprofit’s self-portrayal. This first depended on redefining who, exactly, consumed standardized tests. Whereas ETS believed that its main consumer base consisted of the colleges and universities that used its testing data, and that students were only testing candidates, Nairn argued that the true consumers were the students or parents who purchased these tests. By redefining the consumer, Nairn also placed more emphasis on the act of test-taking (that is, how students used standardized tests) than score analysis (how schools used such tests). Redefining the primary consumer also, in effect, changed the actual commodity under exchange. For ETS, test data had value in its predictive validity and its reliability, or, how closely test scores correlated with a subsequent metric—almost always first-year grades—and how consistently any given test-taker would earn similar scores were they to take multiple administrations of the same test. As Nairn suggested through his use of numerous vignettes in which aspiring students were victimized by ETS practices, the actual commodity was the sense of self rendered through the test-taking process. Just as a great test-taking experience could validate a student’s self-worth, a bad testing experience, even before a student ever received a score, could generate feelings of worthlessness: I am useless because I am stupid because my test scores are too low to go to college.³⁸⁹

To this end, Nairn’s redefinition of the consumer relationship in standardized testing also reoriented consumers as victims of a rigged system. Nairn posed test-taking as “involuntary consumption”: even though students ultimately had the choice to not take standardized

³⁸⁹ Nairn, *Reign of ETS*, chs 1-2.

admissions exams, such a choice would come to the serious detriment of their career plans.³⁹⁰ Even then, students who took the test could have their plans derailed by an underwhelming score. Students, as consumers of tests, were caught in a “contract of adhesion” without any clear understanding of their contractual rights.³⁹¹ Suspicious scores could be canceled by ETS without any recourse for affected test-takers—unless a student had limitless resources to go against ETS’s team of lawyers. Test-takers were not only victimized by psychometric logic that cast discontents as “those who simply have a painful time accepting their proper role in life,” but also by the well-funded and legally-protected means ETS used to keep psychometric systems of educational measurement in place.³⁹²

Underlying all of these claims was Nairn’s argument that Educational Testing Service’s tests did not measure students’ aptitude but, instead, reflected their class position. After illustrating that the SAT’s predictive validity was quite weak, even when combined with other predictors (such as a student’s high school record), Nairn used SAT score data to argue that only one clear correlation could be drawn with testing data: between a student’s test scores and their parents’ income. Regardless of region or ethnic group, the higher the score range, the wealthier the average household income. Nairn insisted this “score-income correlation” was “more a reflection of socio-economic status than their actual potential for future accomplishment.”³⁹³ By framing test scores as a reflection of class, Nairn in effect narrowed the range of critique; ethnic

³⁹⁰ Nairn, *Reign of ETS*, 261.

³⁹¹ Nairn, *Reign of ETS*, 265. According to Cornell Law’s Legal Information Institute Wex Legal Dictionary, in a contract of adhesion, the weaker party, typically a consumer, “must adhere to the contract and therefore does not have the power to negotiate or modify the terms of the contract”; see: “Adhesion Contract (Contract of Adhesion), Legal Information Institute Wex Dictionary, https://www.law.cornell.edu/wex/adhesion_contract_contract_of_adhesion (last accessed March 14, 2017). This contractual relationship is usually found in large purchases reliant on credit or insurance.

³⁹² Nairn, *Reign of ETS*, 23.

³⁹³ Nairn, *Reign of ETS*, 210.

and cultural biases were downplayed for the economic conditions that allowed students to have the specific type of education rewarded by ETS's tests, a type of education historically reserved for white men privately-educated in the Northeast.

This score-income correlation also blurred who, exactly, was given the greater grounds to complain through their position as consumers: poor families, who often received the lowest scores but may have other identities through which to stake their testing grievances; or middle-class families, whose relatively average scores generated a personal sense of mediocrity at odds with prevailing social narratives for middle-class success. As Nairn maintained, only the wealthy benefitted from ETS's model of merit and aptitude, a model that did not match up with an ETS study that found "actual accomplishments outside the classroom did not correlate with income either."³⁹⁴ By individualizing the rhetoric of test-taking—as a measurement of how one performed against all test-takers, devoid of demographic contexts—ETS could continue to commodify the self-assessments students generated through the test-taking process, all while evading claims of systemic bias.

Nairn maintained that ETS was able to have such tremendous influence because its rarely monitored nonprofit status granted it legal protection from consumer-focused scrutiny and complaint. Nairn particularly emphasized ETS's long-overdue paperwork for nonprofit status in New York State, where the organization was based. As a non-profit, ETS operated "exempt from the scrutiny of the [Federal Trade Commission]"—and as an educational enterprise, it benefitted from weak federal oversight.³⁹⁵ (Although the Department of Health, Education, and Welfare would be split into the Department of Health and Human Services and Department of Education

³⁹⁴ Nairn, *Reign of ETS*, 206.

³⁹⁵ Nairn, *Reign of ETS*, 280.

by the end of the 1970s, this separate Cabinet-level upgrade didn't necessarily come with tremendous additional federal-level controls over educational practices.) Nairn argued that ETS was, for all purposes, a monopoly; other testing companies did exist, but they operated at nowhere near as large a scope or scale as ETS did. Because so many institutions and organizations—governmental, educational, and otherwise—relied on ETS-produced tests for hiring, admissions, and classification, the nonprofit existed in a realm beyond legal scrutiny. For Nairn, the scenario was bureaucratic absurdity: a virtual monopoly that evaded trustbusters, providing educational products outside the reach of federal government, with its revenues and operations protected by its nonprofit business designation.

The solution, as Nairn concluded, was to make ETS conform to the rules of business through transparency legislation—an effort also undertaken by the New York PIRG. Nairn detailed New York PIRG and State Senator Kenneth P. LaValle's three-year effort to enact "truth-in-testing" legislation in New York State.³⁹⁶ Signed into law of July 1979 and effective the beginning of the following year, New York State's Educational Testing Act of 1979 positioned test-takers as student-consumers entitled to special protections. Primarily, the law mandated testing companies "give students specific information on what their scores meant and how their scores would be presented to academic and other institutions" as well as "provide students—on request—with a copy of the questions, correct answers and the student's own answers thirty days after the scores had been received."³⁹⁷ By requiring ETS and other test-makers to adhere to certain regulations similar to other corporations, such truth-in-testing

³⁹⁶ Nairn, *Reign of ETS*, 370.

³⁹⁷ Nairn, *Reign of ETS*, 371; see also Appendix V (pp. 530-547) for the actual language of subsequent editions of truth-in-testing legislation. See also: Shawn G. Kennedy, "Albany Bill Challenging Secrecy in Testing," *New York Times*, May 27, 1979.

legislation allowed students and parents to observe and challenge psychometric principles from the position of dissatisfied customers.

Although truth-in-testing became law in New York State at the turn of the 1980s, this victory not without pushback from test-makers. Claiming the law would undermine test security and create financial ruin through skyrocketing research and development costs, the Association of American Medical Colleges and American Dental Association threatened to stop providing their respective admissions exams in New York State. The AAMC later secured an injunction to prevent MCAT administrations in New York.³⁹⁸ The Psychological Corporation followed suit, pulling the Miller Analogies Test from the state. But despite test-makers' full-throated opposition to statewide and national transparency legislation, ETS and the College Board—by far the two largest bodies in standardized testing—reversed their opposition to New York's Law days before it took effect, deciding to follow disclosure procedures in exchange for significantly fewer test administration dates.³⁹⁹ Within a year, the College Board made disclosure a nationwide policy, following in the steps of the Law School Admissions Council and other bodies that oversaw admissions testing.⁴⁰⁰

³⁹⁸ Edward B. Fiske, "Colleges May Lose 20 Entrance Exams: New York 'Truth' Law Makes Cost Prohibitive, Publishers Say," *New York Times*, October 8, 1979; Ralph Nader, "Foes of 'Truth in Testing'" (letter to the editor), *New York Times* August 3, 1979; Dena Kleiman "Medical Colleges Win Injunction Against Testing Law," *New York Times* January 22, 1980; Dan Hulbert, "Q. & A. on 'Truth in Testing,'" *New York Times*, January 6, 1980

³⁹⁹ Edward B. Fiske, "College Entrance 'Truth in Testing' Effort Assailed," *New York Times*, September 11, 1979; Edward B. Fiske, "Student Questions and Answers on 'Truth in Testing' Provisions," *New York Times* October 10, 1979; "S.A.T. Schedule Will Be Reduced Beginning Jan. 1," *New York Times*, October 10, 1979; Irvin Molotsky, "Testing Law Opposed in Congress By Makers of Exams: Sees New York as 'Hostage,'" *New York Times*, October 28, 1979; Edward B. Fiske, "5 Exam Groups Ease Opposition To Law on Tests," *New York Times* December 30, 1979.

⁴⁰⁰ Dena Kleiman, "'Truth in Testing Law' Faces Legality Test," *New York Times*, November 9, 1979; Edward B. Fiske, "'Truth in Testing' to Be Nationwide," *New York Times* March 27, 1981.

Other figures with a vested interest in college admissions remained critical of truth-in-testing after disclosure legislation took effect in New York. Educational historian-turned-policy wonk Diane Ravitch lambasted New York’s truth-in-testing law as “consumer fraud” in itself.⁴⁰¹ Ravitch accused transparency advocates of having little sense of standardized testing’s expenses, thereby pushing for a situation that would only harm consumers in the long run. Testing companies would be forced to raise the price of entrance exams considerably in order to recoup questions lost by disclosure policies, effectively reducing poor students’ ability to even take standardized admissions exams.

Stanford Dean of Admissions Fred Hargadon was particularly critical of truth-in-testing legislation, and went so far as to suggest that the push for transparency was emblematic of middle-class moral bankruptcy. As quoted in the *New York Times*, Hargadon accused transparency advocates of shielding behind the cries of the underprivileged to cloak their concerns over the composition of the middle class:

“It’s one thing to make it on your own on merit. [...] ‘It’s another, as parents, to have the same measures applied to your children. A lot of self-made people aren’t going to hold their children to the same standards they applied to themselves.’”⁴⁰²

New York Board of Regents member Kenneth B. Clark echoed these sentiments. (The Board of Regents was responsible for granting and overseeing ETS’s status as a nonprofit educational institution.) Clark opined that the law’s provision for students to request and review testing materials was a “deceptive and meaningless exercise” that would only reveal the obvious: that test scores reflected class-based disparities in educational opportunity, disparities that

⁴⁰¹ Diane Ravitch, “The Law on Tests Is Bad,” *New York Times* 15 August, 1979.

⁴⁰² “Fred M. Hechinger, “Aptitude Tests Stir a Debate—But is it Moot?” *New York Times*, 4 March 1980.

privileged students could broaden through their ability to access additional opportunities.⁴⁰³ By Clark's assessment, truth-in-testing was a well intentioned but misguided "placebo," one that erroneously aimed to solve the social inequities reflected in test scores by offering procedures students could only follow after the fact. The new protocols offered little more than a legally sanctioned way to vent frustrations rather than upend the system.⁴⁰⁴

One irony of the debate over testing is that, at the time, certain education writers suspected the fight over transparency signaled the end of standardized college testing altogether. *New York Times* education writer Fred Hechinger suspected the SATs would "almost certainly be an academic paper tiger within a few years."⁴⁰⁵ By Hechinger's estimation, colleges and universities had over-expanded to the point where they could not be too picky about who fills their classrooms, so long as their tuition checks cleared. Completing the SAT would become less of a hurdle to quality postsecondary education and more a minor nuisance, an examination with little value in itself except for those students who achieved statistically abnormal results. This "shift from selection to placement on the campus" would reinforce consumers' wishes rather than fulfill institutions' missions.⁴⁰⁶ Indeed, as ETS and the College Board endured several high-profile scoring gaffes exposed by students who took advantage of disclosure policies, the standardized testing industry seemed prone for obsolescence: What good would standardized

⁴⁰³ Kenneth B. Clark, "This So-Called Truth-in-Testing Law Is a Placebo" (letter to the editor), *New York Times*, August 18, 1979.

⁴⁰⁴ *Ibid.*

⁴⁰⁵ Hechinger, "Aptitude Tests Stir a Debate—But is it Moot?"

⁴⁰⁶ *Ibid.*

tests be if they were riddled with errors and when college had become a literal and metaphoric buyers' market?⁴⁰⁷

The Legitimation and Expansion of the Test-Prep Industry

As consumer advocates pushed for greater transparency, standardized test manufacturers also felt increasing pressure from standardized test preparation (“test-prep”) companies. Test-prep companies had offered courses and study aids for standardized college entrance exams since the beginning of the postwar period, but they were largely concentrated in the greater New York metropolitan area and New England. Nonetheless, the College Board devoted considerable time in the 1950s and 1960s commissioning studies that discredited the effectiveness of “test coaching,” claiming that the only effective form of test-prep was a thorough education. To this end, test-prep for college or postgraduate standardized test was ineffective because foundational skills had already been set years earlier: aptitude for college success had already been set far earlier in a student’s education.⁴⁰⁸

Despite the College Board’s insistence that coaching was a waste of time, test-prep companies benefitted from several shifts to the standardized testing and American educational practices during the late 1960s and early 1970s. Quite simply, far more young people were entering college than in the early postwar period, and colleges and universities increasingly turned to standardized admissions exams to make sense of their growing applicant pool. These colleges and universities—particularly those in the Southern and Midwestern parts of the United

⁴⁰⁷ For an early example, see: Edward B. Fiske, “Youth Outwits Merit Exam, Raising 240,000 Scores,” *New York Times*, March 17, 1981; Edward B. Fiske, “A Second Student Wins Challenge on Answer to Math Exam Problem,” *New York Times*, March 24, 1981; Edward B. Fiske, “Pyramids of Test Question 44 Open a Pandora’s Box,” *New York Times*, April 14, 1981; and Edward B. Fiske, “Soul-Searching in the Testing Establishment,” *New York Times* April 28, 1981.

⁴⁰⁸ See in particular, College Entrance Examination Board, *Effects on Coaching on Scholastic Aptitude Test Scores* (New York: College Entrance Examination Board, 1965), which compiled seven studies between 1953 and 1962 to discredit the effectiveness of short-term intensive test preparation.

States—increasingly turned to the American College Testing (ACT) Program as a standardized admissions exam. While this growing test market may have been stressful for the College Board and Educational Testing Service, it spelled security for the test-prep industry: more people taking more standardized tests, whatever the test, created a greater demand for college hopefuls to successfully clear the testing hurdle.⁴⁰⁹

Given the vast expansion of the college-going and college-hopeful populations during the previous two decades, as well as the spectacular claims that certain test-prep companies made about their ability to raise test scores, the Federal Trade Commission felt it necessary to probe the \$60-million industry to assess its ethicality and legality. Although plans to investigate the test-prep industry were first announced in late 1976, the commission's report would not be finished for another two and a-half years. The results, however, would directly alter the relationship between standardized test manufacturers, test-coaches, and the American public.⁴¹⁰

Although the FTC's Bureau of Consumer Protection sought to "estimate statistically the impact of commercial coaching on SAT scores," the Bureau explicitly stressed its investigation could not be categorized as experimental.⁴¹¹ The bureau avoided using a group of non-coached students as a control group. While using a control may have elicited more statistically and scientifically sound results, the Bureau stressed that in this instance, a control group would have been financially and ethically unsound. The Bureau felt it could not "[deny] access to

⁴⁰⁹ Lemann, *Big Test*, 102-108; Shepherd, 86-94; "College Testing Program Started," *Science* 130, no. 3387 (November 27, 1959): 1456; "New Testing Plans for Colleges Set," *New York Times*, September 9, 1959: L19; "College Test Plan is Rated a Success," *New York Times* 31 May, 1960: L17; Frank R. Peters and Eugenia L. Plog, "The Effectiveness of the ACT for Selection and Placement at the Ohio State University," *Educational Research Bulletin* 40, no. 9 (December 13, 1961): 232-241, 252; Edward B. Fiske, "Finding Fault with the Testers," *New York Times* November, 18, 1979: SM40.

⁴¹⁰ "Trade Commission to Investigate Coaching for College-Entry Tests," *New York Times*, November 3, 1976; Fred M. Hechinger, "About Education: Coaching Courses Studied by F.T.C.," *New York Times* 19 Dec., 1978.

⁴¹¹ Federal Trade Commission Bureau of Consumer Protection, "Effects of Coaching on Standardized Admission Examinations" (Washington, D.C.: Federal Trade Commission, 1979), 1.

commercial coaching to students who want[ed] it.”⁴¹² Should coaching have proven effective, parents may have been incensed over a policy that hobbled their children’s chances on the SAT.

At least one of the companies under investigation was reluctant to relinquish its clientele’s personal information, leading to arbitration in the United States District Court for Massachusetts. In *Federal Trade Commission v. Stanley H. Kaplan Educational Center, Ltd., et al*, the FTC petitioned the court to have Kaplan comply with its *subpoena duces tecum*.⁴¹³ Presiding judge D.J. Tauro deferred in part to the precedent set by *FCC v. Schreiber*, in which the Supreme Court decided that “courts should not substitute their judgment for that of regulatory agencies more familiar with the industries they are charged with regulating.”⁴¹⁴ The court would thus not honor Kaplan and others’ request to prevent the subpoena because the FTC had a better understanding of what would be necessary for a thorough and valid investigation.

Once obtaining the files, the bureau worked around its investigative shortcomings by designing ersatz experimental and control groups. The bureau corroborated customer information for the two schools under investigation against ETS and College Board files for all SAT candidates who resided in the same postal-code hubs as the majority of the coaching schools’ clientele.⁴¹⁵ The files for these coaching clients were then corroborated against ETS’s list of SAT candidates from the New York metropolitan area for all academic years between 1974 and 1977. The Bureau winnowed this pool of identified coaching clients students from 1,568 to roughly 1,000, discounting clients who never took the SAT, took it outside the New York City

⁴¹² FTC, “Effects of Coaching,” 3.

⁴¹³ *Federal Trade Commission v. Stanley H. Kaplan Educational Ltd., Et Al*, 433 F. Supp. 989 (1977), 990.

⁴¹⁴ *Federal Trade Commission v. Stanley H. Kaplan Educational Ltd., Et Al*, 433 F. Supp. 989 (1977), 993 .

⁴¹⁵ FTC, “Effects of Coaching,” 4. As noted in the report, the investigators began with three coaching schools, but soon dropped one of these programs from its investigation due to its overly small clientele pool.

metropolitan area, or took the ACT instead.⁴¹⁶ Finally, the Bureau forged a control group by randomly selected the SAT-score records for 2,500 un-coached students from ETS files.

Although the investigation could not allow for a genuine experimental group, test coaching was prevalent enough in the New York metropolitan area by the late 1970s to create large simulations of experimental and control populations.

The investigators also used a standardized demographic device to clarify any preexisting discrepancies between their coached and uncoached groups. The bureau compared coached and uncoached students' respective responses on the Student Demographic Questionnaire (SDQ), a voluntary survey administered during each SAT administration.⁴¹⁷ Investigators determined that students who attended coaching schools tended to have higher course marks and PSAT scores than their un-coached peers. Coached students were more often than not "A" students in their English and math courses, while un-coached students were more often than not "B" students in these same subjects. The bigger difference between coached and un-coached students, however, could be found in their family's income figures. Over two-fifths of the coached students came from families whose parental incomes exceeded \$30,000 per year, while nearly one-half of the un-coached students came from households with incomes less than \$18,000 per year. Even when accounting for self-selection factors that would encourage more youth from more affluent families to consider college (and thus be more likely to take the SAT), coached students were considerably more likely to come from uppermost socioeconomic brackets.⁴¹⁸

⁴¹⁶ FTC, "Effects of Coaching," 5.

⁴¹⁷ FTC, "Effects of Coaching," 6

⁴¹⁸ FTC, "Effects of Coaching," 8-11; "Historical Employment and Wages, 1975-2000," *New York State Department of Labor* website, <https://labor.ny.gov/stats/employ/county.shtm>.

Investigators employed multiple regression analyses in order to best gauge the genuine impact of coaching on test scores. This statistical technique was preferred to others because it allowed the “ability to analyze the impact of one variable on another variable while [... holding constant] the effects of several other factors.”⁴¹⁹ Given the sheer variance within the SAT candidacy in any given exam administration, the Bureau found it necessary determine the effect of test-prep on SAT scores by first controlling for gender, income, high school marks, class rank, high school type, PSAT scores, and distance between PSAT and SAT sittings.

The Bureau determined “that coaching was effective at one of the two schools”: Stanley H Kaplan Educational Centers produced, on average, a fifty-point increase in customers’ composite SAT scores.⁴²⁰ However, this assertion was somewhat misleading as the average rates of improvement attributable to coaching among the four different testing dates used in the study varied from 35.1 points to 71 points. When factoring in confidence intervals, the rate of improvement attributable to Kaplan’s coaching ranged from 31 to 67 points, and among individual test dates, confidence intervals ranged from 2 points to 96 points. The other coaching school under investigation, Test-Preparation Center, Inc., had drearier results. On average, students coached by Test-preparation Center had a mere 3.6-point improvement on their composite SAT scores, and their performances on the verbal section tended to slightly decline.⁴²¹

The difference in test-coaching results between the two companies could be attributed to markedly different operating practices. Brooklyn-based Stanley H Kaplan Educational Centers refused to operate without preponderant control over its material—and the namesake owner went so far as to engage in nepotistic behavior during the early expansion of his business in order to

⁴¹⁹ FTC, “Effects of Coaching,” 3.

⁴²⁰ FTC, “Effects of Coaching,” ii.

⁴²¹ FTC, “Effects of Coaching,” 18-22.

ensure that every office outside of New York City operated according to his standards. As seen in a 1975 ad placed in the *New York Times*, the Bronx-based Test-Preparation Center did not necessarily engage in such tight-knit management:

PART TIME [:] Hi profit biz oppty [sic] created by & for educators. Administer and educational program in your community. Contact TEST-PREPARATION CENTERS, Inc., 3701 Henry Hudson Pkwy, Suite 'D', Riverdale, N.Y. [:] 212-796-1076⁴²²

For a “low investment,” one could manage a Test-preparation Centers branch.⁴²³ The company also directly courted teachers in the classifieds section of the Times, offering math teachers fifteen dollars per hour for part-time after-school work on Wednesday afternoons.⁴²⁴ Although the company openly advertising the wages it would pay its employees, it did not use its advertisements to specify how much its SAT-prep and speed-reading courses cost—only stressing there would be a “DISCOUNT for Early Registration.”⁴²⁵

Although the bureau asserted test coaching worked in at least some instances, it backtracked on this claim in the report’s second half. The investigators instead insisted the claim “that School A [Kaplan] ‘works’ and School B [Test-preparation] does not” was half-baked.⁴²⁶ Any gains made by students enrolled in a test-coaching course had to be measured against their preexisting levels of underachievement. According to the investigators, if students did poorly on

⁴²² “Classified Ad 714,” *New York Times*, March 16, 1975: E13

⁴²³ “Classified Ad 97,” *New York Times*, April 30, 1978: F41; and “Classified Ad 152,” *New York Times*, May 7, 1978: F47

⁴²⁴ “Classified Ad 357,” *New York Times*, October 10, 1975; “Classified Ad 302,” *New York Times*, September 12, 1976: E11; and “Classified Ad 436,” *New York Times*, August 21, 1977: E12.

⁴²⁵ “Classified Ad 340,” *New York Times*, September 14, 1975: E10; “Classified Ad 619” *New York Times* January 18, 1976: E10 (emphasis in originals). See also: “Classified As 301,” *New York Times*, September 19, 1976: E8.

⁴²⁶ FTC, “Effects of Coaching,” 22.

the SAT “given their demographics and other personal characteristics” before being coached, then coaching may not have caused genuine improvement.⁴²⁷ More bluntly, if a coaching school led students to live up to expectations, then such improvement was not from genuine coaching. Such students purchased coaching through self-selection—that is, recognizing the range of their alleged capabilities and living up to them—rather than the desire to attain outstanding SAT scores.

Had the Bureau unequivocally reported that certain coaching schools could consistently produce genuine improvement in students’ SAT scores, Educational Testing Service might have found itself in an unenviable position. However, ETS may have been partially responsible for the tone of the Bureau’s final product. R. Jeffrey Smith alleged that the Federal Trade Commission was “[h]arassed by [ETS] and nervous about a document lending credence to a suspect industry.”⁴²⁸ The FTC decided to release the document only after a considerable delay and with several pointed caveats—and following the resignation of Arthur Levine, who oversaw the investigation.⁴²⁹ When questioned, *Consumer Affairs* staffers remarked that the findings did not “provide enough information to question the legitimacy of standardized tests themselves”; the study lacked a scientifically genuine control group of students, and thus could not be trusted beyond a minimal degree.⁴³⁰

However tepid, the FTC’s findings were enough for the test-prep industry to claim a measure of legitimacy. Following the report, Stanley Kaplan used rhetorical deference to give his

⁴²⁷ FTC, “Effects of Coaching,” 23.

⁴²⁸ R. Jeffrey Smith, “Test Coaching Dispute Lingers,” in “‘Truth-in-Testing’ Attracts Diverse Support,” *Science* 205, no. 4411 (September 14, 1979): 1114 (sidebar).

⁴²⁹ Larry Kramer, “FTC Sees Cram Courses as Helpful,” *Washington Post*, March 28, 1979: D7.

⁴³⁰ “Cram Courses May Lift Aptitude Test Scores of Some, Study Finds,” *Wall Street Journal*, May 30, 1979; 10.

company extra credibility. By asserting “cram courses do not help,” Kaplan distanced his company from its competition, but also implied to the public that the test-prep industry should not be seen as a cohort of enterprises with a shared agenda or product, but merely a broad spectrum of companies dependent upon the continued use of standardized exams.⁴³¹ Although the FTC study sought to determine if the test-prep industry had engaged in consumer fraud, the results may have ironically spurred some justification for certain test-prep companies to demand high prices for their goods: Kaplan charged over three times as much for SAT prep that Test-Preparation Center, and even factoring in the amount of time spent in direct instruction, a client would have paid twice as much per hour at Kaplan.⁴³² This per-hour cost comparison may seem somewhat misleading because Kaplan had devised dozens of hours of supplemental audio material students could use while at a Kaplan center. But, the use of such material was optional, as a client could use as much or as little of Kaplan’s Test-n-Tape system as he or she desired.⁴³³

Regardless, the FTC report may have reinforced and popularized Kaplan’s method of operation: one paid considerably more for test-prep when offered myriad possible ways to prepare using a company’s materials. One was no longer strictly a student, but a student-consumer, whose purchase signaled not merely a desire to improve, but the monetary means one was willing to expend to reach those ends. If a student were to customize and fully use the supplementary material, then Kaplan and its analogues could reinforce the claim that they were engaging in long-term, multidimensional coaching—and that such holistic learning did not come cheaply. Conversely, clients who ignored the various additional materials could be dismissed as

⁴³¹ Kramer, “FTC Sees Cram Courses as Helpful.”

⁴³² Larry Kramer, “Test Results Affected by Cram Courses; FTC Says Cram Courses May Help,” *Washington Post*, May 30, 1979: E1

⁴³³ Stanley H. Kaplan and Anne Farris, *Test Pilot: How I Broke Testing Barriers for Millions of Students and Caused a Sonic Boom in the Business of Education* (New York: Simon & Schuster, 2001), 42.

careless students and thoughtless consumers, since the company's numerous provisions mitigated that company's burden in guaranteeing significant test score improvement.

As the test-prep industry benefitted from a boost to its credibility, it also became more competitive. During the 1980s, more test-prep companies gained national recognition—none more than The Princeton Review (“TPR”). Founded in New York City in 1981 by John Katzman, TPR held an openly antipathetic posture toward the SAT and similar standardized tests. Katzman maintained that how well or how poorly a person scored reflected nothing more than their ability to take that particular standardized test.⁴³⁴ The SAT was not an exam one should slavishly devote countless hours toward, Katzman assured customers, but an unremarkable and thoroughly predictable test that could easily be bested. For several hundred dollars, students took TPR courses to learn how to game the SAT's structure in order to attain markedly higher scores.⁴³⁵ Both Katzman and his associate Adam Robinson had experience in the test-prep industry prior to The Princeton Review, both having worked at Bob Scheller's test-coaching company Pre-Test Review.⁴³⁶ But, by developing a corporate persona attractive to teenage cynicism, Katzman was able to expand his own business from a boutique Manhattan service for affluent students to a wildly successful set of nationwide franchises and trade paperback prep-books. This corporate persona, when coupled with its two-tier system of

⁴³⁴ Nicholas Lemann, *The Big Test: The Secret History of the American Meritocracy* (New York: Farrar, Straus, & Giroux, 2000), 228. Lemann likens Katzman to Allan Nairn through their social background and rebellion against standardized testing. However, this comparison does not consider that Nairn's polemic was against ubiquity of both standardized testing and ETS itself, whereas Katzman sought a profit from the perceived weaknesses of exams such as the SAT. See also: Shepherd, 112-121.

⁴³⁵ Jonathan Rowe, “Preparing for the SAT,” *Christian Science Monitor*, September 19, 1986: Ideas 21; microfilm. In the story, a TPR course was said to typically cost “between \$500 and \$600,” which would be roughly equivalent to \$1,110 to \$1350 in 2017; see Bureau of Labor Statistics, CPI Inflation Calculator, <https://data.bls.gov/cgi-bin/cpicalc.pl> (last accessed March 15, 2017).

⁴³⁶ David Owen, “Adam and John say put your pencil down; The SAT is easier that you: a) fear; b) think; c) guess; d) all of the above,” *Rolling Stone* 444 (March 28, 1985): 74, 77; microfilm.

expensive courses and inexpensive prep books, allowed TPR to serve the educational class demands of an elite clientele while still profiting from the aspirations of a middle-class book-buyers.⁴³⁷

Even in its earliest advertisements, TPR used a glib and sarcastic tone to establish its stance toward standardized admissions exams. One early ad, printed in the *New York Times* in late summer, 1981, hooked readers with the remark “Some people have an Uncle on the Admissions Board...and some people have The Princeton Review.”⁴³⁸ Although the ad noted that TPR courses were taught “by a staff of Ivy League alumni and undergraduates” using the latest computer technology, it remained silent on how an eight-week SAT prep course actually cost.⁴³⁹ Such ads subtextually commended potential clientele for not being from an older order of established money and power, while casually ignoring that these customers had enough disposable income to pay for boutique test-prep.⁴⁴⁰ TPR also reinforced this position through its slogan. By declaring itself “the nemesis of standardized testing” in early advertisements, TPR

⁴³⁷ Joshua Hammer, “Cram Scam: Fighting Educational Injustice for Fun and Profit,” *The New Republic* April 24, 1989: 15-18. Hammer noted that about 20,000 students took TPR courses among 1.7 million SAT test-takers in 1989—reflecting a small but highly profitable set of students angling for spaces in elite schools. At \$645 dollars a course (15), this clientele likely generated over \$12 million in revenues that year alone. Katzman also profited from TPR’s franchise structure, in which each of TPR’s 40 franchises had to “pay Katzman between \$100,000 and \$200,000 up front, buy course materials from him, and hand over eight percent of revenues” (17).

⁴³⁸ “Display Ad 1349 -- no Title.” *New York Times (1923-Current File)*, Aug 30, 1981. 488, <https://search.proquest.com/docview/121681292?accountid=14745> (last accessed March 15, 2017). For a similar early ad from TPR, please see: “Classified Ad 128 -- no Title.” *New York Times (1923-Current File)*, Sep 01, 1981. 1, <https://search.proquest.com/docview/121678606?accountid=14745>; and “Classified Ad 126 -- no Title.” *New York Times (1923-Current File)*, Sep 06, 1981. 1, <https://search.proquest.com/docview/121596977?accountid=14745> (last accessed March 15, 2017).

⁴³⁹ *Ibid.*

⁴⁴⁰ For an example of a TPR ad with similar sensibilities, see: “Display Ad 569 -- no Title.” *New York Times (1923-Current File)*, May 18, 1986. 1, <https://search.proquest.com/docview/110938448?accountid=14745> (last accessed March 15, 2017). This ad asserts “[i]f people have told you that you can’t significantly increase your SAT score [...] THEY LIED!” This exclamation sits between two cartoons: on the right, a white student carrying books, curiously reading the hook; on the left, an older white man in a three-piece suit, with folded arms and an upturned nose. This imagery and rhetoric suggests TPR sought to present itself as a company in opposition to an older order of academic and social elite—without necessarily alienating those who could afford the course.

could signal that it had no fondness for testing practices, while sidestepping the obvious fact that its existence dependent upon the sustained popularity of the SAT and other entrance exams.⁴⁴¹

The Princeton Review also used its advertisements to set itself apart as cooler and more effective than its competition—particularly Kaplan. A series of TPR ads printed in 1985 triumphantly declared “WE SCORE MORE!”⁴⁴² Above this exclamation lay a bar graph representing the alleged improvement to students’ SAT scores after various test-prep methods. The Princeton Review’s 150-point increase not only dwarfed Kaplan’s modest 50-point gains, but burst through the graph itself. Ads from later in the decade took more direct jabs. One, printed in the *Wall Street Journal* in 1988, quipped “If You Can’t Afford The Princeton Review, Stanley Kaplan Is An Acceptable Compromise.”⁴⁴³ This line of advertising mocked those who chose inferior, understaffed, and outdated Kaplan SAT prep courses for the sake of maintaining their family budgets. TPR courted its clientele by posing its SAT courses as an investment with tremendous payoff—without ever suggesting that its target demographic could purchase classes without making a financial sacrifice, that Kaplan had the same target demo, or that all students consuming test-prep material had the same college (and class) aspirations.

Some TPR ads employed several clever techniques at once, including deceptive sources of praise. One 1988 ad in *Wall Street Journal* joked, “Not everyone tests well,” below which lay six comically drawn figures illustrating the evolution from chimp to modern man.⁴⁴⁴ Labeled “Week 1” to “Week 6,” the figures suggested TPR could help students become “highly evolved”

⁴⁴¹ *Ibid.*

⁴⁴² "Display Ad 35 -- no Title." *New York Times (1923-Current File)*, Feb 12, 1985. 1, <https://search.proquest.com/docview/111241858?accountid=14745> (last accessed March 15, 2017).

⁴⁴³ Display Ad, *Wall Street Journal*, August 24, 1988: 10; microfilm. For a reprint of this ad, please see: Display Ad, *Wall Street Journal*, September 14, 1988: 8; microfilm.

⁴⁴⁴ Display Ad, *Wall Street Journal*, February 23, 1988: 10; microfilm.

test-takers. The ad copy included snippet endorsements from *Rolling Stone* and *The Christian Science Monitor*, appealing to the sensibilities of both clients and their parents. But the testimonial attributed to *Rolling Stone* was misleading. As printed, it stated:

Rolling Stone reports, “Princeton Review Students are being admitted to college that wouldn’t have considered them before...When you raise a kid’s score 200 or 300 points, ...you change his whole outlook about himself.”⁴⁴⁵

These statements, while attributable to David Owen’s glowing profile of Katzman and Robinson for *Rolling Stone*, were not written by Owen but said by Katzman himself. The second ellipses simply erased any indication of who said what.⁴⁴⁶ For TPR, even self-generated praise could be turned into a marketing material.

By the middle of the 1980s, Katzman expanded TPR’s product line to include test-prep workbooks—courting a far wider consumer base than students who could afford expensive prep courses. Released in 1986, *Cracking the SAT* built off of TPR’s cynical corporate persona, attracting students who also believed the SAT was an obnoxious hassle but may not have shared the same Ivy League aspirations as the company’s earlier clientele. For ten dollars, any student could learn how to manipulate the mechanics and structure of the SAT to their own benefit, and editorialists praised the bestseller for “exposing the gimmicks and flaws of [... the] SAT.”⁴⁴⁷ To this end, any student who purchased *Cracking the SAT* could also feel as through they had tapped

⁴⁴⁵ *Ibid.* See also: *Educational Testing Service v. John Katzman and The Princeton Review*, 631 F. Supp. 550 (1986) 551, note 1; available via LexisNexis Academic, <http://www.lexisnexis.com.ezproxy.lib.usf.edu/hottopics/lnacademic/> (accessed April 21, 2013). Judge Maryanne Trump Barry noted that another advertisement for The Princeton Review may have given the false impression that the company, and not Katzman’s work experience in the test prep industry, had dated back to 1977.

⁴⁴⁶ Display Ad, *Wall Street Journal*, February 23, 1988: 10; microfilm. See also, Owen, “Adam and John say put your pencil down,” 80. Owen’s fondness for TPR has been well documented; see: Hammer, “Cram Scam,”; and Gerald W. Bracey, “ETS as Big Brother: An Essay-Review,” *Phi Delta Kappan* 67, no. 1 (September 1985): 75-79.

⁴⁴⁷ Michael Kinsey, “‘SATs Are a) Flawed[;] b) Comical[;] c) Both of the Above,’” *Wall Street Journal*, September 4, 1986: 27; microfilm.

into coaching secrets typically reserved for upper- and upper-middle class teens. *Cracking the SAT* offered the sheen of privilege without any of its material perks.

Cracking the SAT used a variety of rhetorical and pedagogical techniques to make its readers feel like savvy consumers who, with TPR's help, could easily avoid the College Board's tricks and traps. David Owen, who wrote the book's foreword, insisted that preparing for the SAT was a way to safeguard against dimwitted admissions officers, who "won't understand how to interpret the scores you send them" but nonetheless "[speculate] freely about your intelligence and even your personality on the basis of how you do."⁴⁴⁸ Katzman likewise assured customers that "the SAT is not written by Nobel Prize-winning physicists or Pulitzer Prize-winning journalists," just "ordinary folks who had ordinary test scores."⁴⁴⁹ Because normal people wrote the SAT, one did not have to be a genius to do well. Rather, test takers needed to remember a few basic rules: always guess rather than leave an answer blank, no matter what ETS or your gut says; mark up the test booklet as much as you need; and rack up points by focusing on the earlier questions in a set because they were invariably easier. Clever test-takers, Katzman stressed, only bothered memorizing the 200 or so vocabulary words the College Board routinely used, and worked backwards to save time on algebra questions. Readers of *Cracking the SAT* also learned to avoid acting like Joe Bloggs, a thoroughly average student whose "hunches are *always* wrong on on difficult questions."⁴⁵⁰ Test-takers should always go for answers Joe Bloggs would choose early in each verbal and math section, Katzman noted, and avoid all answers Joe would be lured by at the end of each section. Once students mastered avoiding Joe Bloggs's pitfalls for each

⁴⁴⁸ David Owen, foreword for *Cracking the System: The SAT*, by Adam Robinson and John Katzman (New York: Villard Books, 1986), v.

⁴⁴⁹ Katzman, *Cracking the SAT*, xiii.

⁴⁵⁰ Katzman, *Cracking the SAT*, 27.

question type, they could have their hand at a full-length practice test—which, for an additional seven dollars, the Princeton Review would score and offer detailed computer-generated analysis. Through *Cracking the System*, Katzman turned TPR into a brand that made customers at all price points feel too clever to be average.

By the mid-1980s, Katzman was an ascendant figure who used relatively controversial behavior to his company’s advantage. Such braggadocio occasionally crossed into illegal actions, leading to a series of lawsuits and appeals between Katzman and Educational Testing Service during the mid-1980s. The rulings in these cases further entrenched the test-prep industry in a contentious but ultimately mutually beneficial relationship with standardized test makers over the legality of certain business practices. Legal actions began in August, 1985, when ETS filed a civil action against both Katzman and The Princeton Review in the United States District Court for New Jersey. ETS aimed to forbid the test-prep company from using several hundred “confidential test questions” allegedly pilfered by Katzman and his employees for TPR clients.⁴⁵¹ Katzman and TPR were accused of coaching clientele with “facsimile” versions of SAT-Math, SAT-Verbal, and other College Board test questions--causing ETS to discontinue nearly 300 “secure” test questions slated for use in the 1985 versions of the SAT.⁴⁵² ETS maintained Katzman’s theft of secure test questions not only compromised the integrity of the test but, more importantly, was an act of copyright infringement.

⁴⁵¹ *Educational Testing Service v. John Katzman and The Princeton Review, Inc.*, 626 F. Supp. 527 (1985), 528.

⁴⁵² *Educational Testing Service v. John Katzman, The Princeton Review, Inc., Robert S. Scheller, and Pretest Review, Inc.; John Katzman and The Princeton Review, Inc., Appellants*, 793 F.2d 533 (1986), 536; see also: *Educational Testing Service v. John Katzman and the Princeton Review* 631 F.Supp 550 (1986), at 551, in which Judge Marianne Trump Barry explains in her decision: “The SAT and the Achievement Tests are copyrighted by ETS and, pursuant to 37 C.F.R. § 202.20, are secured tests which are only viewed by the public during the administration of those tests to registered examinees.”

As the case opinions illustrated, ETS and already reached a détente with Katzman over earlier instances of theft, when TPR was a fledgling sole proprietorship. In 1982, Katzman came under the possession of “certain Math and English Composition Achievement Tests,” and distributed material from these exams to his clientele prior to their official use by ETS.⁴⁵³ Katzman refused to reveal how he attained the material, but reached an accord with ETS the next year with the following terms:

to return all copies of the purloined tests, to refrain from copying or distributing any ETS copyrighted or copyrightable materials or registering for or attending any test administered by ETS unless it was for bona fide purposes, and to notify ETS if any unlawfully obtained ETS tests came into [his] possession and provide ETS with information as to their source.⁴⁵⁴

ETS alleged that Katzman betrayed that agreement by copyrighting test-prep questions too similar to actual SAT questions. The test-maker also indicated that in early 1985, “Katzman attempted to take the SAT as a standby candidate at Dwight-Englewood School in Englewood, New Jersey,” a tactic also used by other TPR employees after the earlier agreement.⁴⁵⁵ These combined actions indicated to ETS that Katzman sought to profit off of the test-maker’s creations and had no intention to stop.

The judges for these multiple cases had to determine whether the state of New Jersey was the proper jurisdiction for ETS to file suit against Katzman or TPR and, most pressingly, whether objects like standardized test questions were novel enough items to merit a copyright infringement suit. New Jersey District Court Judge Maryanne Trump Barry deduced that, despite being a New York resident, “Katzman [was] amenable to suit in New Jersey” because his 1983

⁴⁵³ *Educational Testing Service v. John Katzman and The Princeton Review*, 631 F. Supp. 550, (1986), 551.

⁴⁵⁴ *Educational Testing Service v. John Katzman, The Princeton Review, Inc., Robert S. Scheller, and Pretest Review, Inc.*, 793 F.2d 533 (1986), 536

⁴⁵⁵ *Educational Testing Service v. John Katzman and The Princeton Review*, 631 F. Supp. 550 (1986), 552.

arrangement with Princeton-based Educational Testing Service was valid precedent for his contact in the state, even if TPR had been incorporated in another state the interim.⁴⁵⁶ Third Circuit Court of Appeals Judge Dolores Sloviter upheld that Katzman’s distribution of “facsimile” test questions was unjustifiable. Sloviter found both of Katzman’s arguments about ETS’s choice of copyright lacking. Even if ETS copyrighted their tests as compilations, Sloviter noted, the test questions within such compilations were still protected by copyright.⁴⁵⁷ The judge also dismissed the claim that ETS questions were uncopyrightable ideas rather than copyrightable expressions of ideas, and declared TPR’s educational benefits were too minimal to consider the academic fair use loopholes for breach of copyright.⁴⁵⁸ But, although Sloviter found some of TPR’s “facsimile” questions to be so similar to ETS originals to essentially be unwarranted copies, the judge also declared that the preliminary injunction overreached. Sloviter ruled that the earlier prohibition against “‘adapting’ ETS’ materials may [have encompassed] permissible use of ETS’ material,” as Katzman and TPR were unduly prevented by the

⁴⁵⁶ *Educational Testing Service v. John Katzman and The Princeton Review*, 631 F. Supp. 550 (1986), 561. Barry’s opinion on Katzman’s motion to dismiss provides a highly detailed but well-written interpretation of minimum contacts analysis.

⁴⁵⁷ *Educational Testing Services v. John Katzman, The Princeton Review, Inc., Robert S. Scheller and PreTest Review, Inc.; John Katzman and The Princeton Review, Inc., Appellants*, 793 F.2d 533 (1986), 538-539. Citing *Northwestern Bell Telephone Co. v. Bedco of Minnesota, Inc.*, (501 F. Supp. 299, 301-02; D. Minn. 1980), Sloviter explained that in some instances, most notably telephone books, a copyright for a compilation protects the collection even if none of the component parts (such as addresses and telephone numbers) passed the minimum standard for copyright. Compilation copyright, Sloviter noted, also extended to the material within a compilation created by the author of the compilation. ETS filed for copyright compilation because new versions of the SAT included test items from earlier versions—and although these older test questions’ copyright was not altered by the compilation’s copyright, their copyright did not suddenly disappear by basis of being included within a compilation.

⁴⁵⁸ *Educational Testing Services v. John Katzman, The Princeton Review, Inc., Robert S. Scheller and PreTest Review, Inc.; John Katzman and The Princeton Review, Inc., Appellants*, 793 F.2d 533 (1986), 539, 543. As Sloviter details, what Katzman and the other defendants seemed to base their defense on was the merger principle, in which the number of ways to express an idea are so limited that an expression itself is essentially in merger with its underlying idea and thus not subject to copyright. Neither approach swayed Sloviter, who quipped “We are, quite frankly, unpersuaded that the number of questions that can be devised to test students on their knowledge of square roots or dangling participles is so limited that ETS’ questions designed for this purpose represent a merger with the underlying ideas” (540).

injunction from potential legitimate business practices.⁴⁵⁹ Whatever Educational Testing Service's special needs were, it could not impede upon the business of testing.

What complicated Educational Testing Service's litigious behavior was that test manufacturers had already entered the test-prep market. Between truth-in-testing legislation and the FTC's qualified legitimization of test-coaching, standardized testing manufacturers found themselves in a precarious position at the beginning of the 1980s. How could they continue functioning if correct test answers had to be publicly disclosed, and if test-coaching companies could continue expanding operations? Such changes could have created prohibitively large research and development budgets. Test manufacturers found a solution to both problems in the same place: by taking material rendered obsolete by disclosure policies and selling it in the marketplace, thereby upending truth-in-testing advocates' class-based complaints and consumer politics.

By releasing old material, test-makers' could reap both profits and public esteem while, at the same time, maintaining their position that short-term practice was largely ineffective. The College Board's first mass-market preparatory product, *4 SATs*, was a direct consequence of truth-in-testing legislation. Its popularity led to the commercially sold *10 SATs* in 1983.⁴⁶⁰ For the first time, students could become acquainted with the SAT through previous versions of the exam. But the organization also stressed that little good would actually come from sitting down and taking all ten tests—nor thinking anything beyond long-term preparation would do much

⁴⁵⁹ *Educational Testing Services v. John Katzman, The Princeton Review, Inc., Robert S. Scheller and PreTest Review, Inc.*; *John Katzman and The Princeton Review, Inc., Appellants*, 793 F.2d 533 (1986), 545.

⁴⁶⁰ David Owen, "SAT Coaching Guides: Do They Work?," *Washington Post*, August 12, 1984: ER5; microfilm. *10 SATs*, Owen notes, was published by Scribner Book Companies, Inc. See also: Peter Spiro, "Test Prep Panic," *The New Republic* December 20, 1982: 13. See also: Shepherd, 127-131.

good.⁴⁶¹ As some critics noted at the time, many commercial SAT-prep products on the market were ill suited to prepare students for the actual test. Many prep guides either mangled the logical structure of question types or contained “bizarrely superfluous” vocabulary lists and reference guides with little relevance to the SAT.⁴⁶² Other products relied on new computer technology, and offered practice tests stored on floppy disk, sometimes for as much as \$300.⁴⁶³ The College Board’s commercial material, by contrast, automatically held market appeal because it was guaranteed to be as similar as possible to the actual test material; the SAT’s validity and reliability depended on relative consistency between subsequent versions of a test, so consumers could rest assured that SAT released because of transparency laws were close to what they would encounter on test day.⁴⁶⁴ What else could compare to the real thing, particularly when it cost less than ten dollars? Commercially publishing obsolete material allowed the College Board to convey that truth-in-testing legislation would not hobble the quantity or quality of its testing content, while signaling to test-prep companies that even the release of genuine test questions would not undermine entire system.

Although The Princeton Review ultimately lost its legal battle against Educational Testing Service, Katzman’s company demonstrated that the test-prep industry’s legitimacy during the 1980s hinged in part upon companies displaying distinguishable attitudes toward the

⁴⁶¹ Mark J. Sherman, "The College Board Joins Publishers of S.A.T. 'Cram' Books." *New York Times*, Jan 08, 1984; Lisa Belkin, "Preparing for S.A.T.'s: The Crunch in On," *New York Times* October 6, 1984: 48.

⁴⁶² Owen, "SAT Coaching Guides: Do They Work??"

⁴⁶³ Belkin, "Preparing for S.A.T.'s,"; Dianna Harms, "A Student Tests S.A.T. Software," *Antic* 4, no. 6 (October, 1985), Antic archive, <http://www.atarimagazines.com/v4n6/SATSoftware.html> (last accessed 16 March, 2017). The price point was not nearly as high for new computer-based SAT prep guides by the end of the decade; see: L.R. Shannon, "Peripherals: S.A.T. Guide on Disk," *New York Times*, September 13, 1988: C8.

⁴⁶⁴ The College Board did release its own line of computer-based prep, too, but TestSense was originally intended for use in secondary schools as a device to track hundred of students’ potential performance on the PSAT; see: Peter H. Lewis, "Peripherals: Help with the S.A.T.'s" *New York Times*, October 28, 1986: C6.

SAT and standardized testing in general. (Katzman claimed his legal defeat was actually a victory; the two years and hundreds of thousands of dollars ETS invested in its lawsuit only led to a reward of “\$52,000 in damages” and a massive growth in The Princeton Review’s business.)⁴⁶⁵ Unlike Kaplan, The Princeton Review publicly held the SAT in contempt—and Katzman would later liken the exam to “a cancer.”⁴⁶⁶ But Katzman’s disgust for the very product that made his company possible makes sense when seen as an act of creating a corporate persona. By publicly sharing a disdain for the exam with TPR’s key demographic, Katzman allowed clients to believe that their choice in test-prep could also reflect their view toward standardized testing—creating handsome profits for Katzman and his company in the process.

The Knotty Logic of the Contemporary Educational Marketplace

In the ensuing decades, certain test-prep companies have developed into massive enterprises that court additional groups of educational consumers. No test-prep company has experienced as much change as Kaplan, which grew into a for-profit educational provider that now keeps an entire media conglomerate afloat. In November 1984, the Washington Post Company purchased Kaplan, indicating a desire to broaden the test-prep company’s scope of services and market penetration.⁴⁶⁷ Although Kaplan earned \$35 million from nearly 95,000 clients that year, the company reached “only a small percentage” of high school students who took the SAT at the time.⁴⁶⁸ Despite a steady growth in clientele, market reach, and revenues,

⁴⁶⁵ Lee Mitgang, “Educational Testing Service Wins Copyright Victory Over Coaching Firm,” *Associated Press Wire*, December 23, 1987.

⁴⁶⁶ Tony Schwartz. “The Test Under Stress.” *New York Times* January 10, 1999. This is not the most vitriolic term Katzman used for the SAT, once dismissing it outright as “bullshit”; see: Katzman, quoted in Owen, “Adam and John Say Put Your Pencil Down,” 80.

⁴⁶⁷ David Vise. “Post to Buy Educational Prep Firm,” *Washington Post*, November 21, 1984: C1; microfilm.

⁴⁶⁸ Vise, “Post to Buy Educational Prep Firm,” C2; Phillip H. Wiggins “Washington Post to Buy Tutoring Centers Chain,” *New York Times* November 21, 1984: D5; Stanley H. Kaplan, *Test Pilot: How I Broke Testing*

Kaplan remained a relatively minor holding for Washington Post Company until the turn of the century.⁴⁶⁹ Beginning in the late 1990s, Kaplan segued into offering its own higher educational services, purchasing online postsecondary programs such as Concord Law School that had already received accreditation. Kaplan thus plugged into the nascent for-profit higher education market without investing too much on infrastructure.⁴⁷⁰ By 2005, these higher education ventures overtook test-prep as Kaplan's most revenue-generating sector.⁴⁷¹ It was the Great Recession, however, that made Washington Post Company "more dependent than ever on a single business: Kaplan"—and in particular, its higher education division.⁴⁷² As economic calamity drove the newspaper industry into a tailspin, it also pushed economically insecure Americans into for-profit colleges, allowing the Washington Post Company to weather the fall of print journalism at the cost of becoming dependent on trends in higher education.

Kaplan's expansion into higher education also illustrates the complicated relationship between test-prep companies, their sources of profit, and the state. Of the \$1.9 billion dollars in revenue Kaplan, Inc. generated in 2015, about 45% came from its Higher Education division.⁴⁷³

Barriers for Millions of Students and Caused a Sonic Boom in the Business of Education (New York: Kaplan Publishing, 2001), 141; Lemann, *Big Test*, 227; Shepherd, 121-127, 140-145.

⁴⁶⁹ Between 1995 and 1999, Kaplan's operating revenues nearly tripled, from \$89 million to \$259 million. But its revenues for 1999 were still only 12% of total operating revenues for the Washington Post Company that year. See: Washington Post Company, *1999 Annual Report to Stockholders*, iii, 31, http://media.corporate-ir.net/media_files/irol/62/62487/reports/ar99/wpo_000331_200_120_n.pdf

⁴⁷⁰ Washington Post Company, January 3, 1999 Form 10-K, (filed March 26, 1999), hardcopy p.13, U.S. Securities and Exchange Commission EDGAR Database, <http://www.sec.gov/Archives/edgar/data/104889/0000950133-99-000908.txt>; Washington Post Company, December 31, 2000 Form 10-K (filed March 23, 2001), hardcopy p.17; EDGAR, <http://www.sec.gov/Archives/edgar/data/104889/000095013301001017/w46916e10-k.txt> (last accessed March 17, 2017).

⁴⁷¹ Washington Post Company, 2007 10-K, 46; Washington Post Company, 2004 10-K, 30, both available at Graham Holdings Investor Relations, <http://www.ghco.com/phoenix.zhtml?c=62487&p=irol-reportsannual>.

⁴⁷² Donald E. Graham, "To Our Shareholders," Washington Post Company 2009 Annual Report, 2.

⁴⁷³ The Washington Post Company changed its name to Graham Holdings in 2013 following the sale of the newspaper to Amazon founder Jeff Bezos; see: Debbi Wilgoren, "Washington Post Co. Renamed Graham Holdings Company to Mark Sale of Newspaper," *Washington Post* November 18, 2013.

While this was considerably lower than its 2010 peak of \$2.9 billion, Kaplan nonetheless provided a greater share of Graham Holding's profits than ever before, accounting for nearly three-quarters of the conglomerate's revenue that year.⁴⁷⁴ Kaplan University, which ran for-profit degree programs for nearly 40,000 students nationwide, derived its greatest source of funds from federal student financial aid, particularly Title IV loans. These loans provided Kaplan University about 628 million dollars by 2015—and by extension, nearly a quarter of all Graham Holdings operating revenue. Kaplan University took in enormous revenues even after the federal government had reached its breaking point with for-profit colleges' student loan profit schemes; the Department of Education had already caught Kaplan padding tuition bills and matriculation fees by thousands of dollars, and had punished the school for producing graduates with extremely high debt-to-earnings ratios and abysmal loan repayment rates.⁴⁷⁵ Although test-prep

⁴⁷⁴ Graham Holdings, 2015 Annual Report, ii, 1, 3; Graham Holdings, 2015 10-K, 1; Washington Post Company, 2012 10-K, 1; Washington Post Company, 2009 10-K, 1; all available at Graham Holdings Investor Relations, <http://www.ghco.com/phoenix.zhtml?c=62487&p=irol-reportsannual> (last accessed March 17, 2017).

⁴⁷⁵ Graham Holdings, 2015 10-K, 1-11; Chris Kirkham, "At Kaplan University, 'Guerilla Registration' Leaves Students Deep In Debt," *Huffington Post*, December 22, 2010, http://www.huffingtonpost.com/2010/12/22/kaplan-university-guerilla-registration_n_799741.html; U.S. Attorney's Office, Western District of Texas, "For-Profit College Kaplan to Refund Federal Financial Aid Under Settlement With United States," press release, January 5, 2015, <https://www.justice.gov/usao-wdtx/pr/profit-college-kaplan-refund-federal-financial-aid-under-settlement-united-states>; Jacob Alderdice, "The Informed Student-Consumer: Regulating For-Profit Colleges by Disclosure," *Harvard Civil Rights-Civil Liberties Law Review* 50 (2015): 216-254; David Wessel, "Lots Riding on Ed Dept Standard for Student-Loan Forgiveness," *Brookings Institute*, June 18, 2015, <https://www.brookings.edu/opinions/lots-riding-on-ed-dept-standard-for-student-loan-forgiveness/>; and Adam Looney and Constantine Yannelis, "A Crisis in Student Loans?: How Changes in the Characteristics of Borrowers and in the Institutions They Attended Contributed to Rising Loan Defaults," *Brookings Papers on Economic Activity* Fall 2015: 1-89, <https://doi.org/10.1353/eca.2015.0003>.

Washington Post Company, responding to what it felt were unfair intrusions on its business practices, conducted its own analysis of student loan use and repayment across all sectors of higher education. Its analyst concluded that for-profit higher education providers such as Kaplan actually had a harder burden to carry because they served "nontraditional" students who typically had lower incomes and less access to capital—and thus, were less likely to repay their loans to a degree satisfactory to existing federal standards. Washington Post maintained that the federal government's turn toward stricter student loan policies would in effect harm historically underserved communities because tougher loan repayment guidelines would compel for-profit schools to turn away students who had no ability to pay upfront. The real solution, they argued, was less regulation—not only in the rules that allegedly prevented for-profit schools from lowering tuition (or, rather, removing the subsidies that kept public and nonprofit private schooling relatively inexpensive), but also removing the 90-10 cap that mandated at least 10% of school tuition revenue come from sources outside federal loans (barring the well-exploited loophole that did not count

continues to generate hundreds of millions in annual revenues for Kaplan, for-profit education has become its most profitable enterprise, deeply ensnaring the corporation within education policy and federal funds.

As certain test-prep companies took on other educational roles, some test-makers started acting more like sales-focused corporations. Over time, the College Board's changes to the SAT seemed less like efforts to fine-tune its flagship test and more like image-conscious rebranding efforts. Addressing longstanding concerns with the integrity of the test score scale, the College Board completely recentered the SAT in 1995. Although recentering the test eliminated the convoluted statistical process by which the College Board equated new versions of the SAT back to a version created in 1941, some critics (notably educational historian Diane Ravitch) suggested the most visible side effect—a dramatic rise in SAT Verbal scores—was a way to placate students while sidestepping declining nationwide standards in language and reading education.⁴⁷⁶ Other efforts by the College Board, such as changing the test's name from “Scholastic Aptitude Test” to simply “SAT,” seemed more overtly focused on improving public relations.⁴⁷⁷ Even major changes to the composition of the SAT allowed the College Board new opportunities for profit. Before debuting an overhauled version of the SAT in 2005, the College

veteran benefits as federal financial aid). See: Washington Post Company, 2012 10-K, 1-10; Mark Kantrowitz, “Student Aid Policy Analysis: the Impact of Loan Repayment Rates on Pell Grant Recipients,” September 1, 2010, and Donald E. Graham, “2010 Letter to Shareholders,” February 23, 2010, Graham Holdings Investor Relations, <http://www.ghco.com/phoenix.zhtml?c=62487&p=irol-reportsannual>.

⁴⁷⁶ College Board Office of Research and Development, “The Effects of SAT Scale Recentering on Percentiles,” Research Summary RS-05, May 1999, <http://files.eric.ed.gov/fulltext/ED563025.pdf>; Sage Stossel, “Hard Lessons: Diane Ravitch Argues for a Return to Academic Rigor in Our Nation's Public Schools,” *The Atlantic*, November 1, 2000, <https://www.theatlantic.com/past/docs/unbound/interviews/ba2000-11-01.htm> (last accessed March 18, 2017); Neil J. Dorans, *The Recentering of SAT Scales and Its Effects on Score Distributions and Score Interpretations* (New York: College Entrance Examination Board, 2002), <https://www.ets.org/Media/Research/pdf/RR-02-04-Dorans.pdf>

⁴⁷⁷ Peter Applebome, “Insisting It's Nothing, Creator Says SAT, Not S.A.T.,” *New York Times*, April 2, 1997: B6.

Board released *The Official SAT Study Guide for the New SAT*, which included eight practice tests to familiarize students with the test’s new essay section and grammar questions.⁴⁷⁸ By placing test-prep companies in a reactive position, the College Board could reassert itself as the best possible source for SAT prep—banking on reliability and affordability to fuel student-consumers’ satisfaction for their line of test-prep and, by extension, the SAT itself.

The College Board has since continued to blur the line between test-makers and test-prep companies, posing their organization as the only ethical test-prep provider for those who aspire for a secure spot in the middle class. The organization rolled out a revamped SAT in March 2016, touting the redesign as part of a larger “opportunity agenda” to preserve the bond between college education and middle-class employment.⁴⁷⁹ Gone were the requisite essay, the traditional guessing penalty, and “vocab you’ll never use again.”⁴⁸⁰ The College Board instead assured teens that the new SAT focused on “the stuff you’ve been learning in high school, the stuff you’ll need to succeed in college,” such as graphic interpretation, contextual meaning, and evidentiary support.⁴⁸¹ Citing trenchant, widespread rates of remedial college coursework—and affirming that college education was increasingly necessary for stable employment—the College Board premised the SAT overhaul on its belief that, if refined correctly, the standardized test could serve as a potent conduit between secondary education and middle-class life. In the

⁴⁷⁸ Lucinda Dyer, “S.A.T.—Adding the Write Stuff,” *Publishers Weekly*, August 2, 2004: 35, 36, 40, 42, 44, 46; Tamar Lewin, “College Board to Revise SAT After Criticism by University,” *New York Times* March 23, 2002: A 10.

⁴⁷⁹ College Board, *Test Specifications for the Redesigned SAT* (College Board, 2015), 197; <https://collegereadiness.collegeboard.org/pdf/test-specifications-redesigned-sat-1.pdf>.

⁴⁸⁰ “SAT,” *College Board* website, <https://collegereadiness.collegeboard.org/sat?navid=gh-nsat>.

⁴⁸¹ “Math Test,” *College Board* website, <https://collegereadiness.collegeboard.org/sat/inside-the-test/math>

College Board’s view, revising the SAT could redress systemic problems in both education and the economy.⁴⁸²

Central to the College Board’s revamped agenda for the SAT was the organization’s partnership with the online nonprofit educational service Khan Academy. Khan Academy agreed to design “free, world-class prep materials” to mitigate the class-based disparities that many American students faced when preparing for the SAT.⁴⁸³ This preparatory material—which included “hundreds of previously unreleased questions[,] videos with step-by-step instructions” and “adaptive and game-based online instructional offerings”—aligned with Khan Academy’s existing series of autodidactic materials.⁴⁸⁴ Noting its “close collaboration with the authors of the SAT themselves,” as well as its development of “sophisticated, interactive software,” Khan Academy assured future test-takers that its preparatory material would be both legitimate and affordable.⁴⁸⁵ In remarks to the *New York Times*, College Board President David Coleman portrayed the relationship as a part of larger class struggle, asserting “[i]t is time for the College Board to say in a clearer voice that the culture and practice of costly test preparation that has arisen around admissions exams drives the perception of inequality and injustice in our country.”⁴⁸⁶ Coleman insisted the College Board had an obligation to undermine the way

⁴⁸² In 2013, 1.7 million students took the SAT. While it had long been the most popular standardized college entrance exam, it has recently slipped just behind the ACT. See: Tamar Lewin, “A New SAT Aims to Realign With Schoolwork.” *New York Times*, March 5, 2014, <http://nyti.ms/1icHkcn>

⁴⁸³ “Khan Academy Partners With College Board to Provide free SAT Prep,” Khan Academy, <https://www.khanacademy.org/test-prep/sat/overview-sat-prep/a/khan-academy-partners-with-college-board-to-provide-free-sat-prep>.

⁴⁸⁴ Benjamin Herold, “Khan Academy Recruited to Provide Online SAT Prep,” *Education Week* 33, no. 24 (March 12, 2014): 14.

⁴⁸⁵ “Let’s Level the Playing Field for SAT Prep,” Khan Academy blog, <https://www.khanacademy.org/about/blog/post/78670138358/lets-level-the-playing-field-for-sat-prep>.

⁴⁸⁶ David Coleman, quoted in Lewis, “A New SAT Aims to Realign With Schoolwork.”

“some test-prep providers intimidate parents at all levels of income.”⁴⁸⁷ Salman Khan echoed these sentiments when interviewed for *Time*, insisting “If the test-prep companies think they can add value, I’m sure they’ll try. But from my point of view, this is the College Board making sure it’s not about money.”⁴⁸⁸ The College Board’s partnership with Khan Academy was the next rhetorical step in its relationship with student-consumers: the College Board was the only source students-consumers could trust to guide them through College Board exams and into a middle-class life path.

The College Board’s partnership with Khan Academy sparked a heated (albeit unsurprising) turf war with the standardized test-prep industry. Test-prep companies lobbed dismissive statements about the College Board’s plans. Paul Kanarek, an executive for the Princeton Review, admitted that “any test prep provider who [was] focused on teaching content [was] now ruined,” but that his company had little to worry about, as its strategy-based “philosophy [had] never been content-oriented.”⁴⁸⁹ One representative for Kaplan Test Prep acknowledged that SAT prep, whatever its form, was crucial—but that consumers will still flock to trusted companies such as Kaplan to secure an advantage.⁴⁹⁰ Even competing test manufacturers inserted their opinion. Executives from rival standardized test manufacturer ACT diminished the College Board’s plans, asserting that ACT had long valued creating testing

⁴⁸⁷ Coleman, quoted in Herold, “Khan Academy Recruited.”

⁴⁸⁸ Salman Khan, quoted in Charlotte Adler, “Free SAT Class Could Flink Prep Companies,” *Time*, March 6, 2014; <http://time.com/13390/sat-test-prep-khan-kapaln-college-board>.

⁴⁸⁹ Jia Lynn Yang, “The College Board is Giving Away Test Prep For Free. Why That Won’t Change Much.” Wonkblog, *Washington Post* website, March 5, 2014, <http://www.washingtonpost.com/blogs/wonkblog/wp/2014/03/05/why-giving-away-test-prep-for-free-wont-change-much/> (last accessed December 5, 2014); see also: Felix Gillette, “Why the Test-Prep Industry Isn’t Afraid of Khan Academy’s SAT Freebies,” *Bloomberg Businessweek*, March 6, 2014; <http://www.businessweek.com/articles/2014-03-06/why-the-test-prep-industry-isnt-afraid-of-khan-academys-sat-freebies> (last accessed December 5, 2014).

⁴⁹⁰ Lewis, “A New SAT Aims to Realign With Schoolwork”; Stephanie Simon, “To Thwart Test Prep Firms, SAT Gets Makeover,” *Politico*, March 5, 2014; <http://www.politico.com/story/2014/03/sat-test-prep-makeover-104291.html>

material both relevant to collegiate needs and aligned with various secondary-school curricula.⁴⁹¹

Whereas test-makers used to pose the illegitimacy of test-prep material on psychometric principles—that test-prep courses and guidebooks were illegitimate because they fruitlessly crammed students’ heads with information that could only be learned from years of schooling—they now pose their disdain for test-prep companies *as* a critique of consumption. Or, as the College Board and Khan Academy now quip, “Skills Aren’t Bought.”⁴⁹² Certainly, consumer advocates did not envision an educational environment in which test-makers retained incredible sway over the course of American students’ education, nor one in which socioeconomic biases reflected and reinforced through standardized testing regimens would endure as clearly now as they did in the 1970s. But these tests remain, as do their many biases. Over the past four decades, standardized test manufacturers have shifted the public justification of their exams—from psychometric devices that predicted the likelihood of test-takers’ educational success to products that could help young people make sound educational choices that would secure middle-class comfort. Standardized testing companies took advantage of the blurry relationship consumer advocates posed between the U.S. middle class and public interest, and thrived in a new environment in which disclosure and transparency could be turned into profit. It was this shift that enabled many of the features of standardized testing we take for granted today—and generated, at least for some Americans, an uneasy sense of self when weighing test scores against their ambitions.

⁴⁹¹ Simon, “To Thwart Test Prep Firms.”

⁴⁹² SAT,” *College Board* website, <https://collegereadiness.collegeboard.org/sat?navid=gh-nsat>.

CONCLUSION: STANDARDIZED TESTING REGIMES IN AN UNCERTAIN EDUCATIONAL FUTURE

Standardized testing represented the promises and perils of the postwar liberal state. For those lucky enough to be in the middle class, particularly white men in the middle class, standardized testing was a logical instrument meant to give order to individuals' talents and steer youth into productive academic and vocational pursuits. For those who were not of this privileged sort—if one were a woman, or black, or working class, or neurologically atypical, or any combination thereof—standardized testing was often a marginalizing device, a tool used to provide data that would serve as evidence for existing ideologies and prejudices. Standardized testing has thus always been an oppressive technology for some even as it has helped many others: such is the horror of modern science and technology. What kept the practice afloat for decades, however, was that the group who benefitted most from standardized testing did not feel that the practice, however widespread, breached their sense of the acceptable reach of the state. By acting as a force for business, state, and educational actors to cultivate mutual interests, standardized testing kept all three entities in check—and also largely outside of scrutiny of the white middle class. Hence, by serving the needs of several disparate institutional bases, standardized testing could maintain a broad bipartisan appeal, even as it kept millions of Americans outside of power.

The wholehearted federal embrace of accountability testing programs at the turn of the twenty-first century changed the dynamic between the white middle class and the joint

corporate-state-education standardized testing complex. Although federal consideration of accountability testing had some roots in the Reagan Administration's fear-mongering *A Nation At Risk* and the Clinton Administration's Goals 2000 initiative, the practice didn't really take flight until the G.W. Bush Administration's No Child Left Behind program, which pegged the idea of student success to proficiency models.⁴⁹³ Although No Child Left Behind was designed to ensure students from the most disadvantaged socioeconomic backgrounds and historically neglected communities received the funding and attention necessary to meet national standards in reading and mathematics, the policy more often than not created a glut of standardized testing in America's least-equipped schools. Despite No Child Left Behind's stated goal of ensuring American students would be completely proficient in assessed subjects by 2014, no vested parties ever spelled out what, exactly, proficiency meant, why the terms for proficiency could vary from state to state, or how on earth 100 percent proficiency was even logically possible.⁴⁹⁴

While No Child Left Behind aimed to make educational opportunity more equitable, it often left

⁴⁹³ National Commission on Excellence in Education, *A Nation At Risk: The Imperative for Educational Reform* (Washington, D.C.: U.S. Government Printing Office, 1983). In its report *A Nation at Risk*, the National Commission on Excellence in Education fretted over "a rising tide of mediocrity that threaten[ed] our very future as a Nation and a people" (5), citing the long-term decline of SAT scores a sign of national intellectual decline; the commission recommended a robust system of "nationwide (but not Federal) system of State and local standardized tests" (28) as part of its proposed array of solutions. Despite the commission's reluctance to embrace a federal testing system, it still set a precedent for the desire to coordinate standards-based accountability testing at a national level. Bill Clinton expressed a desire to develop a nationwide testing program as part of Goals 2000, but these efforts ultimately failed. See also: "Clinton Proposes National Exam, Again," FairTest, <http://fairtest.org/clinton-proposes-national-exam-again>; "The Clinton Presidency: Expanding Educational Opportunity," White House website archival pages, <https://clintonwhitehouse5.archives.gov/WH/Accomplishments/eightyears-05.html>; William J. Clinton, "Remarks on Goals 2000 Educational Reform," May 16, 1994, American Presidency Project, <http://www.presidency.ucsb.edu/ws/?pid=50181>; and Michael M. Heise, "Goals 2000: Educate America Act: The Federalization and Legalization of Educational Policy," *Fordham Law Review* 63, no. 2 (1994): 345-381.

⁴⁹⁴ No Child Left Behind Act of 2001 (Public Law 107-110), January 8, 2002; Tom Loveless, "The Peculiar Politics of No Child Left Behind," Brookings Institute, August 2006, https://www.brookings.edu/wp-content/uploads/2016/06/08k12education_loveless.pdf; David J. Hoff, "Not All Agree on Meaning of NCBL Proficiency," *Education Week*, April 17, 2007, <http://www.edweek.org/ew/articles/2007/04/18/33proficient.h26.html>; Thomas S. Dee and Brian A. Jacob, "The Impact of No Child Left Behind on Students, Teachers, and Schools," *Brookings Papers on Economic Activity*, Fall 2010, 149-207; Anya Kamenetz, "It's 2014. All Children Are Supposed to be Proficient. What Happened?" *nprEd*, October 11, 2014, <http://www.npr.org/sections/ed/2014/10/11/354931351/it-s-2014-all-children-are-supposed-to-be-proficient-under-federal-law>;

the most disadvantaged students wading through a sea of standardized tests without many clearly articulated educational goals set in place.

What really made white suburbanites turn against No Child Left Behind, however, were Obama-era policy amendments. Common Core sought to streamline the annual standardized tests students encountered all while giving states greater leeway in coordinate nationwide efforts. These changes put suburban schools under scrutiny in ways that exposed the creative accounting used by many to maintain high ratings. The reaction against federal accountability testing programs reflected the fears among the white middle-class that the standardized testing system was no longer designed to create a social order in which their children were promoted through school and guided into a collegiate path that will replicate their middle-class existence. White parents this felt their children were victimized by the very system they once embraced and the very system designed to benefit them.⁴⁹⁵ This white suburban opposition to Common Core reforms combined with the racial animus at the foundation of American far-right politics. Whatever critiques of standardized testing and accountability systems may have existed—and many valid criticisms did exist—became swept under a larger racist fear in Common Core represented the tyrannical whims of an overbearing and unacceptably black federal government.

Widespread discontent with Common Core, imbued with racist and anti-state fantasies, helped propel Donald Trump into the White House, creating the nauseating educational policy

⁴⁹⁵ Stephanie Simon, “Parents Protest Surge in Standardized Testing,” *Reuters* June 12, 2012, <http://www.reuters.com/article/us-usa-education-testing-idUSBRE85B0EO20120612>; Trevor Tompson, Jennifer Benz, and Jennifer Agiesta, “Parents’ Attitudes on the Quality of Education in the United States,” The Associated Press-NORC Center for Public Affairs Research, August 2013, http://www.apnorc.org/PDFs/Parent%20Attitudes/AP_NORC_Parents%20Attitudes%20on%20the%20Quality%20of%20Education%20in%20the%20US_FINAL_2.pdf; Tim Murphy, “Inside the Mammoth Backlash to Common Core,” *Mother Jones*, September/October 2014, <http://www.motherjones.com/politics/2014/09/common-core-education-reform-backlash-obamacare/>; Sophie Quinton, “Is the Anti-Common Core Movement Just ‘Suburban White Moms’?” *The Atlantic*, June 2, 2015, <https://www.theatlantic.com/politics/archive/2015/06/is-the-anti-common-core-movement-just-suburban-white-moms/432084/>

moment that we now inhabit. In addition to stoking xenophobia and white nationalism, Trump ran for president on the promise of abolishing Common Core, time and again insisting it was federal overreach without either acknowledging that states voluntarily opted into the program or offering any suggestion of what standards-based systems could replace it. Despite no imminent threat to the traditional structure of educational oversight, and despite the Every Student Succeeds Act already rolling back the heavy-handed federal testing policies of No Child Left Behind, the lip service paid to the idea that education must be under local control was enough for Trump's voting base. Several months into the administration, however, the question of how to undo Common Core remains unanswered. Trump's appointee for Secretary of Education, Betsy DeVos—a billionaire who seems unable to comprehend fundamental concepts of contemporary American educational policy such as “proficiency” and “growth”—has instead fashioned her agenda around undermining federal protections for students piece by piece. For DeVos, the entire federal educational apparatus must be refashioned wholesale as a platform for the amorphous and fanciful notion of “choice,” a concept that has largely amounted thus far to rolling back protections for students from minority and marginal communities.⁴⁹⁶ Steering the

⁴⁹⁶ Cory Turner, “Can a President Trump Get Rid of Common Core,” *nprED* November 10, 2016, <http://www.npr.org/sections/ed/2016/11/10/501426803/can-president-trump-get-rid-of-common-core>; Gabby Morrongiello, “DeVos Vows to end Common Core at Michigan Rally with Trump,” *Washington Examiner* December 9, 2016, <http://www.washingtonexaminer.com/devos-vows-to-end-common-core-at-michigan-rally-with-trump/article/2609326>; U.S. Department of Education, “Transitioning to the Every Student Succeeds Act (ESSA): Frequently Asked Questions,” January 18, 2017, <https://www2.ed.gov/policy/elsec/leg/essa/essatransitionfaqs11817.pdf>; Andrew Ujifusa, “Trump Will Repeal Common Core, Says Kellyanne Conway (He Can't),” *Education Week* February 8, 2017, http://blogs.edweek.org/edweek/campaign-k-12/2017/02/trump_will_repeal_common_core_kellyanne_conway.html; Joy Resmovits, “DeVos' New Transgender Student Guidance Changes Course On Bathroom Access Cases,” *Los Angeles Times* June 16, 2017, <http://www.latimes.com/local/education/la-essential-education-updates-southern-this-is-the-trump-administration-s-new-1497632892-htmlstory.html>; Joy Resmovits, “Trump Administration Narrows Civil Rights Investigations,” *Los Angeles Times* June 16, 2017, <http://www.latimes.com/local/education/la-essential-education-updates-southern-trump-s-education-department-pulls-back-1497572246-htmlstory.html>; Susan Berry, “Trump at CEO Business Town Hall: ‘Common Core... We Have to End It,’” *Breitbart* April 4, 2017, <http://www.breitbart.com/big-government/2017/04/04/trump-town-hall-common-core-end-it/>

department away from investigating civil rights abuses doesn't do a thing Common Core, of course—but it does send a clear signal to Trump's fans who the state is supposed to serve.

Last year, Americans found themselves at a fork in the educational road—with neither turn particularly good. On one path was the familiar world, of accountability testing and neoliberal centrist charter school systems, built upon *Waiting for Superman* and Michelle Rhee and Arne Duncan. The other path—the one the United States seems to have chosen for the moment—is far more chaotic, with potentially far fewer federal structures yet with no actual sense that testing will go away. An educational landscape with far fewer standardized tests is possible, but it requires adopting politics committed to investing more heavily in public school funding and teacher salaries. This requires money. It also requires a commitment by Americans themselves not to rely on private solutions where public ones could work. There is something profoundly cowardly in Americans' willingness to write off entire major urban districts to charter school wizardry and Silicon Valley technocratic 'disruption' for the sake of lower taxes in upper income brackets. To this end, standardized testing regimes never truly have to change, so long as those in the white middle class feel that the state is not overstepping its bounds by making its presence felt in their everyday lives. Rather than making choices that would undo the ways the state uses standardized testing to govern our lives, Americans have chosen to undermine the state and leave standardized testing to whoever can pick up the slack. For a question with no right answers, Americans still made the wrong choice.

REFERENCES

Archival Collections

Gerald R. Ford Presidential Library, Ann Arbor, Michigan.

Harry S Truman Presidential Library, Independence, Missouri.

United States National Archives, College Park, Maryland.

Corporate Annual Reports and Financial Statements

Carnegie Corporation of New York. *Annual Report for the Fiscal Year Ended September 30, 1953*. New York: Carnegie Corporation of New York, 1954.

Carnegie Corporation of New York. *Annual Report for the Fiscal Year Ended September 30, 1954*. New York: Carnegie Corporation of New York, 1955.

Educational Testing Service. *Annual Report to the Board of Trustees, 1948-1949*. Princeton: Educational Testing Service, 1949.

Educational Testing Service. *Annual Report to the Board of Trustees, 1949-50*. Princeton: Educational Testing Service, 1950.

Ford Foundation. *The Ford Foundation Report for 1954: January 1, 1954-September 30, 1954*. New York: Ford Foundation, 1955.

Ford Foundation. *Ford Foundation Report: October 1, 1954 to September 30, 1955*. New York: Ford Foundation, 1956.

Graham Holdings. 2015 10K. Graham Holdings Investor Relations.
<http://www.ghco.com/phoenix.zhtml?c=62487&p=irol-reportsannual>.

Graham Holdings. 2015 Annual Report. Graham Holdings Investor Relations.
<http://www.ghco.com/phoenix.zhtml?c=62487&p=irol-reportsannual>.

Graham Holdings. 2016 Annual Report. Graham Holdings Investor Relations.
<http://www.ghco.com/phoenix.zhtml?c=62487&p=irol-reportsannual>.

Washington Post Company. 1999 10-K. U.S. Securities and Exchange Commission EDGAR Database. <http://www.sec.gov/Archives/edgar/data/104889/0000950133-99-000908.txt>

Washington Post Company, 2000 10-K. U.S. Securities and Exchange Commission EDGAR Database. <http://www.sec.gov/Archives/edgar/data/104889/000095013301001017/w46916e10-k.txt>.

Washington Post Company. 2004 10-K. Graham Holdings Investor Relations. <http://www.ghco.com/phoenix.zhtml?c=62487&p=irol-reportsannual>.

Washington Post Company. 2007 10-K; Graham Holdings Investor Relations. <http://www.ghco.com/phoenix.zhtml?c=62487&p=irol-reportsannual>.

Washington Post Company. 2009 10-K. Graham Holdings Investor Relations. <http://www.ghco.com/phoenix.zhtml?c=62487&p=irol-reportsannual>.

Washington Post Company. 2012 10-K. Graham Holdings Investor Relations. <http://www.ghco.com/phoenix.zhtml?c=62487&p=irol-reportsannual>.

Washington Post Company. 1999 Annual Report. Graham Holdings Investor Relations. <http://www.ghco.com/phoenix.zhtml?c=62487&p=irol-reportsannual>.

Washington Post Company. 2009 Annual Report. Graham Holdings Investor Relations. <http://www.ghco.com/phoenix.zhtml?c=62487&p=irol-reportsannual>.

Court Cases

Debra P. v. Turlington, 474 F. Supp. 244 (M.D. Fla. 1979).

Debra P., a Minor, by Irene P., Her Mother and Next Friend, et al., Plaintiffs- Appellants, v. Ralph D. Turlington, Individually and As Commissioner Of Education, et al., Defendants-appellees, 730 F.2d 1405 (11th Cir. 1984).

DeFunis v. Odegaard, 416 U.S. 312 (1974).

Donohue v. Copiague Union Free Sch. Dist., 47 N.Y.2d 440 (1979).

Educational Testing Service v. John Katzman and The Princeton Review, Inc., 626 F. Supp. 527 (1985).

Educational Testing Service v. John Katzman and The Princeton Review, 631 F. Supp. 550 (1986).

Educational Testing Service v. John Katzman, The Princeton Review, Inc., Robert S. Scheller, and Pretest Review, Inc.; John Katzman and The Princeton Review, Inc., Appellants, 793 F.2d 533 (1986).

Federal Trade Commission v. Stanley H. Kaplan Educational Ltd., Et Al, 433 F. Supp. 989 (1977).

Grutter v. Bollinger, 539 U.S. 306 (2003).

Hobson v. Hansen, 269 F. Supp. 401 (D.D.C. 1967).

Larry P. v. Riles, 495 F. Supp 926 (1979).

Marjorie Webster Jr. Col., Inc. v. Middle States Association of Colleges and Secondary Schools, Inc., 302 F. Supp. 459 (D.D.C. 1969).

Marjorie Webster Junior College v. Middle States Association of Colleges and Secondary Schools 432 F.2d 650 (D.C. Cir. 1970).

Parents in Action on Special Education v. Hannon, 506 F.Supp. 831 (1980).

Peter W. v. S.F. Unified Sch. Dist., 131 Cal. Rptr. (Ct. App. 1976).

Regents of Univ. of California v. Bakke, 438 U.S. 265 (1978).

Sharif by Salahuddin v. New York State Educational Department, 709 F. Supp. 345 (S.D.N.Y. 1989).

Executive Orders

U.S. President. Executive Order. "Establishing the War Manpower Commission (Executive Order 9139)." April 18, 1942.

U.S. President. Executive Order. "Amending the Selective Service Regulations (Executive Order 12320)." 31 March, 1951.

Government Publications

Armsby, Henry H. *Engineering, Science, and Management War Training: Final Report*. Washington, D.C.: U.S. Government Printing Office, 1946.

Barlow, Esther. *Abstracts of Personnel Research Reports: VIII. 1954-1968* (AFHRL-TR-68-124). Lackland Air Force Base, Texas: Air Force Systems Command, 1968.

Bush, Vannevar. *Science: The Endless Frontier*. Washington, D.S.: U.S. Government Printing Office, 1945.

- Coleman, James S. et al. *Equality of Educational Opportunity*. Washington, D.C.: U.S. Government Printing Office, 1966.
- De Witt, Nicholas. *Soviet Professional Manpower*. Washington, D.C.: National Science Foundation, 1955.
- Federal Interagency Committee on Education. *Toward a Federal Strategy for Protection of the Consumer of Education*. Washington, D.C.: Department of Health, Education, and Welfare, 1975.
- Federal Security Agency, Office of Education. *Guidance Problems in Wartime--Education and National Defense Series Pamphlet No. 18*. Washington, D.C.: US Government Printing Office, 1942.
- Federal Trade Commission Bureau of Consumer Protection. "Effects of Coaching on Standardized Admission Examinations." Washington, D.C.: Federal Trade Commission, 1979. <http://files.eric.ed.gov/fulltext/ED174645.pdf>.
- McLaughlin, Kenneth F., editor. *Understanding Testing: Purposes and Interpretations for Pupil Development*. Washington, DC: U.S. Government Printing Office, 1960.
- McReynolds, Jane. "Development of motivation keys for the Armed Forces Qualification Test Forms 3 and 4." (AFPTRC-TN-56-60). Lackland Air Force Base, Texas: Personnel Laboratory, Wright Air Development Center, May, 1956.
- National Research Council. *Research—A National Resource, vol. II—Industrial Research*. Washington, D.C.: U.S. Government Printing Office, 1941.
- National Resources Planning Board. *Progress Report, 1940-1941*. Washington, D.C.: U.S. Government Printing Office, 1941.
- National Resources Planning Board. *Our National Resources: Facts and Problems*. Washington, D.C.: U.S. Government Printing Office, 1941.
- National Resources Planning Board. *After Defense—What?* Washington, D.C.: U.S. Government Printing Office, 1942.
- National Resources Planning Board. *Post-War Planning*. Washington, D.C.: U.S. Government Printing Office, 1942.
- National Resources Planning Board. *Demobilization and Readjustment: Report of the Conference on Postwar Readjustment of Civilian and Military Personnel*. Washington, D. C.: U.S. Government Printing Office, 1943.

- National Resources Planning Board. *National Resources Development Report for 1943—Part II: Wartime Planning for War and Post War*. Washington, D.C.: U.S. Government Printing Office, 1943.
- U.S. Army. “Differential Patterns of Item Attempts on the Army General Classification Test Exhibited By Grade IV and V Men Tested at the Reception Center, Fort Leavenworth, Kansas, 1944.” AGO Report No. 519. April 29, 1944. www.dtic.mil/get-tr-doc/pdf?AD=ADB951525.
- U.S. Army. “Interpretations of AGCT Test Scores of Negro and White Selectees,” AGO-PRS Report No. 307[2?]. April 7, 1942. <http://www.dtic.mil/dtic/tr/fulltext/u2/b951522.pdf>.
- U.S. Congress. Senate. Committee on Education and Labor. *Investigation of Manpower Resources: Hearings Before a Subcommittee of the Committee on Education and Labor on S. Res. 291, A Resolution Authorizing the Appointment of a Special Committee to Investigate the Manpower Resources in the United States, part 1* 77th Congress, 2nd Session, October 15-November 20, 1942. Washington, D.C.: U.S. Government Printing Office, 1942.
- U.S. Congress. Senate. Committee on Appropriations. *Hearings Before a Subcommittee of the Committee on Appropriations on Investigation of Manpower, United States Senate, 78th Congress, 1st Session. Report of Senator Theodore Francis Green, a member of the subcommittee.* 78th Congress, 1st Session. Washington, D.C.: U.S. Government Printing Office, 1943.
- U.S. Congress. House. Committee on Military Affairs. *Hearings Before the Committee on Military Affairs on H.R. 2239, H.R. 1742, H.R. 1728 and H.R. 992—Bills Relating to the Full Utilization of Manpower.* 78th Congress, 1st Session, March 25, 26, 30, 31, April 1, 2, 5, 6, 7, 8, 9, 12, 13, 14, 15, 16, 17, 19, 21, 22, 23, May 6, and 7, 1943. Washington, D.C.: U.S. Government Printing Office, 1943.
- U.S. Congress. Senate. Committee of Finance. *Hearings Before a Subcommittee of the Committee on Finance, United States Senate, 78th Congress, 2nd Session, on S. 1617: A Bill to Provide Federal Government Aid For the Readjustment in Civilian Life of Returning World War II Veterans.* 78th Congress, 2nd Session, January 14, 15, 21, 24, February 11, 14, 28, March 8, 10, 1944. Washington, D.C.: US Government Printing Office, 1944.
- U.S. Congress. House. Committee on World War Veterans’ Legislation. *Hearings Before the Committee on World War Veterans’ Legislation, House of Representatives, 78th Congress, 2nd Session, on H.R. 3917 and S. 1767: To Provide Federal Government Aid for the Readjustment in Civilian Life of Returning World War II Veterans.* 78th Congress, 2nd Session, January 11, 12, 13, 17, 18, February 24, and March 9, 10, 27, 28, 29, 30, 31, 1944. Washington, D.C.: US Government Printing Office, 1944.

- U.S. Congress. Senate. Committee on Labor and Public Welfare. *Hearings before United States Senate Committee on Labor and Public Welfare meeting on Science and Education for National Defense*, 85th Congress, 2nd Session, January 21, 1958. Washington, D.C.: US Government Printing Office, 1958.
- U.S. Congress. Senate. Committee on Government Operations. *Hearings Before the Subcommittee on Executive Reorganization and Government Research of the Committee on Government Operations*. 92nd Congress, 1st Session. Washington, D.C.: US Government Printing Office, 1972.
- U.S. Congress. House. Judiciary Committee. *Sex and Race Differences on Standardized Tests: Oversight Hearings before the Subcommittee on Civil and Constitutional Rights of the Committee on the Judiciary*, 100th Congress, 1st Session, April 23, 1987. Washington, U.S. Government Printing Office, 1989.
- U.S. Selective Service System. *Report of the Director of Selective Service for the Fiscal Year 1953 to the Congress of the United States*. Washington, D.C.: U.S. Government Printing Office, 1954.
- U.S. Department of Labor, Office of Policy Planning and Research. *The Negro Family: The Case for National Action*. By Daniel Patrick Moynihan. Washington, D.C.: United States Department of Labor, 1965. <https://www.dol.gov/oasam/programs/history/webid-meynihan.htm>.
- Training Within Industry. *Job Instruction: A Manual for Shop Supervisors and Instructors*. Washington, D.C.: War Production Board, 1942.
- Training Within Industry. *Second Annual Progress Report*. Washington, D.C.: War Production Board, 1942.
- Training Within Industry. *Job Methods Training: Training Sessions Outline and Reference Material*. Washington, D.C.: War Manpower Commission, 1943.
- Training Within Industry. *Training Within Industry Materials: Bulletins Issued by Training Within Industry and Outlines of the Training Within Industry Programs for War Plants and Essential Services*. Washington, D.C.: War Manpower Commission, 1945.
- War Manpower Commission, Bureau of Training. *The Training of Vocational Counselors*. Washington, D.C.: War Manpower Commission, 1944.
- Zeidner, Joseph, and Arthur J. Drucker. *Behavioral Science in the Army: A Corporate History of the Army Research Institute*. Alexandria, VA: United States Army Research Institute for the Behavioral and Social Sciences, 1988.

Laws

Selective Service Act of 1940, U.S. Code 50 (2006), §§ 301 to 318

Servicemen's Readjustment Act of 1944, Public Law 78-436, U.S. Statutes At Large 58 (1944): 283-284.

National Defense Education Act of 1958, Public Law 85-864, U.S. Statutes At Large 72 (1958): 1572-1580.

No Child Left Behind Act of 2001, Public Law 107-110, U.S. Statutes At Large 115 (2001): 1425-2095.

Periodicals and News Services

American Psychologist

Antic

Atlantic

Associated Press Wire

AWM Newsletter

Bioscience

Bloomberg Businessweek

Breitbart

Brookings Institute

Challenge

Change

Change in Higher Education

Christian Science Monitor

Clearing House

Defense Information Bulletin

Denver Post

Education Week
Educational Horizons
Fox News Opinion
Harper's
Hechinger Report
High School Journal
Huffington Post
Inside Higher Ed
Journal of Reading
Journal of Educational Research
Journal of Clinical Psychology
Journal of Psychology
Los Angeles Times
Nation
New York Times
Newsweek
Mother Jones
New Republic
National Public Radio
Pacific Standard
Phi Delta Kappan
Politico
Publishers Weekly

Psychological Bulletin

Psychology Today

Quartz

Review of Educational Research

Rolling Stone

Reuters

Saturday Review

Science

Science of US

Science News-Letter

Scientific American

Tampa Bay Times

Time

USA Today

Wall Street Journal

Washington Examiner

Washington Post

Public Statements and Presented Papers

Flores, Penelope V. "How Dick and Jane Perform Differently in Geometry: Test Results on Reasoning, Visualization, Transformation, Applications, and Coordinates." Paper presented at Annual Meeting of the American Educational Research Association. Boston: April 16-20, 1990. <http://files.eric.ed.gov/fulltext/ED320915.pdf>.

Fordham, Signithia. "Black Students' School Success: Coping with the 'Burden of 'Acting White.'" Paper presented at the Annual Meeting of the American Anthropological Association. Washington, D.C., December 3-7, 1985. <http://files.eric.ed.gov/fulltext/ED281948.pdf>

Jensen, Arthur R. "How Much Can We Boost IQ and Scholastic Achievement?" Speech given before annual meeting of California Advisory Council of Educational Research. San Diego: October, 1967.

Shibley Hyde, Janet. "Biological, Social, and Organizational Components of Success for Women in Academic Science and Engineering." Paper presented at the National Academies Convocation on Maximizing the Success of Women in Science and Engineering. Washington, D.C.: December 9, 2005. <https://www.ncbi.nlm.nih.gov/books/NBK23762/>.

Summers, Lawrence H. "Remarks at NBER Conference on Diversifying the Science & Engineering Workforce." Cambridge, January 14, 2005. http://www.harvard.edu/president/speeches/summers_2005/nber.php.

U.S. Attorney's Office, Western District of Texas. "For-Profit College Kaplan to Refund Federal Financial Aid Under Settlement With United States." Press release. January 5, 2015. <https://www.justice.gov/usao-wdtx/pr/profit-college-kaplan-refund-federal-financial-aid-under-settlement-united-states>

Published Sources—Articles

Acker, Joan. "Inequality Regimes: Gender, Class, and Race in Organization." *Gender and Society* 20, no. 4. August, 2006): 441-464.

Adams, Julia. "The Puzzle of the American State...and Its Historians." *American Historical Review* 115, no. 3. June, 2010): 786-791.

Adams, Michael C.C. "Who Didn't Use the G.I. Bill? Notes on a Lingering Question." *Studies in Popular Culture* 23, no. 2. October, 2000): 65-74.

Adelstein, Richard P. "'The Nation as an Economic Unit': Keynes, Roosevelt, and the Managerial Ideal." *Journal of American History* 78, no. 1. June, 1991): 160-187.

Akard, Patrick J. "Bringing the Economy Back in. Again): Conceptions of the Capitalist State and Their Relevance for Public Policy." *Mid-American Review of Sociology* 11, no. 2. Winter, 1986): 75-100.

Albury, W.R. "Politics and Rhetoric in the Sociobiology Debate." *Social Studies of Science* 10, no. 4. November, 1980): 519-536.

Alderdice, Jacob. "The Informed Student-Consumer: Regulating For-Profit Colleges by Disclosure." *Harvard Civil Rights-Civil Liberties Law Review* 50. 2015): 216-254.

Aldrich, Mark. "The Gender Gap in Earnings During World War II: New Evidence." *ILR Review* 42, no. 3. April, 1989): 415-429.

- Altus, William D. "Personality correlates of Q-L variability on the ACE." *Journal of Consulting Psychology* 16, no. 4. August, 1952): 284-291.
- Anderson, Karen Tucker. "Last Hired, First Fired: Black Women Workers during World War II." *Journal of American History* 69, no. 1. June, 1982): 82-97.
- Andrews, Suzanna. "Elizabeth Warren: A Warrior for the Middle Class." *More* May, 2014. <http://www.more.com/news/womens-issues/elizabeth-warren-warrior-middle-class>.
- Angrist, Joshua, and Alan B. Kruger, "Why do World War II Veterans Earn More than Nonveterans?" *Journal of Labor Economics* 12, no. 1. January, 1994): 74-97.
- Backman, Margaret E. "Patterns of Mental Abilities: Ethnic, Socioeconomic, and Sex Differences." *American Educational Research Journal* 9, no. 1. Winter, 1972): 1-12.
- Baker, Scott. "The Paradoxes of Desegregation: Race, Class, and Education, 1935-1975." *American Journal of Education* 109, no. 3. May, 2001): 320-343.
- Baker, Scott. "Desegregation, Minimum Competency Testing, and the Origins of Accountability: North Carolina and the Nation." *History of Education Quarterly* 55, no. 1. February, 2015): 33-57.
- Bailey, Beth, and David Farber. "The 'Double-V' Campaign in World War II Hawaii: African Americans, Racial Ideology, and Federal Power." *Journal of Social History* 26, no. 4. Summer, 1993): 817-843.
- Beady, Jr., Charles H., and Stephen Hansell. "Teacher Race and Expectations for Student Achievement." *American Educational Research Journal* 18, no. 2. Summer, 1981): 191-206.
- Beidleman, David C. and Christine L. Cole. "Scholastic Aptitude Test Gender Gap." *American Secondary Education* 19, no. 2. 1991): 2-5.
- Beit-Hallahmi, Benjamin. "Science, Ideology, and Ideals: The Social History of IQ Testing." *Centennial Review* 38, no. 2. Spring, 1994.
- Béland, Daniel, François Vergnoille de Chantal, and Sarah-Louise Raillard, "The American State: Between Political Invisibility and Institutional Fragmentation." *Revue Française de Science Politique* 64, no. 2. 2014): 1-14.
- Belasco, Andrew S., Kelly O. Rosinger, and James C. Hearn. "The Test-Optional Movement at America's Selective Liberal Arts Colleges: A Boom for Equity of Something Else?" *Education Evaluation and Policy Analysis* 37, no. 2. June, 2015): 206-223.
- Berezowsky, Sherrin. "Statistical Criticism and the Eminent Man in Francis Galton's *Hereditary Genius*." *Victorian Literature and Culture* 43. 2015): 821-839.

- Boake, Corwin. "From the Binet-Simon to the Wechsler-Bellvue: Tracing the History of Intelligence Testing." *Journal of Clinical and Experimental Neuropsychology* 24, no. 3. 2002): 385-386.
- Bond, Horace Mann. "Talent—and Toilets." *Journal of Negro Education* 28, no. 1. Winter, 1959): 3-14.
- Blum, Lenore. "A Brief History of the Association for Women in Mathematics: The Presidents' Perspectives." Association for Women in Mathematics website. <http://www.awm-math.org/articles/notices/199107/blum/>.
- Bradway, Katherine. "Jung's Psychological Types." *Journal of Analytical Psychology* 9, no. 2. July, 1964): 129-135.
- Brandt, Deborah. "Drafting U.S. Literacy." *College English* 66, no. 5. May, 2004): 485-502.
- Braun, Henry. "Taming Inflation is Never Easy." *Measurement* 13. 2015): 31-34.
- Breen, T.H. "'Baubles of Britain': The American and Consumer Revolutions of the Eighteenth Century." *Past & Present* 119. May, 1988): 73-104.
- Breen, William J. "Social Science and State Policy in World War II: Human Relations, Pedagogy, and Industrial Training, 1940-1945." *Business History Review* 76, no. 2. Summer, 2002): 233-266.
- Brigh, Mary. "'We Cannot Afford to Hurry': Training Within Industry Applied to Nursing." *American Journal of Nursing* 44, no. 3. March, 1994): 223-226.
- Brouillette, J. W. "Brain Power in Education." *Peabody Journal of Education* 23, no. 2. September 1945): 77-78.
- Brown-Collier, Elba K., and Bruce E. Collier. "What Keynes Really Said about Deficit Spending." *Journal of Post Keynesian Economics* 17, no. 3. Spring, 1995): 341-355.
- Bucholtz, Mary. "The Whiteness of Nerds: Superstandard English and Racial Markedness." *Journal of Linguistic Anthropology* 11, no. 1. 2001): 84-100.
- Buffington, Catherine, *et al.* "STEM Training and Early Career Outcomes of Female and Male Graduate Students: Evidence from UMETRICS Data Linked to the 2010 Census." *American Economic Review: Papers & Proceedings* 106, no. 5. 2016): 333-338.
- Bussert-Webb, Kathy. "The Teacher's Testing Panopticon." *Teaching & Learning* 18, no. 3. Summer, 2004): 98-122.

- Cardozo, Mike. "P.S. on Marjorie Webster." *American Bar Association Journal* 57. October, 1971): 1051-1052.
- Carson, John. "Mental Testing in the Early Twentieth Century: Internationalizing the Mental Testing Story." *History of Psychology* 17, no. 3. 2014): 249-255.
- Cech, Erin and Mary Blair-Low, "Perceiving Glass Ceilings? Meritocratic versus Structural Explanations of Gender Inequality among Women in Science and Technology." *Social Problems* 57, no. 3. August, 2010): 371-397.
- Chamberlain, S.D. "The Use of 'Aptitude Tests' in Sales Management." *Journal of Marketing* 8, no. 2. October, 1943): 159-164.
- Cicciola, Elisabetta, Renato Foschi, and Giovanni Pietro Lombardo. "Making Up Intelligence Scales: De Sanctis's and Binet's Tests, 1905 and After." *History of Psychology* 17, no. 3. 2014): 223-236.
- Claussen, Dane S. "A Brief History of Anti-Intellectualism in American Media." 97, no. 3. June, 2011): 8-13.
- Conwill, William L. "The Inheritance of IQ and Scholastic Achievement: Further Comments on the Jensen Article." *Journal of Negro Education* 49, no. 1. Winter, 1980): 97-104.
- Coombs, Jerrold R. "Can Minimum Competency Testing Be Justified?" *High School Journal* 62, no. 4. January, 1979): 175-180.
- Crenshaw, Kimberlé. "Mapping the Margins: Intersectionality, Identity Politics, and Violence Against Women of Color." *Stanford Law Review* 43. July 1991): 1241-1299.
- Crocker, Linda, and Jeri Benson. "Achievement, Guessing, and Risk-Taking Behavior under Norm Referenced and Criterion Referenced Conditions." *American Educational Research Journal* 13, no. 3. Summer, 1976): 207-215.
- Cross, Robert D. "The Historical Development of Anti-Intellectualism in American Society: Implications for the Schooling of African Americans." *Journal of Negro Education* 59, no. 1. Winter, 1990): 19-28.
- Cuordileone, K.A. "'Politics in An Age of Anxiety': Cold War Political Culture and the Crisis of Masculinity, 1949-1960." *Journal of American History* 87, no. 2. September, 2000): 515-545.
- Daniels, Norman. "IQ, Heritability, and Human Nature." *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science*. 1974): 143-180.

- Davey, Gareth, Chuan De Lian, and Louise Higgins. "The University Entrance Examination System in China." *Journal of Further and Higher Education* 31, no. 4. November, 2007): 385-396.
- Davidson, Claire. "Snapshot: Future Homemakers of Tomorrow." Betty Crocker website. <http://www.bettycrocker.com/menus-holidays-parties/mhplibrary/parties-and-get-togethers/vintage-betty/snapshot-future-homemakers-of-tomorrow>.
- deJong-Lambert, William. "Rethinking Little Rock: The Cold War Politics of School Integration in the United States." *European Education* 38, no. 4. Winter 2006-7): 65-81.
- Dee, Thomas S. and Brian A. Jacob, "The Impact of No Child Left Behind on Students, Teachers, and Schools." *Brookings Papers on Economic Activity*. Fall, 2010): 149-207.
- "Deliver Us From the Racial Hypocrisy of the National Merit Scholarship Program." *Journal of Blacks in Higher Education* 39,. Spring, 2003): 61-63.
- Denton, Georgina. "Neither Guns Nor Bombs—Neither the State Nor God—Will Stop us From Fighting For Our Children: Motherhood and Protest in 1960s and 1970s America." *The Sixties: A Journal of History, Politics, and Culture* 5, no. 2. November, 2012): 205-228.
- Dowd, Nancy E. "Bakke and Weber: The Concept of Societal Discrimination." *Loyola University Law Journal* 11, no. 2. Winter, 1980): 297-326.
- Epstein, Ann S. *et al.* "Preschool Assessment: A Guide to Developing a Balanced Approach." *Preschool Policy Matters* 7. July, 2004. <http://www.doe.in.gov/sites/default/files/earlylearning/preschool-assessment-guide-developing-balanced-approach.pdf>
- Fass, Paula S. "The IQ: A Cultural and Historical Framework." *American Journal of Education* 88, no. 4. August, 1980): 431-458.
- Feng, Yuan. "National College Entrance Examinations: the Dynamics of Political Centralism in China's Elite Education." *Journal of Education* 181, no. 1. 1999): 39-57.
- Fennema, Elizabeth. "Sex Differences in Mathematics Learning: Why?" *The Elementary School Journal* 75, no. 3. December, 1974): 183-190.
- Fennema, Elizabeth. "Mathematics Learning and the Sexes: A review." *Journal for Research in Mathematics Education* 5, no. 3. December, 1974): 304-309.
- Fennema, Elizabeth and Julia A. Sherman. "Fennema-Sherman Mathematics Scale: Instruments Designed to Measure Attitudes toward the Learning of Mathematics by Females and Males." *Journal for Research in Mathematics Education* 7, no. 5. November, 1976): 324-326.

- Fennema, Elizabeth and Julia A. Sherman. "Sexual stereotyping and Mathematics Learning." *The Arithmetic Teacher* 24, no. 5. May, 1977): 369-372.
- Fennema, Elizabeth and Julia A. Sherman. "Sex-Related Differences in Mathematics Achievement, Spatial Visualization and Affective Factors." *American Educational Research Journal* 14, no. 1. Winter, 1977): 51-71.
- Fennema, Elizabeth H. and Julia A. Sherman. "Sex Differences in Mathematics Achievement and Related Factors: A Further Study." *Journal for Research in Mathematics Education* 9, no. 3. May, 1978): 189-203.
- Finnigan, Katherine M. and Katherine S. Corker. "Do Performance Avoidance Goals Moderate the Effect of Different Types of Stereotype Threat on Women's Math Performance?" *Journal of Research in Personality* 63. 2016): 36-43.
- Friedman, Lynn. "Mathematics and the Gender Gap: A Meta-Analysis of Recent Studies on Sex Differences in Mathematical Tasks." *Review of Educational Research* 59, no. 2. Summer, 1989): 185-213.
- Gaddis, S. Michael, and Douglas Lee Lauen. "School Accountability and the Black-White Test Score Gap." *Social Science Research* 44. 2014): 15-31.
- Garcia, D. Linda. "Standard Setting in the United States: Public and Private Sector Roles." *Journal of the American Society for Information Science*. 1986-1998 series), 43, no. 8. September, 1992): 531-537.
- Garcia, Matthew. "Labor, Migration, and Social Justice in the Age of the Grape Boycott." *Gastronomica* 7, no. 3. Summer, 2007): 68-74.
- Gifford, Bernard R. "The Political Economy of Testing and Opportunity Allocation." *The Journal of Negro Education* 59, no. 1. Winter, 1990): 58-69.
- Gilens, Martin. "'Race Coding' and White Opposition to Welfare." *American Political Science Review* 90, no. 3. September, 1996): 593-604.
- Gillham, Nicholas W. "The Battle Between the Biometricians and the Mendelians: How Sir Francis Galton's Work Caused his Disciples to Reach Conflicting Conclusions About the Hereditary Mechanism." *Science and Education* 24. 2015): 62-64.
- Glauber, Rebecca. "Marriage and the Motherhood Wage Penalty among African Americans, Hispanics, and Whites." *Journal of Marriage and Family* 69, no. 4. November, 2007): 951-961.
- Glickman, Lawrence B. "The Strike in the Temple of Consumption: Consumer Activism and Twentieth-Century American Political Culture." *Journal of American History* 88, no. 1. June, 2001): 99-128.

- Gomberg, Paul. "IQ and Race: A Discussion of Some Confusions." *Ethics* 85, no. 3. April, 1975): 258-266.
- Gordon, Margaret T. "A Different View of the IQ-Achievement Gap." *Sociology of Education* 49, no. 1. January, 1976): 4-11.
- Grissom, Jason A., and Christopher Redding. "Discretion and Disproportionality: Explaining the Underrepresentation of High- Achieving Students of Color in Gifted Programs." *AERA Open* 2, no. 1. January-March 2016): 1-25.
- Hacking, Ian. "Biopower and the Avalanche of Printed Numbers." *Humanities in Society* 5. 1982): 279-95.
- Hacking, Ian. "Making Up People." *London Review of Books*, 28, no. 16. August 17, 2006. <https://www.lrb.co.uk/v28/n16/ian-hacking/making-up-people>
- Haddon, Phoebe A. and Deborah W. Post. "Misuse and Abuse of the LSAT: Making the Case for Alternative Evaluative Efforts and a Redefinition of Merit." *St. Johns Law Review* 80. 2006): 41-105.
- Halpern, Diane F., et al. "The Science of Sex Differences in Science and Mathematics." *Psychological Science in the Public Interest* 8, no. 1. August, 2007): 1-51.
- Hansen, G. Eric. "Intellect and Power: Some Notes on the Intellectual as a Political Type." *Journal of Politics* 31, no. 2. May, 1969): 311-328.
- Haney, Walt. "Testing Reasoning and Reasoning About Testing." *Review of Educational Research* 54, no. 4. Winter, 1984): 597-654.
- Hannum, Emily, Xuehui An, and Hua-Yu Sebastian Cherng. "Examinations and Educational Opportunity in China: Mobility and Bottlenecks for the Rural Poor." *Oxford Review of Education* 37, no. 2. April, 2011): 267-305.
- Hansen, Michael. "Teaching to the Test: Hype or Reality?" Brown Center Chalkboard blog, Brookings Institute. <https://www.brookings.edu/blog/brown-center-chalkboard/2016/05/16/teaching-to-the-test-hype-or-reality/>.
- Hayman, Jr., Robert L. "Presumptions of Justice: Law, Politics, and the Mentally Retarded Parent." *Harvard Law Review* 103, no. 6. April, 1990): 1201-1271.
- Heise, Michael M. "Goals 2000: Educate America Act: The Federalization and Legalization of Educational Policy." *Fordham Law Review* 63, no. 2. 1994): 345-381.
- Hendrickson, Robert M. "Rethinking Affirmative Action: Redefining Compelling State Interest and Merit in Admission." *Peabody Journal of Education* 76, no. 1. 2001): 117-135.

- Herzer, Frank, Julia Wendt, and Alfons O. Hamm. "Discriminating Clinical from Nonclinical Manifestations of Test Anxiety: A Validation Study." *Behavior Therapy* 45. 2014): 222-231.
- Hilton, Matthew. "Social Activism in an Age of Consumption: The Organized Consumer Movement." *Social History* 32, no. 2. May, 2007): 121-143.
- Hilton, Matthew. "The Death of a Consumer Society." *Transactions of the Royal Historical Society* ser. 6, vol. 18. 2008): 211-236.
- Horn, Joseph M. "Truth, Gender, and the SAT." *Academic Questions* 3. 1989-90): 35-39.
- Hochschild, Jennifer L. "The Strange Career of Affirmative Action." *Ohio State Law Journal* 59, no. 3. 1998): 997-1037.
- Hoskin, Keith. "The Examination, Disciplinary Power and Rational Schooling." *History of Education* 8, no. 2. 1979): 135-146.
- Hoskin, Keith W. and Richard H. Macve. "Accounting and the Examination: A Genealogy of Disciplinary Power." *Accounting, Organizations and Society* 11, no. 2. 1986): 105-136.
- Hoyt, Kendall. "Vaccine Innovation: Lessons from World War II." *Journal of Public Health Policy* 27, no. 1. 2006): 38-57.
- Humes, Edward. "How the GI Bill Shunted Blacks into Vocational Training." *The Journal of Blacks in Higher Education* no. 53. Autumn 2006): 92-104.
- Hutt, Ethan and Aaron Tang. "The New Education Malpractice Litigation." *Virginia Law Review* 99, no. 3. May, 2013): 428-440.
- Isaac, Joel. "The Human Sciences in Cold War America." *The Historical Journal* 50, no. 3. September, 2007): 725-746.
- Jacob, Brian A. "Accountability, Incentives, and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools." *Journal of Public Economics* 89. 2005): 761-796.
- Jacobs, Meg. "How About Some Meat?: The Office of Price Administration, Consumption Politics and State Building from the Bottom Up, 1941-1946." *Journal of American History* 84, no. 3. December, 1997): 910-941.
- Jacobs, Meg. "State of the Field: the Politics of Consumption." *Reviews in American History* 39, no. 3. September, 2011): 561-573.

- Jennings, Jennifer L. and Jonathan Marc Bereak. "‘Teaching to the Test’ in the NCLB Era: How Test Predictability Affects Our Understanding of Student Performance." *Educational Researcher* 43, no. 8: 381-389.
- Jensen, Laura S. "Government, the State, and Governance." *Polity* 40, no. 3. July, 2008): 379-385.
- Jones, Jeffrey M. "Race, Ideology, and Support for Affirmative Action." Gallup website, August 23, 2005. <http://www.gallup.com/poll/18091/race-ideology-support-affirmative-action.aspx>.
- Jones, Leslie Ann. "Interview with a One-Time ‘Betty Crocker Homemaker of Tomorrow.’" *the hairpin*, June 13, 2013. <http://thehairpin.com/2013/06/interview-with-a-one-time-betty-crocker-homemaker-of-tomorrow/>.
- Kafka, Judith. "Shifting Authority: Teachers’ Role in the Bureaucratization of School Discipline in Postwar Los Angeles." *History of Education Quarterly* 49, no. 3. August, 2009): 323-346.
- Kahn, Joan R., Javier García-Manglano and Suzanne M. Bianchi. "The Motherhood Penalty at Midlife: Long-Term Effects of Children on Women’s Careers." *Journal of Marriage and Family* 76. February, 2014): 56-72.
- Kane, Johnathan M. and Janet E. Mertz. "Debunking Myths About Gender and Math Performance." *Notices of the AMS* 59, no. 1. January, 2012): 10-21.
- Kane, Thomas J. and Douglas O. Staiger, "The Promise and Pitfalls of Using Imprecise School Accountability Measures." *Journal of Economic Perspectives* 16, no. 4. Autumn, 2002): 91-114.
- Katznelson, Ira, and Suzanne Mettler. "On Race and Policy History: A Dialogue About the G.I. Bill." *Perspectives on Politics* 6, no. 3. September, 2008): 519-537.
- Kendi, Ibram X. "Why the Academic Achievement Gap is a Racist Idea." *Black Perspectives*, October 20, 2016, <http://www.aaihs.org/why-the-academic-achievement-gap-is-a-racist-idea/>.
- Kidder, William C. "Does the LSAT Mirror or Magnify Racial and Ethnic Differences in Educational Attainment? A Study of Equally Achieving ‘Elite’ College Students." *California Law Review* 89, no. 4. July, 2001): 1055-1124.
- King, Desmond, and Robert C. Lieberman. "Finding the American State: Transcending the ‘Statelessness’ Account." *Polity* 40, no. 3. July, 2008): 368-378.
- King, Desmond, and Marc Stears. "How the U.S. State Works: A Theory of Standardization." *Perspectives on Politics* 9, no. 3. September, 2011): 505-518.

- Kimmel, Michael S. "Masculinity as Homophobia: Fear, Shame, and Silence in the Construction of Gender Identity." In *Theorizing Masculinities*, edited by Harry Brod and Michael Kaufman. New York: Sage, 1991): 119-141.
- Kossoudji, Sherrie A., and Laura J. Dresser. "Working Class Rosies: Women Industrial Workers during World War II." *Journal of Economic History* 52, no. 2. June, 1992): 431-446.
- Krieger, Nancy. "Who and What is a 'Population'? Historical Debates, Current Controversies and Implications for Understanding 'Population Health' and Rectifying Health Inequities." *Milbank Quarterly* 90, no. 4. December, 2012): 634-681.
- Kunert, Joachim, Astrid Montag, and Sigrid Pöhlmann. "The Quincunx: History and Mathematics." *Statistical Papers* 42. 2001): 143-169.
- Lambert, Caroline and Eric Pezet. "Accounting and the Making of *Homo Liberalis*." *Foucault Studies* 13. May, 2012): 67-81.
- Lambert, Nadine M. "Psychological Evidence in *Larry P. v. Wilson Riles*: Evaluation by a Witness for the Defense." *American Psychologist* 36, no. 9. September, 1981): 937-952.
- Lambert, Philip. "Mathematical Ability and Masculinity." *The Arithmetic Teacher* 7, no. 1. January, 1960) 19-21.
- Lambert Zarrel V. and Fred W. Kniffin. "Consumer Discontent: A Social Perspective." *California Management Review* 18, no. 1. Fall, 1975): 36-44.
- Lamoree, Elizabeth. "Gambling on Grapes: Management, Marketing, and Labor in California Agribusiness." *Agricultural History* 86, no. 3. Summer, 2012): 104-127.
- Lampland, Martha. "False Numbers as Formalizing Practices." *Social Studies of Science* 40, no. 3. June, 2010): 377-404.
- Lawler, James. "IQ: Biological Fact or Methodological Construct?" *Science & Society* 41, no. 2. Summer, 1977): 208-218.
- Lazear, Edward P. "Speeding, Terrorism, and Teaching to the Test." *Quarterly Journal of Economics* 121, no. 3. August, 2006): 1029-1061.
- Lecklider, Aaron S. "Inventing the Egghead: The Paradoxes of Brainpower in Cold War American Culture." *Journal of American Studies* 45, no. 2. May, 2011): 245-265.
- Leighton, Richard J. "Consumer Protection Agency Proposals: The Origin of the Species." *Administrative Law Review* 25, no. 3. Summer, 1973) 269-312.

- Leighton, Richard J. "The Consumer Advocacy Agency Proposal...Again." *Administrative Law Review* 27, no. 2. Spring, 1975): 149-164.
- Levine, Arthur, and Keith R. Wilson. "Student Activism in the 1970s: Transformation Not Decline." *Higher Education* 8, no. 6. November, 1979): 627-640.
- Liberman, Robert C. "Finding the American State: Transcending the 'Statelessness' Account." *Polity* 40, no. 3. July, 2008): 368-378.
- Light, Jennifer S. "When Computers Were Women" *Technology and Culture* 40, no. 3. July, 1999): 455-483.
- Linn, Robert L. "Assessments and Accountability." *Educational Researcher* 29, no. 2. March, 2000): 4-16. <http://www.education.umd.edu/EDMS/MARCES/mdarch/pdf/1000052.pdf>.
- Lipman-Blumen, Jean, and Ann R. Tickamyer. "Sex Roles in Transition: A Ten-Year Perspective." *Annual Review of Sociology* 1. 1975): 297-337.
- Liu, Goodwin. "The Causation Fallacy: *Bakke* and the Basic Arithmetic of Selective Admissions." *Michigan Law Review* 100, no. 10. March, 2002): 1045-1107.
- Looney, Adam and Constantine Yannelis. "A Crisis in Student Loans?: How Changes in the Characteristics of Borrowers and in the Institutions They Attended Contributed to Rising Loan Defaults." *Brookings Papers on Economic Activity* Fall, 2015: 1-89. <https://doi.org/10.1353/eca.2015.0003>.
- Loss, Christopher P. "'The Most Wonderful Thing Has Happened to Me in the Army': Psychology, Citizenship, and American Higher Education in World War II." *The Journal of American History* 92, no. 3. December, 2005): 864-891.
- Lufi, Dubi, and Lina Darliuk. "The Interactive Effect of Test Anxiety and Learning Disabilities Among Adolescents." *Educational Research* 43. 2005): 236-249.
- Lunz, Mary E. and Betty A. Bergstrom. "The Effect of Review on Student Ability and Test Efficiency for Computerized Adaptive Tests." *Applied Psychological Measurement* 16, no. 1. March, 1992): 33-40.
- Madaus, George F. and Thomas Kellaghan. "Testing as a Mechanism of Public Policy: A Brief History and Description." *Measurement & Evaluation in Counseling & Development* 26, no. 1. April, 1993.
- Marks, Susan. "Betty Crocker Search for the American Housewife of Tomorrow Scholarship." *Finding Betty Crocker* blog. October 26, 2008. <http://findingbettycrocker.blogspot.com/2008/10/betty-crocker-search-for-american.html>;

- Martin, Karin A. "Becoming a Gendered Body: Practices of Preschools." *American Sociological Review* 63 no. 4. August, 1998): 494-511.
- Mason, James Hocker. "The Educational Milieu 1874-1911: College Entrance Requirements and the Shaping of Education." *The English Journal* 68, no. 4. April, 1979): 40-45.
- Matross, Ronald P. "A Profile of Public Interest Research Groups." *University of Minnesota Office of Student Affairs Research Bulletin* 17, no. 9. February 10, 1977): 10-13.
- Mccauley, Mary H. "The Myers-Briggs Type Indicator: A Measure for Individuals and Groups." *Measurement and Evaluation in Counseling and Development* 22, no. 4. January, 1990): 181-195.
- Merriam, Charles E. "The National Resources Planning Board: A Chapter in American Planning Experience." *American Political Science Review* 38, no. 6. December, 1944): 1075-1088.
- Meyer, John W., David Tyack, Joane Nagel, and Audri Gordon. "Public Education as Nation-Building in America: Enrollments and Bureaucratization in the American States, 1870-1930." *American Journal of Sociology* 85, no 3. November, 1979): 591-613.
- Mayer, Robert N. "The Socially Conscious Consumer—Another Look at the Data." *Journal of Consumer Research* 3, no. 2. September, 1976): 113-115.
- McGuire, Phillip. "Desegregation of the Armed Forces: Black Leadership, Protest, and World War II, *Journal of Negro History* 68, no. 2. Spring, 1983): 147-158.
- McClelland, David C. "Encouraging Excellence." *Daedalus* 90, no. 4. Fall, 1961): 711-724.
- Mennicken, Andrea, and Peter Miller. "Accounting, Territorialization and Power." *Foucault Studies* 13. May, 2012): 4-24.
- Metzger, Gillian E. "Agencies, Polarization, and the States." *Columbia Law Review* 115, no. 7. November, 2015): 1739-1787.
- Milkman, Roger. "A Simple Exposition of Jensen's Error." *Journal of Educational Statistics* 3, no. 3. Autumn, 1978): 203-208.
- Milkman, Ruth. "Redefining 'Women's Work': The Sexual Division of Labor in the Auto Industry during World War II." *Feminist Studies* 8, no. 2. Summer, 1982): 336-372.
- Miller, Peter and Ted O'Leary. "Accounting and the Construction of the Governable Person." *Accounting, Organizations and Society* 12, no. 3. 1987): 235-265.
- Millman, Jason, Carol H. Bishop, and Robert Ebel. "An Analysis of Test-Wiseness." *Educational and Psychological Measurement* 25, no. 3. 1965): 707-726.

- Moos, Grant. "Remembering Betty Crocker Homemakers of Tomorrow." *Taste of General Mills* blog. General Mills website. April 9, 2013.
<http://blog.generalmills.com/2013/04/remembering-betty-crocker-homemakers-of-tomorrow/>.
- Morning, Ann. "Everyone Knows It's a Social Construct": Contemporary Science and the Nature of Race." *Sociological Focus* 40, no. 4. November, 2007): 436-454.
- Moss, Kary L. "Standardized Tests as a Tool of Exclusion: Improper Use of the SAT in New York." *Berkeley Journal of Gender, Law & Justice* 4, no. 2. 1989): 230-244.
<http://scholarship.law.berkeley.edu/cgi/viewcontent.cgi?article=1027&context=bglj>.
- Mülberger, Annette. "The Need for Contextual Approaches to the History of Mental Testing." *History of Psychology* 17, no. 3. 2014): 177-186.
- Muncy, Robyn. "Cooperative Motherhood and Democratic Civic Culture in Postwar Suburbia, 1940-1965." *Journal of Social History* 38, no. 2. Winter, 2004): 285-310.
- Murray, Carolyn Bennett, and James S. Jackson. "The Conditioned Failure Model of Black Educational Underachievement." *Humboldt Journal of Social Relations* 10, no. 1. Fall/Winter 1982-1983): 276-300.
- Myers, Katherine D. "An Extended History of the Myers-Briggs type Indicator Instrument." *MBTI Type Today* website. 2011. <http://www.mbtitoday.org/downloads/An-Extended-History-of-the-MBTI.pdf>;
- n.a. "Antitrust Laws: College Accrediting Association's Refusal to Evaluate Proprietary College Not a Sherman Act Violation." *Harvard Law Review* 84, no. 8. June, 1971): 1912-1920.
- n.a. "Public Backs Affirmative Action, But Not Minority Preferences." Pew Research Center website. June 2, 2009. <http://www.pewresearch.org/2009/06/02/public-backs-affirmative-action-but-not-minority-preferences/>.
- n.a. "NMSC History and Facts." National Merit Scholarship Corporation website.
<http://www.nationalmerit.org/s/1758/interior.aspx?sid=1758&gid=2&pgid=451>.
- Nam, Charles B. "Impact of the 'GI Bills' on the Educational Level of the Male Population." *Social Forces* 43, no. 1. October, 1964): 26-32.
- Nankervis, Bryan. "Gender Inequity in the National Merit Scholarship Program" *Journal of College Admission* 220. Spring, 2013): 20-25.
- Neal, Derek, and Diane Whitmore Schanzenbach. "Left Behind by Design: Proficiency Counts and Test-Based Accountability." *Review of Economics and Statistics* 92, no. 2. May, 2010): 263-283.

- Nettl, J.P. "The State as a Conceptual Variable." *World Politics* 20, no. 4. July, 1968): 559-592.
- Niederle, Muriel and Lise Vesterlund. "Explaining the Gender Gap in Math Test Scores: The Role of Competition." *Journal of Economic Perspectives* 24, no. 2. Spring, 2010): 129-144.
- Nicolas, Serge, *et. al.* "Sick? Or Slow? On the origins of intelligence as a Psychological Object." *Intelligence* 41. 2013): 699-711.
- Nitko, Anthony J. "Distinguishing the Many Varieties of Criterion-References Tests." *Review of Educational Research* 50, no. 3. Autumn, 1980): 461-485.
- Novak, William J. "The Myth of the 'Weak' American State." *American Historical Review* 113, no. 3. June, 2008): 752-772.
- Ohanian, Lee E. "The Macroeconomic Effects of War Finance in the United States: World War II and the Korean War." *American Economic Review* 87, no. 1. March, 1997): 23-40.
- Olson, Keith W. "The G.I. Bill and Higher Education: Success and Surprise." *American Quarterly* 25, no. 5. December, 1973): 596-610.
- Onkst, David H. "'First a Negro...Incidentally a Veteran': Black World War Two Veterans and the G.I. Bill of Rights in the Deep South, 1944-1948." *Journal of Social History* 31, no. 3. Spring, 1998): 517-543.
- Pallas, Aaron M. and Karl L. Alexander. "Sex Differences in Quantitative SAT Performance: New Evidence on the Differential Course Hypothesis." *American Educational Research Journal* 20, no. 2. Summer, 1983): 165-182.
- Patashnik, Eric M., and Julian E. Zelizer. "The Struggle to Remake Politics: Liberal Reform and the Limits of Policy Feedback in the Contemporary American State." *Perspectives on Politics* 11, no. 4. December, 2013): 1071-1087.
- Perez-Felkner, Lara, Samantha Nix, and Kirby Thomas. "Gendered Pathways: How Mathematics Ability Beliefs Shape Secondary and Postsecondary Course and Degree Choices." *Frontiers in Psychology*, April 06, 2017. <https://doi.org/10.3389/fpsyg.2017.00386>.
- Persson-Benbow, Camilla, and Julian C. Stanley. "Sex Differences in Mathematical Ability: Fact or Artifact?" *Science* 210, no. 4475. December 12, 1980): 1262-1264.
- Persson Benbow, Camilla and Julian Stanley. "Consequences in High School and College of Sex Differences in Mathematical Ability: A Longitudinal Perspective." *American Educational Research Journal* 19, no. 4. Winter 1982): 598-622.
- Persson Benbow, Camilla and Julian Stanley. "Sex Differences in Mathematical Reasoning Ability: More Facts." *Science* 222, no. 4627. December 2, 1983): 1029-1031.

- Persson Benbow, Camilla and Julian Stanley. "Differential Course-Taking Hypothesis Revisited." *American Educational Research Journal* 20, no. 4. Winter, 1983): 469-473.
- Persson Benbow, Camilla. Reply to Jonathan Beckwith and Michael Woodruff, in "Advancement in Mathematics." *Science* 223, no. 4642. March 23, 1984): 1248.
- Persson Benbow, Camilla and Lola Minor. "Mathematically Talented Males and Females and Achievement in the High School Sciences." *American Educational Research Journal* 23, no. 3. Autumn, 1986): 425-436.
- Peters, Frank R. and Eugenia L. Plog. "The Effectiveness of the ACT for Selection and Placement at the Ohio State University." *Educational Research Bulletin* 40, no. 9. December 13, 1961): 232-241, 252.
- Phillips, Christopher J. "The New Math and Midcentury American Politics." *Journal of American History* 101, no. 2. September, 2014): 454-479.
- Pitkin, R.B. "How the First GI Bill Was Written". part 1 of 2. *American Legion Magazine* January, 1969: 24-28, 51-53.
- Pitkin, R.B. "How the First GI Bill Was Written.". part 2 of 2. *American Legion Magazine* February, 1969: 22-26, 48-51.
- Popham, James, and T.R. Husek. "Implications of Criterion-References Measurement." *Journal of Educational Measurement* 6, no. 1. Spring, 1969): 1-9.
- Powell, Marvin, Henry A. O'Connor, and Murray Deutsch. "Are There Really Sex Differences in Achievement?" *Journal of Educational Research* 57, no. 4. December, 1963): 210-212.
- Powers, Patricia. "Social Change: Nader Style." *Journal of Education and Social Work* 13, no. 3. Fall 1977): 63-69.
- Ramo, Simon. "Space or Electronics—which Will Dominate the Century?" in *Proceedings of the National Electronics Conference*, vol. XIV. Chicago: National Electronics Conference, Inc., 1959.
- Reed, Jr., Adolph. "Marx, Race, and Neoliberalism." *New Labor Forum* 22, no. 1. Winter, 2013): 48-57.
- Renshaw, Patrick. "Organized Labour and the United States War Economy, 1939-1945." *Journal of Contemporary History* 21, no. 1. January, 1986): 3-22.
- Renshaw, Patrick. "Was There a Keynesian Economy in the USA between 1933 and 1945?" *Journal of Contemporary History* 34, no. 3. July, 1999): 337-364.

- Renwick, Chris. "From Political Economy to Sociology: Francis Galton and the Social-Scientific Origins of Eugenics." *British Society for the History of Science* 44, no. 3. September, 2011): 343-369.
- Resnick, Daniel P. "Minimum Competency Testing Historically Considered." *Review of Research in Education*, 8. 1980): 3-29.
- Richardson, J.T. "Accountability Incentives and Academic Achievement: Distributional Impacts of Accountability When Standards are Set Low." *Economics of Education Review* 44. February, 2015): 1-16.
- Rockwell, Landon G. "The Planning Function of the National Resources Planning Board." *Journal of Politics* 7, no. 2. May, 1945): 169-178.
- Rossiter, Margaret W. "The Matthew Matilda Effect in Science." *Social Studies of Science* 23, no. 2. May, 1993): 325-341.
- Rury, John L. "Race, Region, and Education: An Analysis of Black and White Scores on the 1917 Army Alpha Intelligence Test." *The Journal of Negro Education* 57, no. 1. Winter, 1988): 51-65.
- Rushton, J. Philippe, and Arthur R. Jensen. "The Totality of Available Evidence Shows the Race IQ Gap Still Remains." *Psychological Science* 17, no. 1. October, 2006): 921-922.
- Ryan, Patrick J. "Unnatural Selection: Intelligence Testing, Eugenics, and American Political Cultures." *Journal of Social History* 30, no. 3. Spring, 1997): 669-685.
- Savage, David G. "The Long Decline in SAT Scores." *Educational Leadership* January, 1978: 290-293. http://www.ascd.org/ASCD/pdf/journals/ed_lead/el_197801_savage.pdf.
- Schiebinger, Londa. "The History and Philosophy of Women in Science: A Review Essay." *Signs* 12, no. 2. Winter, 1987): 305-332.
- Schiebinger, Londa. "Has Feminism Changed Science?" *Signs* 25, no. 4. Summer, 2000): 1171-1175.
- Schiller, Rebel E. "The Administrative State, Front and Center: Studying Law and Administration in Postwar America." *Law and History Review* 26, no. 2. Summer, 2008): 415-427.
- Schudson, Michael S. "Organizing the 'Meritocracy': A History of the College Entrance Examination Board." *Harvard Educational Review* 42, no. 1. February, 1972): 32-69.
- Shultz, Irvin T., and Bentley Barnabas. "Testing for Leadership in Industry." *Transactions of the Kansas Academy of Science* 48, no. 2. September, 1945): 160-164.

- Schwery, Denise, David Hulac, and Amy Schweinle. "Understanding the Gender Gap in Mathematics Achievement: The Role of Self-Efficacy and Stereotype Threat." *School Psychology Forum* 10, no. 4. Winter, 2016): 386-396.
- Seeman, Melvin. "On the Meaning of Alienation." *American Sociological Review* 24, no. 6. December, 1959): 783-791.
- Sen, Sankar, Zeynep Gürhan-Canli and Vicki Morwitz. "Withholding Consumption: A Social Dilemma Perspective on Consumer Boycotts." *Journal of Consumer Research* 28, no. 3. December, 2001): 399-417.
- Sesardic, Neven. "Philosophy of Science That Ignores Science: Race, IQ and Heritability." *Philosophy of Science* 67, no. 4. December, 2000): 580-602.
- Scroop, Daniel. "The Anti-Chain Store Movement and the Politics of Consumption." *American Quarterly* 60, no. 4. December, 2008): 925-949.
- Sheehan, Kate R. and Mary W. Gray. "Sex Bias in the SAT and the DTMA." *Journal of General Psychology* 119, no. 1. January, 1992): 5-14.
- Sherman, Julia, and Elizabeth Fennema. "The Study of Mathematics by High School Girls and Boys: Related Variables." *American Educational Research Journal* 14, no. 2. Spring, 1977): 159-168.
- Shepard, Lorrie. "Norm-References vs. Criterion-Referenced Tests." *Educational Horizons* 58, no. 1. Fall, 1979): 26-32.
- Shepard, Lorrie. "Issues in Minimum Competency Testing." *Review of Research in Education* 8. 1980): 30-82.
- Shockley, Megan Taylor. "Working for Democracy: Working-Class African-American Women, Citizenship, and Civil Rights in Detroit, 1940-1954." *Michigan Historical Review* 29, no. 2. Fall, 2003): 125-157.
- Shuit, Douglas P. "At 60, Myers-Briggs is Still Sorting Out and Identifying People's Types." *Workforce Management*. December, 2003): 72-74.
- Skocpol, Theda. "A Society Without a 'State'? Political Organization, Social Conflict, and Welfare Provision in the United States." *Journal of Public Policy* 7, no. 4. October-December, 1987): 349-371.
- Slayton, Rebecca. "From a 'Dead Albatross' to Lincoln Labs: Applied Research and the Making of a Normal Cold War University." *Historical Studies in the Natural Sciences* 42, no. 4. September, 2012): 255-282.

- Sommer, Markus and Martin E. Arendasy. "Comparing Different Explanations of the Effect of Test Anxiety on Respondents' Test Scores." *Intelligence* 42. 2014): 115-127.
- Speer, Jamin D. "Pre-Market Skills, Occupational Choice, and Career Progression." *Journal of Human Resources* 52, no. 1. January, 2017): 187-246.
- Stafford, Richard E. "Hereditary and Environmental Components of Quantitative Reasoning." *Review of Educational Research* 42, no. 2. Spring, 1972): 183-201.
- Stanley, Marcus. "College Education and the Midcentury GI Bills." *The Quarterly Journal of Economics* 118, no. 2. May 2003): 671-708.
- Star, Susan Leigh, and James R. Griesemer. "Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Berkeley's Museum of Vertebrate Zoology, 1907-39." *Social Studies of Science* 19, no. 3. August, 1989): 387-420.
- Steeh, Charlotte, and Maria Krysan. "Trends: Affirmative Action and the Public, 1970-1995." *Public Opinion Quarterly* 60, no. 1. Spring, 1996): 128-158.
- Stein, Nancy. "Affirmative Action and the Persistence of Racism." *Social Justice* 22, no. 3. Fall, 1995): 28-44.
- Stubblefield, Anna. "'Beyond the Pale': Tainted Whiteness, Cognitive Disability, and Eugenic Sterilization." *Hypatia* 22, no. 2. Spring, 2007): 162-181.
- Sugarman, Stephen D. "Accountability through the Courts." *The School Review* 82, no. 2. February, 1974): 233-259.
- Tan, Anh, *et al.*, "The Human Hippocampus is Not Sexually-Dimorphic: Meta-Analysis of Structural MRI Volumes." *NeuroImage* 124, part A. January, 2016): 350-366.
- Tani, Karen M. "Welfare and Rights Before the Movement: Rights as a Language of the State." *Yale Law Journal* 122, no. 2. November, 2012): 314-383.
- Tatum, Kimberly M., Joyce Coleman Nichols, and Fernaudra Ferguson. "Examining Institutional Barriers that Impede Access to Legal Education." *Negro Educational Review* 59, nos. 1-2. Spring/Summer, 2008): 79-92.
- Terzian, Seven G. and Leigh Shapiro. "Corporate Science Education: Westinghouse and the Value of Science in Mid-Twentieth Century America." *Public Understanding of Science* 24. 2): 147-166.
- Tucker, William H. "Places, Everyone! IQ Heritability, Ideology, and Education." *Journal of Educational Thought* 32, no. 3. December, 1998): 217-240.

- Turner, Sarah, and John Bound. "Closing the Gap or Widening the Divide: the Effects of the G.I. Bill and World War II on the Educational Outcomes of Black Americans." *Journal of Economic History* 63, no. 1. March, 2003): 145-177.
- Tyack, David, and William Tobin. "The 'Grammar' of Schooling: Why Has it Been so Hard to Change?" *American Education Research Journal* 31, no. 3. Autumn, 1994): 453-479.
- Tyson, Karolyn, William Darity, Jr., and Domini R. Castellino. "It's Not 'A Black Thing': Understanding the Burden of Acting White and Other Dilemmas of High Achievement." *American Sociological Review* 70, no. 4. August, 2005): 582-605.
- Vogel, David and Mark Nadel. "Who is a Consumer? An Analysis of the Politics of Consumer Conflict." *American Politics Quarterly* 5, no. 1. January, 1977): 27-56.
- Waldrep, Christopher. "National Policing, Lynching, and Constitutional Change." *Journal of Southern History* 74, no. 3. August, 2008): 589-626.
- Wang, Jessica. "Scientists and the Problem of the Public in Cold War America, 1945-1960." *Osiris* 2nd series, 17. 2002): 323-247.
- Wansink, Brian. "Changing Eating Habits on the Home Front: Lost Lessons from World War II Research." *Journal of Public Policy & Marketing* 21, no. 1. Spring, 2002): 90-99.
- West, Candace and Don Zimmerman. "Doing Gender." *Gender and Society* 1, no. 2. June 1987): 125-151.
- White, Jr., George. "'I Am Teaching Some of the Boys': Chaplain Robert Boston Dokes and Army Testing of Black Soldiers in World War II, *Journal of Negro Education* 81, no. 3. Summer, 2012): 200-217.
- Williams, Arthur R., and Karl F. Johnson. "Race, Social Welfare, and the Decline of Postwar Liberalism: A New or Old Key?" *Public Administration Review* 60, no. 6. November-December, 2000): 560-572.
- Williams, David R., *et al.* "Traditional and Contemporary Prejudice and Urban Whites' Support for Affirmative Action and Government Help." *Social Problems* 46, no. 4. November, 1999): 503-527.
- Willrich, Michael. "The Two Percent Solution: Eugenic Jurisprudence and the Socialization of American Law, 1900-1930." *Law and History Review* 16, no. 1. Spring, 1998): 63-111.
- Wilson, George, and Amie L. Nielson. "'Color Coding' and Support for Social Policy spending: Assessing the Parameters among Whites." *Annals of the American Academy of Political and Social Science* 634. March, 2011): 174-189.
- Winant, Howard. "Race and Race Theory." *Annual Review of Sociology* 26. 2000): 169-185.

Witkowski, Terrence H. "World War II Poster Campaigns: Preaching Frugality to American Consumers." *Journal of Advertising* 32, no. 1. Spring, 2003): 69-82.

Woodcock, Richard W. "New Looks in the Assessment of Cognitive Ability." *Peabody Journal of Education* 77, no. 2. 2002): 6-22.

Wozencraft, Marian. "Are boys better than girls in arithmetic?" *The Arithmetic Teacher* 10, no. 8. December, 1963): 486-490.

Yamamoto, Kaoru. "Psychological Testing: Invasion of Privacy?" *Educational Leadership* February 1966: 363-368.
http://ascd.com/ASCD/pdf/journals/ed_lead/el_196602_yamamoto.pdf.

Yonge, George D. "The Use of Masculinity-Femininity Measures to Account for Sex Differences in Problem Solving." *California Journal of Educational Research* 12, no. 5. November, 1961): 208-212, 220.

Published Sources—Edited Volumes

Evans, Peter B., Dietrich Reuschmeyer, and Theda Skocpol, editors. *Bringing the State Back In*. Cambridge: Cambridge University Press, 1985.

Johanek, Michael C., editor. *A Faithful Mirror: Reflections on the College Board and Education in America*. New York: College Board, 2001.

Grinstein, Louise S., and Paul J Campbell, editors. *Women of Mathematics: A Biobibliographic Sourcebook*. Westport, CT: Greenwood Press, 1987.

Simpson, Christopher, editor. *Universities and Empire: Money and Politics in the Social Sciences During the Cold War*. New York: The New Press, 1998.

Solovey, Mark, and Hamilton Cravens, editors. *Cold War Social Science: Knowledge Production, Liberal Democracy, and Human Nature*. New York: Palgrave MacMillan, 2012.

Published Sources—Memoir

Kaplan, Stanley H., and Anne Farris, *Test Pilot: How I Broke Testing Barriers for Millions of Students and Caused a Sonic Boom in the Business of Education*. New York: Simon & Schuster, 2001.

Published Sources—Monographs

Appel, Toby A. *Shaping Biology: The National Science Foundation and American Biological Research, 1945-1975*. Baltimore: Johns Hopkins University Press, 2000.

- Au, Wayne. *Unequal By Design: High-Stakes Testing and the Standardization of Inequality*. New York: Routledge, 2009.
- Baucom, Ian. *Specters of the Atlantic: Finance Capital, Slavery, and the Philosophy of History*. Durham: Duke University Press, 2005.
- Beadie, Nancy. *Education and the Creation of Capital in the Early American Republic*. Cambridge: Cambridge University Press, 2010.
- Bederman, Gail. *Manliness and Civilization: A Cultural History of Gender and Race in the United States*. Chicago: University of Chicago Press, 1996.
- Bell, Howard. *The Bakke Case: Race Education, and Affirmative Action*. Lawrence: University Press of Kansas, 2000.
- Boehm, Lisa Krissof. *Making a Way Out of No Way: African American Women and the Second Great Migration*. Jackson: University of Mississippi Press, 2009.
- Bleier, Ruth. *Science and Gender: A Critique of Biology and its Theories on Women*. New York: Pergamon Press, 1984.
- Bowles, Frank. *The Refounding of the College Board, 1948-1963: An Informal Commentary and Selected Papers*. New York: College Entrance Examination Board, 1967.
- Brown, Harrison, James Bonner, and John Weir, *The Next Hundred Years*. London: The Scientific Book Club, 1957.
- Butler, Judith. *Bodies That Matter: On the Discursive Limits of Sex*. New York: Routledge, 2011.
- Chapman, Paul Davis. *Schools as Sorters: Lewis M. Terman, Applied Psychology, and the Intelligence Testing Movement 1890-1930*. New York: NYU Press, 1988.
- Cohen-Cole, Jamie. *The Open Mind: Cold War Politics and the Sciences of Human Nature*. Chicago: University of Chicago Press, 2014.
- Cline Cohen, Patricia. *A Calculating People: The Spread of Numeracy in Early America*. New York: Routledge, 1999.
- Clowse, Barbara Barksdale. *Brainpower for the Cold War: The Sputnik Crisis and the National Defense Education Act of 1958*. Westport, CT: Greenwood Press, 1981).
- Cohen, Lizabeth. *A Consumers' Republic: The Politics of Mass Consumption in Postwar America*. New York: Alfred A. Knopf, 2003.

- Cooper, Brian P. *Family Fictions and Family Facts: Harriet Martineau, Adolphe Quetelet, and the Population Question in England, 1798-1859*. London: Routledge, 2007.
- Critchlow, Donald T. *Phyllis Schlafly and Grassroots Conservatism: A Woman's Crusade*. Princeton: Princeton University Press, 2005.
- Dow, Peter B. *Sputnik Politics: Lessons from the Sputnik Era*. Cambridge: Harvard University Press, 1991.
- Downs, Greg. *Declarations of Dependence: The Long Reconstruction of Popular Politics in the South, 1861-1908*. Chapel Hill: UNC Press, 2011.
- Dudziak, Mary. *Cold War Civil Rights: Race and the Image of American Democracy*. Princeton: Princeton University Press, 2011.
- Edwards, Paul N. *The Closed World: Computers and the Politics of Discourse in Cold War America*. Cambridge, MIT Press, 1996.
- Fausto-Sterline, Anne. *Myths of Gender: Biological Theories about Women and Men*. New York: Basic Books, 1985.
- Fields, Karen E. and Barbara J. Fields, *Racecraft: The Soul of Inequality in American Life*. London: Verso, 2014);
- Fine, Cordelia. *Delusions of Gender: The Real Science Behind Sex Differences*. London: Icon, 2010.
- Foster, Amy. *Integrating Women into the Astronaut Corps: Politics and Logistics at NASA, 1972-2004*. Baltimore: Johns Hopkins University Press, 2011.
- Frydl, Kathleen. *The GI Bill*. Cambridge: Cambridge University Press, 2009.
- Garrison, Mark J. *A Measure of Failure: The Political Origins of Standardized Testing*. Albany: SUNY Press, 2009.
- Gilmore, Glenda. *Gender and Jim Crow: Women and the Politics of White Supremacy in North Carolina, 1896-1920*. Chapel Hill: University of North Carolina Press, 1996.
- Glickman, Lawrence. *Buying Power: A History of Consumer Activism in America*. Chicago: University of Chicago Press, 2009.
- Hanford, George H. *Life with the SAT: Assessing Our Young People and Our Times*. New York: College Entrance Examination Board, 1991.

- Haney, David Paul. *The Americanization of Social Science: Intellectuals and Public Responsibility in the Postwar United States*. Philadelphia: Temple University Press, 2008.
- Hartman, Andrew. *Education and the Cold War: The Battle for the American School*. New York: Palgrave Macmillan, 2008.
- Herrnstein, Richard J. and Charles Murray, *The Bell Curve: Intelligence and Class Structure in American Life*. New York: Free Press, 1994.
- Hoffman, Banesh. *The Tyranny of Testing*. New York: Crowell-Collier Press, 1962.
- Hofstadter, Richard. *Anti-Intellectualism in American Life*. New York: Alfred A. Knopf, 1963.
- Igo, Sarah E. *The Averaged American: Surveys, Citizens, and the Making of a Mass Public*. Cambridge: Harvard University Press, 2007);
- Ichisada, Miyazaki. *China's Examination Hell: The Civil Service Examinations of Imperial China*. Translated by Conrad Schirokauer. New York: John Weatherhill, 1976.
- Jacoby, Sanford M. *Employing Bureaucracy: Managers, Unions, and the Transformation of Work in the 20th Century*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers, 2004.
- Johnston, Robert D. *The Radical Middle Class: Populist Democracy and the Question of Capitalism in Progressive Era Portland*. Princeton: Princeton University Press, 2003.
- Jorgensen, Lloyd P. *The State and the Non-Public School*. Columbia: University of Missouri Press, 1987.
- Kamenetz, Anya. *The Test: Why Our Schools are Obsessed with Standardized Testing—But You Don't Have to Be*. New York: PublicAffairs, 2015.
- Kammerer, Gladys M. *Impact of War on Federal Personnel Administration, 1939-1945*. Lexington: University of Kentucky Press, 1951.
- Karabel, Jerome. *The Chosen: The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton*. Boston: Houghton Mifflin, 2005.
- Kempf, Arlo. *The Pedagogy of Standardized Testing: The Radical Impacts of Educational Standardization in the US and Canada*. New York: Palgrave MacMillan, 2016.
- Kline, Wendi. *Building a Better Race: Gender, Sexuality, and Eugenics from the Turn of the Century to the Baby Boom*. Berkeley: University of California Press, 2001.
- Knowles, Asa S., and Robert D. Thomson, *Management of Manpower*. New York: The MacMillan Company, 1943.

- Kryder, Daniel. *Divided Arsenal: Race and the American State During World War II*. Cambridge: Cambridge University Press, 2000.
- Laudiciana, Robert A. and Joseph L. Tramultola, *A Legal Overview of the New Student: As Educational Consumer, Citizen, and Bargainer*. Springfield, IL: Charles C. Thomas, 1976.
- Lemann, Nicholas. *The Big Test: The Secret History of the American Meritocracy*. New York: Farrar, Straus, & Giroux, 2000.
- Levine, Susan. *School Lunch Politics: The Surprising History of America's Favorite Welfare Program*. Princeton: Princeton University Press, 2011.
- Lorber, Judith. *Paradoxes of Gender*. New Haven: Yale University Press, 1995.
- Manahan, John Levi. *A Bibliography of Educational Surveys and Tests*. Charlottesville, VA: The University of Virginia Press, 1916.
- Martinot, Steve. *The Machinery of Whiteness: Studies in the Structure of Racialization*. Philadelphia: Temple University Press, 2010.
- McGerr, Michael. *A Fierce Discontent: The Rise and Fall of the Progressive Movement, 1870-1920*. New York: Free Press, 2003.
- Mertler, Craig A. *Interpreting Standardized Test Scores*. New York: SAGE Publications, 2007.
- Miller, Joseph Hillis, and John S. Allen. *Veterans Challenge the Colleges: The New York Program*. New York: King's Crown Press, 1947.
- Murphy, Marjorie. *Blackboard Unions: The AFT and the NEA, 1900-1980*. Ithaca: Cornell University Press, 1992.
- Nairn, Allan. *The Reign of ETS: The Corporation that Makes Up Minds*. New York: Ralph Nader, 1980.
- O'Brien, Gerald V. *Framing the Moron: The Social Construction of Feeble-Mindedness in the American Eugenic Era*. Manchester: Manchester University Press, 2013.
- Oakes, Jeannie. *Keeping Track: How Schools Structure Inequality*. New Haven: Yale University Press, 1985.
- Omi, Michael, and Howard Winant. *Racial Formation in the United States*, 3rd edition. New York: Routledge, 2015.
- Osen, Lynn M. *Women in Mathematics*. Cambridge: MIT Press, 1974.

- Owen, David. *None of the Myth: Behind the Myth of Scholastic Aptitude*. New York: Houghton Mifflin, 1985.
- Perl, Teri. *Math Equals: Biographies of Women Mathematicians + Related Activities*. Menlo Park, CA: Addison-Wesley Publishing, 1978.
- Perl, Teri, and Joan M. Manning. *Women, Numbers and Dreams: Biographical Sketches and Math Activities*. Santa Rosa, CA: National Women's History Project, 1982.
- Peterson, Paul E. *The Politics of School Reform, 1970-1940*. Chicago: University of Chicago Press, 1985.
- Poole, Mary. *The Segregated Origins of Social Security: African Americans and the Welfare State*. Chapel Hill: UNC Press, 2008.
- Prévost, Jean-Guy, and Jean-Pierre Beaud. *Statistics, Public Debate and the State, 1800-1945: A Social, Political and Intellectual History of Numbers*. London: Pickering and Chatto, 2012.
- Puaca, Laura Micheletti. *Searching for Scientific Womanpower: Technocratic Feminism and the Politics of National Security, 1940-1980*. Chapel Hill: University of North Carolina Press, 2014.
- Reagan, Leslie. *When Abortion Was a Crime: Women, Medicine, and the Law in the United States, 1867-1973*. Berkeley: University of California Press, 1997.
- Reese, William J. *Testing Wars in the Public Schools: A Forgotten History*. Cambridge: Harvard University Press, 2013.
- Reschly, Daniel J. *Disproportionate Minority Representation in General and Special Education: Patterns, Issues, and Alternatives*. Des Moines: Iowa State Department of Education, 1997.
- Ross, Davis R.B. *Preparing for Ulysses: Politics and Veterans During World War II*. New York: Columbia University Press, 1969.
- Rossiter, Margaret W. *Women Scientists in America, Volume 1: Struggles and Strategies to 1940*. Baltimore: Johns Hopkins University Press, 1982.
- Rossiter, Margaret W. *Women Scientists in America, Volume 2: Before Affirmative Action, 1940-1972*. Baltimore: Johns Hopkins University Press, 1995.
- Sahlberg, Pasi. *Finnish Lessons 2.0: What Can the World Learn from Educational Change in Finland?*. New York: Teachers College Press, 2014.

- Sætnan, Ann Rudinow, Heidi Mork Lomell, and Svein Hammer, *The Mutual Construction of Statistics and Society*. New York: Routledge, 2011.
- Scott, Joan Wallach. *Gender and the Politics of History*. New York: Columbia University Press, 1988);
- Selden, Steven. *Inheriting Shame: The Story of Eugenics and Racism in America*. New York: Teachers College Press, 1999.
- Self, Robert O. *All in the Family: the Realignment of American Democracy Since the 1960s*. New York: Hill and Wang, 2012.
- Slawson, Douglas J. *The Department of Education Battle, 1918-1932: Public Schools, Catholic Schools, and the Social Order*. Notre Dame: University of Notre Dame Press, 2005.
- Smith, Dorothy E. *The Everyday World As Problematic: A Feminist Sociology*. Boston: Northeastern University Press, 1987.
- Solovey, Mark. *Shaky Foundations: The Politics-Patronage-Social Science Nexus in Cold War America*. New Brunswick: Rutgers University Press, 2013.
- Stevens, Mitchell L. *Creating a Class: College Admissions and the Education of Elites*. Cambridge: Harvard University Press, 2007.
- Trent, James W. *Inventing the Feeble Mind: A History of Mental Retardation in the United States*. Berkeley: University of California Press, 1994.
- Valentine, John A. *The College Board and the School Curriculum: A History of the College Board's Influence on the Substance and Standards of American Education, 1900-1980*. New York: College Entrance Examination Board, 1987.
- Wang, Jessica. *American Science in an Age of Anxiety: Scientists, Anticommunism, and the Cold War*. Durham: University of North Carolina Press, 1999.
- Wang, Zuoyue. *In Sputnik's Shadow: The President's Science Advisory Committee and Cold War America*. New Brunswick: Rutgers University Press, 2008.
- Weems, Jr., Robert E. *Desegregating the Dollar: African American Consumerism in the Twentieth Century*. New York: NYU Press, 1998.
- Wigdor, Alexandra K. and Wendell R. Garner, *Ability Testing: Uses, Consequences and Controversies*, vol. 1. Washington, D.C.: National Academy Press, 1982.
- Wilson, Edward O. *Sociobiology: The New Synthesis*. Cambridge: Belknap Press of Harvard, 1975.

Published Sources—Reports

- American Association for American Women, “The Simple Truth About the Gender Pay Gap: Spring 2017 Edition,” January 2017, <http://www.aauw.org/resource/the-simple-truth-about-the-gender-pay-gap/>.
- Beede, David, *et al.*, Office of the Chief Economist, U.S. Department of Commerce Economics and Statistics Administration, “Women in STEM: A Gender Gap to Innovation,” ESA Issue Brief #04-11, August 2011, available <http://www.esa.doc.gov/sites/default/files/womeninstemagaptoinnovation8311.pdf>
- Blau, Francine D. and Lawrence M. Kahn, “The Gender Wage Gap: Extent, Trends, and Explanations,” Discussion Paper no. 9656,” Institute for the Study of Labor, January 2016, <http://ftp.iza.org/dp9656.pdf>
- Cole, Jr., Charles C. *Encouraging Scientific Talent: A Study of America’s Able Students Who are Lost to College and of Ways to Attracting Them to College and Science Careers*. New York: College Entrance Examination Board, 1956.
- Chicago Teachers Union, “Debunking the Myths of Standardized Testing: A CTU Position Paper,” CTU website, http://www.ctunet.com/quest-center/research/position-papers/text/CTU_Testing_Position_Brief_web-1.pdf
- College Board Office of Research and Development, “The Effects of SAT Scale Recentering on Percentiles,” Research Summary RS-05, May 1999, <http://files.eric.ed.gov/fulltext/ED563025.pdf>
- College Board, *Test Specifications for the Redesigned SAT*. College Board, 2015), 197; <https://collegereadiness.collegeboard.org/pdf/test-specifications-redesigned-sat-1.pdf>.
- College Entrance Examination Board, *Effects on Coaching on Scholastic Aptitude Test Scores*. New York: College Entrance Examination Board, 1965),
- Tim Davey, “A Guide to Computer Adaptive Testing Systems,” Council of Chief State School Officers, 2011, <http://files.eric.ed.gov/fulltext/ED543317.pdf>;
- Dorans, Neil J. *The Recentering of SAT Scales and Its Effects on Score Distributions and Score Interpretations*. New York: College Entrance Examination Board, 2002), <http://www.ets.org/Media/Research/pdf/RR-02-04-Dorans.pdf>
- Educational Testing Service, *Statistical Studies of Selective Service Testing, 1951-1953*. Princeton: Educational Testing Service, 1955.
- Educational Testing Service: *The ETS Sensitivity Review Process: Guidelines and Procedures*. Princeton, Educational Testing Service, 1986.

- Educational Testing Service, *ETS Sensitivity Review Process: An Overview*. Princeton: Educational Testing Service, 1987.
- Ellison, Glenn, and Ashley Swanson. "The Gender Gap in Secondary School Mathematics at High Achievement Levels: Evidence from the American Mathematics Competitions," July 2009, <http://economics.mit.edu/files/4298>.
- Frusciano, Thomas J. "Student Deferment and the Selective Service College Qualification Test, 1951-1967". ETS-RM-83-1), Princeton: Educational Testing Service, 1980:
- Fryer, Jr., Roland G., *et al*, "Enhancing the Efficacy of Teaching Incentive Through Loss Aversion: A Field Experiment," National Bureau of Economic Research Working Paper 18237. Cambridge: NBER, 2012), <http://www.nber.org/papers/w18237>
- Guddemi, Marcy, and Betsy Case, "Assessment Report: Assessing Young Children, Pearson, 2004, http://images.pearsonassessments.com/images/tmrs/tmrs_rg/Assessing_YoungChildren.pdf?WT.mc_id=TMRS_Assessing_Young_Children
- Haines, Michael R. "Long Term Marriage Patterns in the United States from Colonial Times to the Present," National Bureau of Economic Research Historical Paper 80. March 1996. <http://www.nber.org/papers/h0080.pdf>.
- Jacobs, Paul I. "Effects of Coaching on the College Board English Composition Test," ETS Research Bulletin RB-64-24. Princeton: Educational Testing Service, 1964.
- Langdon, David, *et al.*, Office of the Chief Economist, U.S. Department of Commerce Economics and Statistics Administration, "STEM: Good Jobs Now and for the Future," ESA Issue Brief #03-11, July 2011, available http://www.esa.doc.gov/sites/default/files/stemfinaljuly14_1.pdf
- Maldonado-Rivera, Manuel. *The History of the Puerto Rico and Latin American Office*, edited by Michael Farah, <http://about.collegeboard.org/region-offices/puerto-rico>
- McClintock, Charles Graham. *The Competition in Education: US vs. USSR. Tempo Report RM 58TMP-57*. Santa Barbara, CA: General Electric Company, 1958),
- Messick, Samuel "Test Validity and the Ethics of Assessment,". Princeton: Educational Testing Service, 1979.
- Messick, Samuel, *et al.* *The Effectiveness of Coaching for the SAT: Review and Reanalysis of Research from the Fifties to the FTC*. Princeton: Educational Testing Service, 1980)
- Harold W. Metz, *Is There Enough Manpower?*. Washington, D.C.: The Brookings Institution, 1942.

- Mullins, Ina V. S. *Educational Achievement and Sex Discrimination*. Denver: National Assessment of Educational Progress, 1975), 7, <http://files.eric.ed.gov/fulltext/ED115701.pdf>. last accessed February 19, 2017.
- National Association for College Admission Counseling, “Report of the Commission on the Use of Standardized Tests in Undergraduate Admission”. Arlington, VA: NACAC, 2008. September, 2008, <http://files.eric.ed.gov/fulltext/ED502721.pdf>;
- National Commission on Excellence in Education, *A Nation At Risk: The Imperative for Educational Reform*. Washington, D.C.: U.S. Government Printing Office, 1983.
- National Manpower Council, *Student Deferment and National Manpower Policy*. New York: Columbia University Press, 1952.
- The Ohio State University, War-Time Training for Young Men. Columbus: The Ohio State University, c. 1943.
- Research Institute of America, *Manpower: Solving the Problems of Labor Shortage*. New York: Research Institute of America, 1943)
- Rodgers, William M. and Leslie S. Stratton, “The Male Marital Wage Differential: Race, Training, and Fixed Effects,” Discussion Paper no. 1745, Institute for the Study of Labor, September 2005, <http://repec.iza.org/dp1745.pdf>
- Sandow, Stuart A. “Emerging Education Policy Issues in Law: Fraud “. Syracuse: Educational Policy Research Center, 1970), <http://files.eric.ed.gov/fulltext/ED062910.pdf>
- Shepard, Lorrie, Sharon Lynn Kagan, and Emily Wurtz, eds. “Principles and Recommendations for Early Childhood Assessments.” National Education Goals Panel, February 1999, <http://govinfo.library.unt.edu/negp/reports/prinrec.pdf>.
- Tompson, Trevor, Jennifer Benz, and Jennifer Agiesta, “Parents’ Attitudes on the Quality of Education in the United States,” The Associated Press-NORC Center for Public Affairs Research, August 2013, http://www.apnorc.org/PDFs/Parent%20Attitudes/AP_NORC_Parents%20Attitudes%20on%20the%20Quality%20of%20Education%20in%20the%20US_FINAL_2.pdf;
- Turnbull, William W. “Student Change, Program Change: Why the SAT Scores Kept Falling,” College Board Report No. 85-2. New York: College Entrance Examination Board, 1985.
- University of Michigan, “Academic Achievement of Veterans and Non-veterans in Six Colleges of the University of Michigan”. Ann Arbor: University of Michigan Institute of Human Adjustment, 1946),
- U.S. Bureau of Labor Statistics, “Highlights of Women’s Earnings in 2015,” Report no. 1064, November 2016, <https://www.bls.gov/opub/reports/womens-earnings/2015/pdf/home.pdf>;

Watley, Donovan J. "Career Progress of Merit Scholars". Evanston, IL: National Merit Scholarship Corporation, 1968.

Wisconsin State Department of Public Instruction, *The Wisconsin Educational Program: A Handbook for World War II Veterans—War Workers and Counselors*. Madison, WI: Veterans Recognition Board, 1944)

Published Sources—Test Preparation and Guidance Books

Arco Publishing Company. *Practice for the Army Tests*. New York: Arco Publishing Company, 1942.

Benjamin, Hazel C. *Employment Tests in Industry and Business: A Selected, Annotated Bibliography*. Princeton, Princeton University, 1945.

Bosdell, Betty Jane. *Evaluation of Counseling Treatments with Underachieving High School Students*. Grand Forks, ND: Cooperative Research Project, 1962.

Bricklin, Barry, and Patricia M. Bricklin. *Bright Child—Poor Grades: The Psychology of Underachievement*. New York: Dell, 1967.

Brown, Francis J. *Educational Opportunities for Veterans*. Washington, D.C.: Public Affairs Press, 1946.

Cardall, Alfred J. *A Wartime Guidance Program for Your School*. Chicago: Science Research Associates, 1943.

Crow, Lester D., and Alice Crow. *An Introduction to Guidance: Principles and Practices*. New York: American Book Company, 1951.

Current Book Company. *Air Corps Practice Tests*. New York: Current Book Company, 1942.

Darley, John G. *Clinical Aspects and Interpretation of the Strong Vocational Interest Blank*. New York: The Psychological Corporation, 1941.

Endicott, Frank S. *How to Find and Succeed in Your Post-War Job*. Scranton, PA: International Textbook Company, 1946.

Freehill, Maurice F. *Gifted Children: Their Psychology and Education*. New York: MacMillan Company, 1961.

Hamrin, Shirley A., and Blanche B Paulson. *Counseling Adolescents*. Chicago: Science Research Associates, 1950, 15th printing 1965.

- Hatch Raymond N., and Paul L. Dressel. *Guidance Service in the Secondary School*. Dubuque: W.C. Brown Co., 1953.
- Horchow, Reuben. *How to Get Ahead in the Armed Forces*. New York: Reuben Horchow, 1951.
- Kaplan, Alvin Harold. *Prepare for the Official U.S. Army-Navy Tests*. New York: Capitol Publishing Company, 1942.
- Knapp, Robert Henry. *Practical Guidance Methods for Counselors, teachers, and Administrators*. New York: McGraw Hill, 1953.
- Kuder, G. Frederic. *Revised Manual for the Kuder Preference Record*. Chicago: Science Research Associates, 1946.
- Lyons, George J., and Harmon C. Martin. *The Strategy of Job Finding*. New York: Prentice-Hall, 1944.
- Miller, Leonard M., editor. *Guidance for the Underachiever with Superior Ability*. Washington, D.C.: U.S. Government Printing Office, 1961.
- Morris, Mark, ed., *Career Opportunities*. Washington, D.C.: Progress Press, 1946.
- Munson, Grace E., and Lester J. Schloerb. *High School Course in Self-Appraisal and Careers: Teacher's Manual*. Chicago: Chicago Board of Education Bureau of Child Study, 1941.
- Myers, Isabel Briggs. *The Myers-Briggs Type Indicator: Manual*. Palo Alto, CA: Consulting Psychologists Press, 1962.
- Ohlsen, Merle M. *The Extent to Which Group Counseling Improves the Academic and Personal Adjustment of Under-Achieving Gifted Adolescents*. Urbana: University of Illinois Press, 1960.
- Ohlsen, Merle M. *Appraisal of Group Counseling for Underachieving Bright Fifth Graders and Their Parents*. Urbana: University of Illinois Press, 1964.
- Otis, Jay L., Carroll L. Shartle, and William H. Stead. *Occupational Counseling Techniques: Their Development and Application*. New York: American Book Company, 1940.
- Paulson, Blanche. *The Magic of the Mind*. Chicago: Chicago Board of Education Bureau of Child Study, 1943.
- Pergande Publishing Company. *Full-Length Practice Examinations for U.S. Army-Navy Classification Tests*. Milwaukee, Pergande Publishing Company, 1943.
- Powell, Marvin. *The Psychology of Adolescence*. Indianapolis: Bobbs-Merrill Company, 1963.

- Practice Publishing Company. *Practice for the Soldier's IQ Test*. Cincinnati: Practice Publishing Company, 1943.
- Recktenwald, Lester Nicholas. *Guidance and Counseling with Psychometric Principles: A Basic Treatment*. Washington, D.C.: Catholic University of America, 1953.
- Robinson, Adam and John Katzman. *Cracking the System: The SAT*. New York: Villard Books, 1986.
- Robinson, Adam and John Katzman, editors. *College Admissions: Getting into the College of Your Choice*. New York: Princeton Review, 1993.
- Ryan, David G. *The First Step in Guidance: Self-Appraisal: A Report of the 1940 Sophomore Testing Program*. New York: American Council on Education Cooperative Test Service, 1941.
- Stewart, Maxwell S. *When Will I Get Out? When Will I Find A Job?*. New York: Public Affairs Pamphlets, 1943.
- Suerken, Ernest Henry, editor. *Find Your Vocation*. Revised edition. New York: Hobson Book Press, 1941.
- Theory of Mental Tests*. New York: John Wiley & Sons, 1950.
- Thorndike, Robert. *The Concepts of Over- and Underachievement*. New York: Columbia University, 1963.
- U.S. Department of Health, Education, and Welfare. *The Adolescent in Your Family*. Washington, D.C.: U.S. Department of Health Education, and Welfare Children's Bureau, 1955.
- Vilmos, William. *Instruction for Tests in the Armed Forces*. New York: Grosset & Dunlap, 1943.
- Watters, Henry Eugene. *Youth Makes the Choice: Vocational and Social Self-Guidance*. Nashville, TN: Broadman Press, 1943.
- Wrenn, C. Gilbert, and Willis E. Dugan, *Guidance Procedures in High School: Some Recommended Procedures Based Upon a Survey of Present Practices in Minnesota*. Minneapolis: University of Minneapolis Press, 1950.
- Yale, Burchard W. *Your Job: Selecting It, Securing It, Succeeding in It*. New York: Barnes & Noble, 1947.

Theses

Hodgson, Mary Frances. "A Study of the Discovery and Development of Science Talent in Secondary School Pupils." Master's thesis, Kansas State College of Agriculture and Applied Science, 1956.

Shepherd, Keegan. "Selling 'Dream Insurance': The Standardized Test-Preparation Industry's Search for Legitimacy, 1946-1989." Master's thesis. University of Central Florida. 2011.

Whittell, Florence Emma. "Standardization of the Yerkes Multiple Choice Method for Human Adults." Master's thesis, University of California, 1920.

Websites

ACT Media Relations. <http://www.act.org/content/act/en/newsroom.html>.

American Presidency Project. <http://www.presidency.ucsb.edu>.

Center for Public Education. <http://www.centerforpubliceducation.org>.

College Board. <https://www.collegeboard.org>.

Cornell Law Legal Information Institute. Wex Legal Dictionary.
<https://www.law.cornell.edu/wex>.

Educational Testing Service. <http://www.ets.org>.

FairTest. <https://www.fairtest.org>.

IBM Archives. <http://www-03.ibm.com/ibm/history/>

Khan Academy. <https://www.khanacademy.org/>.

National Archives. "Archived Presidential White House Websites."
<https://www.archives.gov/presidential-libraries/archived-websites>.

National Council of State Boards of Nursing. <https://www.ncsbn.org/>.

New York State Education Department. <http://www.nysed.gov>.

Organization for Economic Cooperation and Development. www.oecd.org/unitedstates/

Oyez. <http://www.oyez.org>.

Pearson News, Media & Blogs. <https://www.pearson.com/corporate/news.html>

U.S. Bureau of Labor Statistics. <https://www.bls.gov>.

U.S. Census Bureau. CPS Historical Time Series. <https://www.census.gov/library.html>

U.S. Department of Education. National Center for Education Statistics. <https://nces.ed.gov>