



Equity and Assessment in the Post-COVID-19 Era

Randy Bennett

Educational Testing Service
Princeton, NJ 08541
rbennett@ets.org

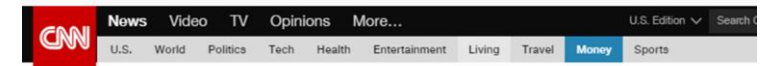
Presented as part of the International Association for Educational Assessment webinar series, International Approaches to Exams Given the Pandemic 2021, April 2021

Overview

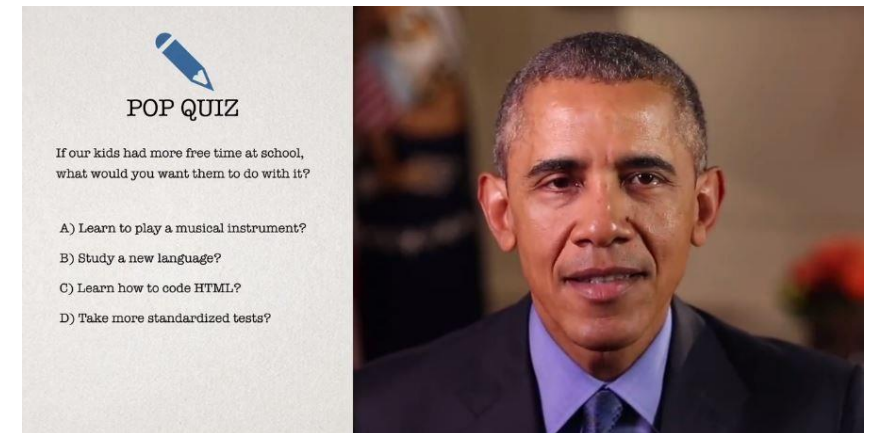
- COVID-19, education, and assessment in the K-12 and postsecondary sectors
- Structural inequity in US society
- Opportunity to learn, achievement, and testing
- Enhancing opportunity to learn
- Summary

K-12 Education: Pre-COVID-19

- Considerable anti-testing sentiment in reaction to amount, uses, and consequences of tests
 - Teacher unions mobilized
 - Students opt-outed
 - President Obama called for limits on mandated testing
 - Many states:
 - Shortened or eliminated particular tests
 - Rescinded requirements for test-based educator evaluation
 - *ESSA* (2015) weakened federal role and encouraged states to explore new accountability assessment approaches through IADA grants
 - NH, LA, NC, GA (2)



Parents all over U.S. 'opting out' of standardized student testing



K-12: The COVID-19 Dynamic

- Spring 2020
 - 46 states closed schools, shifting to remote learning
 - Educators unprepared
 - Student attendance dropped
 - Assessments cancelled in all states
 - USDE announced Competitive Grants for State Assessments Program to further fund innovation in assessment

The New York Times

April 6, 2020

As School Moves Online, Many Students Stay Logged Out

Teachers at some schools across the country report that fewer than half of their students are participating in online learning.

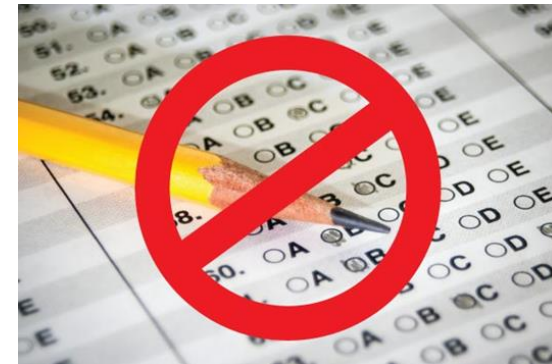
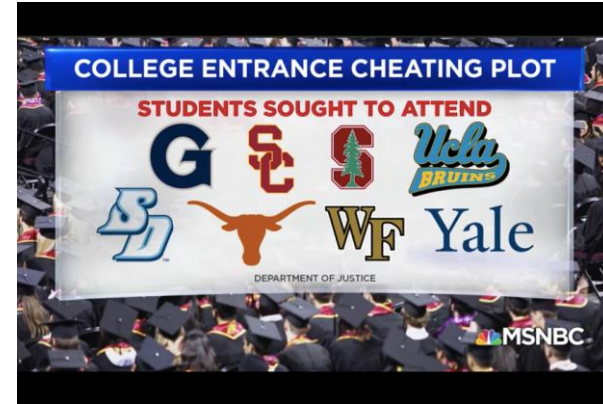
K-12: The COVID-19 Dynamic

- 2020-21 school year
 - Remote or blended learning continued
 - Research suggests significant negative impact on learning
 - Greater for traditionally underserved students
 - Spring assessments results may be affected by confounding factors



Higher Education: Pre-COVID-19

- Selective Institutions (Undergrad and Grad)
 - Longstanding fairness concerns over use of tests and their impact on diversity
 - Varsity Blues Scandal
 - Steadily growing movement for test-optional admissions
- Less-Selective Institutions
 - Changing supply-demand dynamic => test optional admissions



Higher Education: The COVID-19 Dynamic

- Spring 2020
 - Most campuses closed or opened with online classes
 - 2021 testing requirements waived by ~1600 institutions
 - Test-optional recommended by National Association for College Admissions Counseling (NACAC) for all institutions
 - University of California System
 - Announced phase-out of current admissions tests for instate students (5/21/20)



Higher Education: The COVID-19 Dynamic

- 2020-21 school year
 - Enrollments declined, most dramatically for community colleges
 - Remote and hybrid learning continued
 - Test-optional admissions extended to 2021-2022 cycle



**Standardized
test scores
not required**

THROUGH FALL 2021

UAB THE UNIVERSITY OF
ALABAMA AT BIRMINGHAM

COVID-19 Assessment Dynamic

- COVID accelerated existing trends related to dissatisfaction with traditional, single-event tests
 - K-12
 - Trying to innovate its way to greater positive impact
 - Performance assessments
 - Short, on-demand, competency-based assessments
 - End-of-unit, through-course assessments
 - Higher Education
 - Test-optional or test-blind admissions policies
 - Selective undergraduate institutions and graduate programs: Achieve greater diversity
 - Less-selective institutions: Attract more applicants

Structural Inequity

- Definitions

- *Structural racism*: “... a system in which public policies, institutional practices, cultural representations, and other norms work in various, often reinforcing ways to perpetuate racial group inequity. (Aspen Institute Roundtable on Community Change, 2004, p. 11)
 - Systemic racism, Institutional racism
- *Structural inequality*: “ ... occurs when the fabric of organizations, institutions, governments or social networks contains an embedded bias which provides advantages for some members and marginalizes or produces disadvantages for other members. (Wikipedia, 2021)
- *Structural stigma*: “ ... societal-level conditions, cultural norms, and institutional policies that constrain the opportunities, resources, and well-being of the stigmatized.” (Hatzenbeuhler, 2016)

Structural Inequity

- Definitions

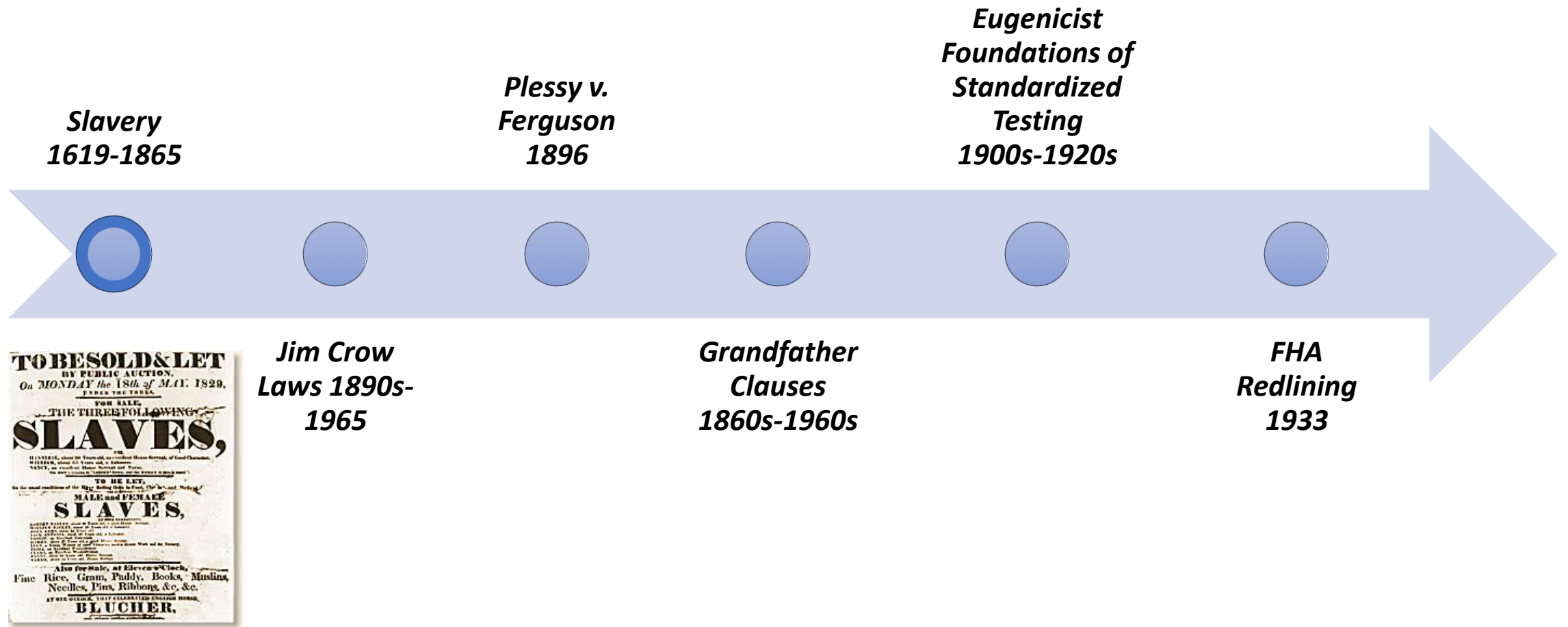
- *Structural racism*: “... a **system** in which **public policies, institutional practices**, cultural representations, and other **norms** work in various, often reinforcing ways to perpetuate racial group inequity. (Aspen Institute Roundtable on Community Change, 2004, p. 11)
 - Systemic racism, Institutional racism
- *Structural inequality*: “ ... occurs when the **fabric of organizations, institutions, governments** or social networks contains an embedded bias which provides advantages for some members and marginalizes or produces disadvantages for other members. (Wikipedia, 2021)
- *Structural stigma*: “ ... **societal-level conditions, cultural norms, and institutional policies** that constrain the opportunities, resources, and well-being of the stigmatized.” (Hatzenbeuhler, 2016)

Structural Inequity

- Definitions

- *Structural racism*: “... a system in which public policies, institutional practices, cultural representations, and other norms work in various, often reinforcing ways to **perpetuate racial group inequity**. (Aspen Institute Roundtable on Community Change, 2004, p. 11)
 - Systemic racism, Institutional racism
- *Structural inequality*: “ ... occurs when the fabric of organizations, institutions, governments or social networks contains an **embedded bias** which provides advantages for some members and **marginalizes or produces disadvantages for other members**. (Wikipedia, 2021)
- *Structural stigma*: “ ... societal-level conditions, cultural norms, and institutional policies that **constrain the opportunities, resources, and well-being of the stigmatized**.” (Hatzenbeuhler, 2016)

Examples of Structural Inequity in US Society



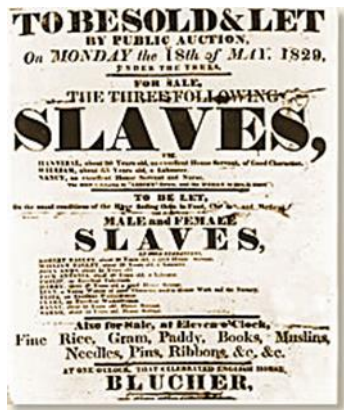
Examples of Structural Inequity in US Society

Slavery
1619-1865



Plessy v. Ferguson
1896

Eugenicist Foundations of Standardized Testing
1900s-1920s



Jim Crow Laws 1890s-1965

Grandfather Clauses 1860s-1960s

FHA Redlining 1933

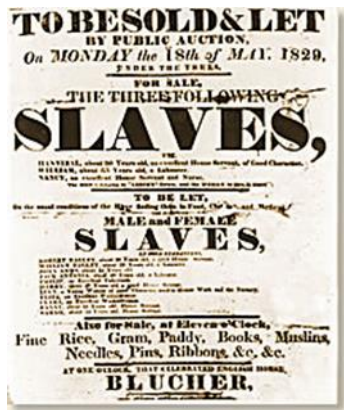
Examples of Structural Inequity in US Society

Slavery
1619-1865



Plessy v. Ferguson
1896

Eugenicist Foundations of Standardized Testing
1900s-1920s



Jim Crow Laws
1890s-1965



Homer Plessy

Grandfather Clauses
1860s-1960s

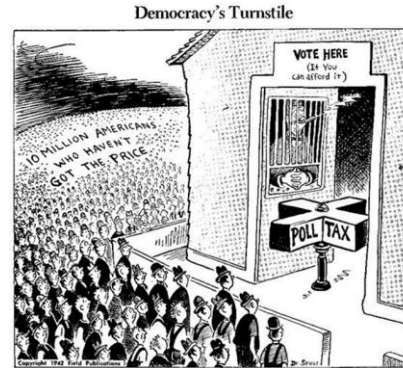
FHA Redlining
1933

Examples of Structural Inequity in US Society

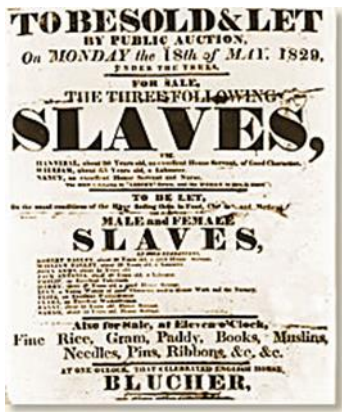
Slavery
1619-1865



Plessy v. Ferguson
1896



Eugenicist Foundations of Standardized Testing
1900s-1920s



Jim Crow Laws
1890s-1965



Homer Plessy

Grandfather Clauses
1860s-1960s

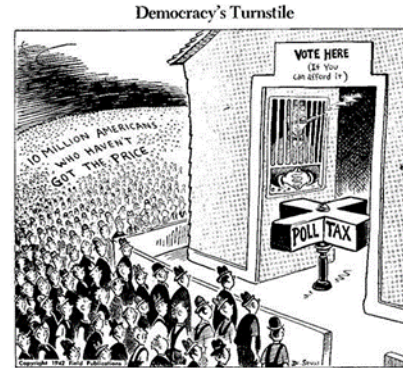
FHA Redlining
1933

Examples of Structural Inequity in US Society

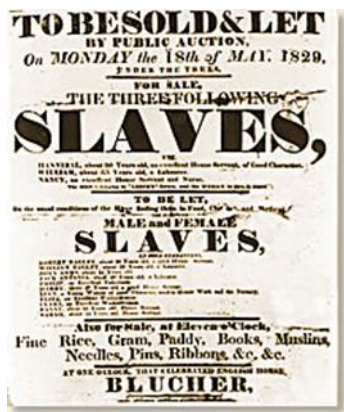
Slavery
1619-1865



Plessy v. Ferguson
1896



Eugenicist Foundations of Standardized Testing
1900s-1920s

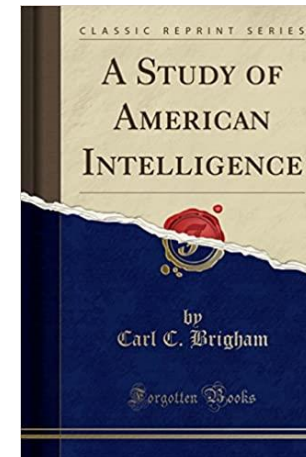


Jim Crow Laws
1890s-1965



Homer Plessy

Grandfather Clauses
1860s-1960s



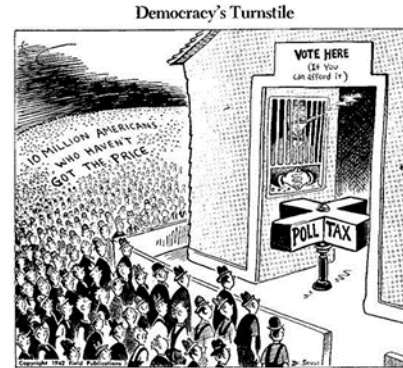
FHA Redlining
1933

Examples of Structural Inequity in US Society

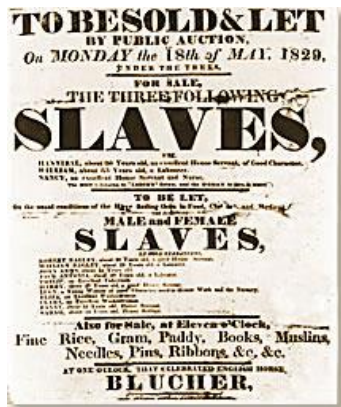
Slavery
1619-1865



Plessy v. Ferguson
1896



Eugenicist Foundations of Standardized Testing
1900s-1920s

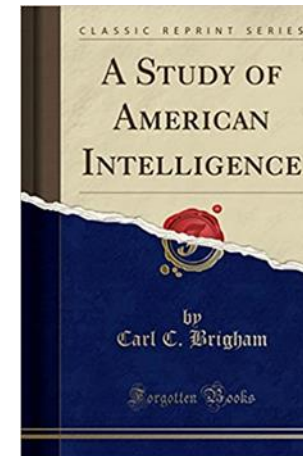


Jim Crow Laws
1890s-1965



Homer Plessy

Grandfather Clauses
1860s-1960s



FHA Redlining
1933

Examples of Structural Inequity's Impact

- Parents' education level (NCES, 2020)
 - Lived in household where highest level of either parent's education was Bachelor's degree or higher
 - 27% of Black students
 - 53% of white students



Examples of Structural Inequity's Impact

- Wealth (US Congress JEC, 2020)
 - Black family median net worth of \$17,150 in 2016 vs \$171,000 for White families
 - 42% of Black families owned homes compared with 73% of White families
 - 20.8% of Blacks vs 8.1% of Whites living below poverty level in 2018



Examples of Structural Inequity's Impact

- Crime
 - Victim status (Warnken & Lauritsen, 2019)
 - Blacks' risk of serious violence
 - ~1.5-2x greater than that of Whites
 - Steady over past four decades
 - Prison terms (Rehavi & Starr, 2014)
 - Black males get longer prison terms than whites arrested for the same federal offenses with the same prior records



Examples of Structural Inequity's Impact

- Health

- Black infant mortality rates 10.9 per 100,000 vs 4.7 per 100,000 for Whites (US Congress JEC, 2020)
- Blacks have average life expectancy at birth of 74.9 years, Whites 78.5 (US Congress JEC, 2020)
- 2018 uninsured rate of 11.5% for Blacks and 7.5% for Whites (KFF, 2021)
- Blacks more likely to have diabetes, high blood pressure, obesity, stroke, and early death from all causes (CDC, 2017)



Structural Inequity in Education

- Public school segregation 1968-2016

Percentage of Black Students Attending Schools that are Mostly Non-White

In 1968, 64% of U.S. black students attended public schools with fewer than 1 in 10 white students. By 2016, that proportion had dropped to 40%. These trends are largely driven by changes in the South.

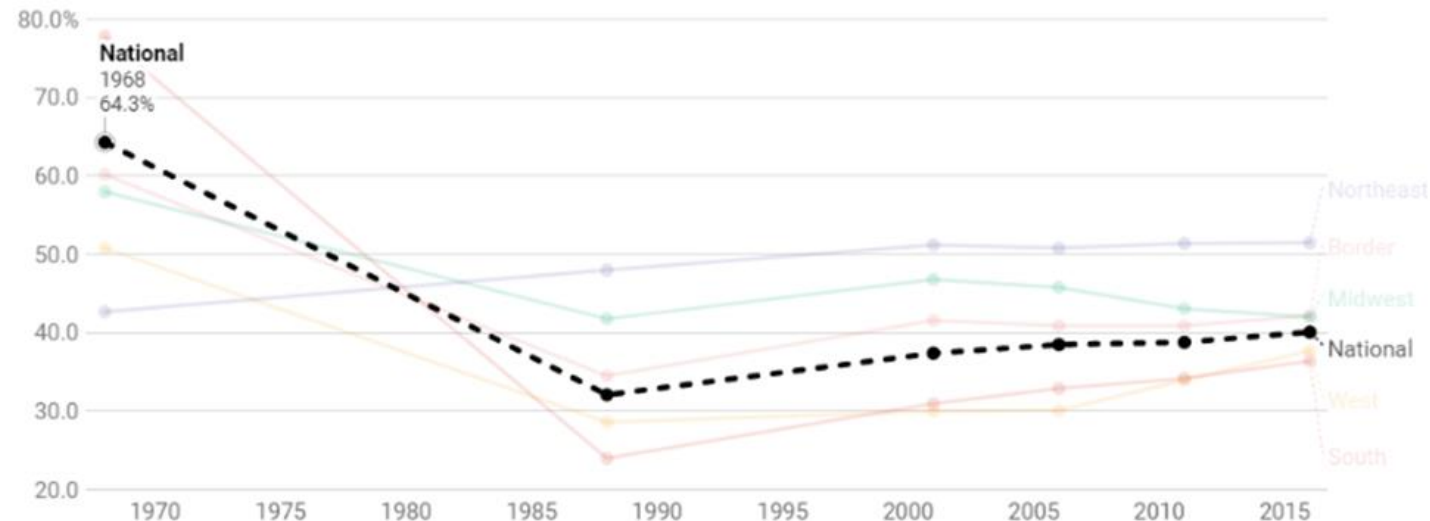


Chart: The Conversation, CC-BY-ND • Source: [The Civil Rights Project](#) • [Get the data](#)

Structural Inequity in Education Today

- Public school resources
 - In 2017, 45% of Black students attended high-poverty schools, compared with 8% of White students (NCES, 2020)
 - School funding via *local* property tax creates inequity
 - Dependence on property tax varies across states (NCES, 2020)
 - 2016–17
 - NH school districts received 59% of revenue
 - VT 0%
 - Equalization varies (EdTrust, 2018)
 - NJ gives the highest poverty districts 439% more per student than the lowest poverty districts
 - MT and MI offer ~0% differential

Structural Inequity in Education Today

- School funding via *local* property tax creates inequity
 - Average nonwhite district receives \$2,226 less per pupil than a white district, ~84% (EdBuild, 2016-2021)
- Wealthier school districts have higher voluntary contributions (Nelson & Gazley, 2014)
- Funding inequities reflected in educator salaries, instructional resources, extracurricular offerings, facilities, maintenance



Structural Inequity in Education Today

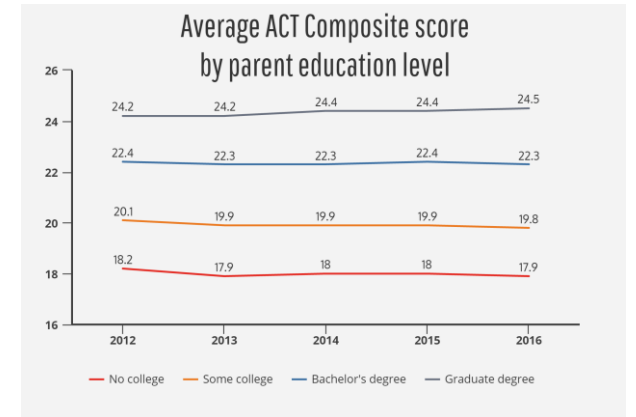
- Teacher-quality differences (NCES, 2017)
 - Percentage of 8th-grade public-school students who had a state-certified math teacher
 - 86% of Black students vs. 91% of White students
 - 80% of Black students attending a school with high minority enrollment (>75%) vs. 88% of White students attending such a school
 - Percentage of 8th-grade public school students who had a math teacher with more than 5 years experience
 - 70% of Black students vs. vs. 78% of White students

Structural Inequities Have Impacted

- Parents' education level
- Wealth
- Crime-victim status and criminal justice
- Health
- School segregation
- Public-school resources
- Teacher quality

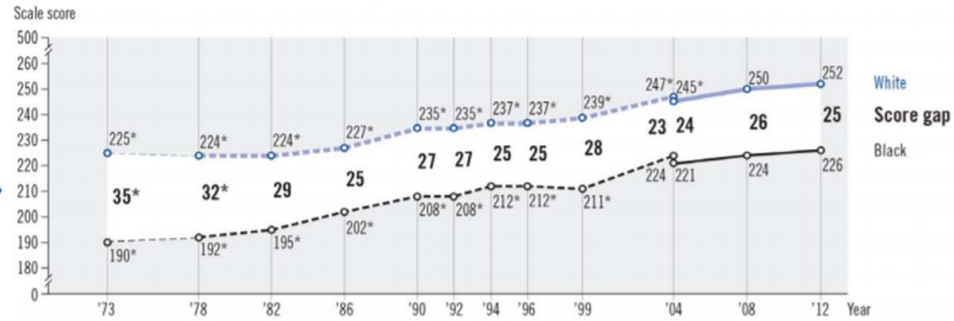
Structural Impacts Have Created

- Dramatic differences in opportunities to learn
 - In school
 - At home
 - In the community



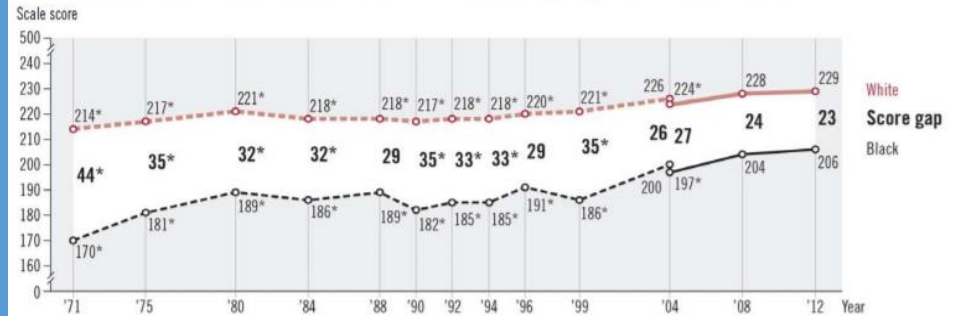
Score Differences Over Generations

Figure 23. Trend in NAEP mathematics average scores and score gaps for White and Black 9-year-old students



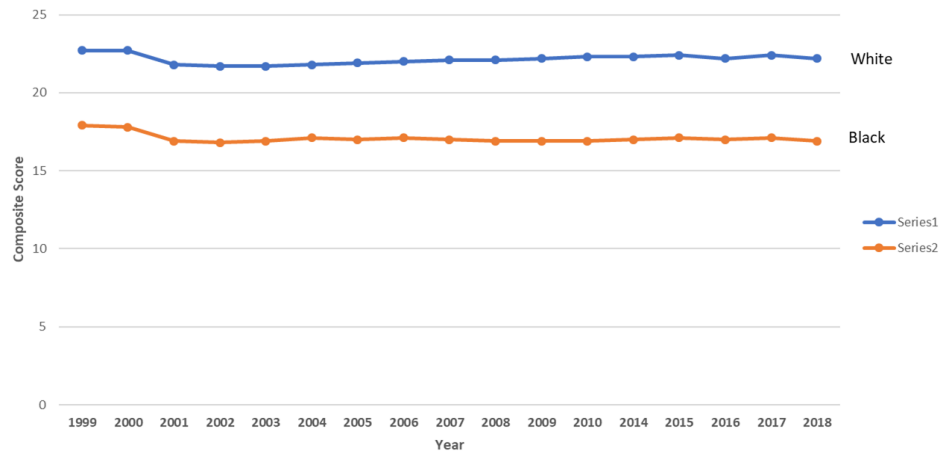
* Significantly different ($p < .05$) from 2012.
 † Extrapolated data adjusting for the limited number of questions from the 1973 mathematics assessment in common with the assessments that followed.
 NOTE: Black includes African American. Race categories exclude Hispanic origin. Score gaps are calculated based on differences between unrounded average scores.

Figure 7. Trend in NAEP reading average scores and score gaps for White and Black 9-year-old students

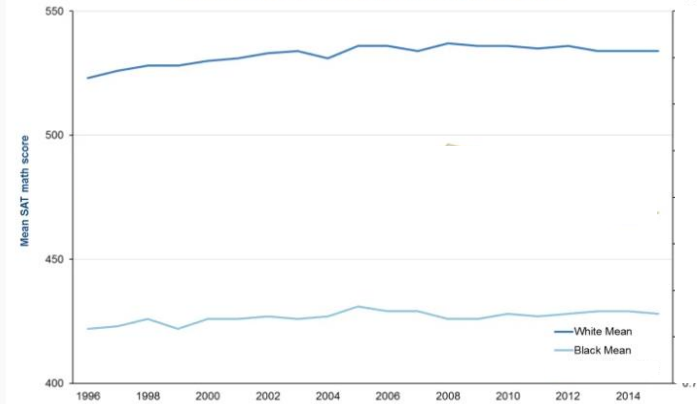


* Significantly different ($p < .05$) from 2012.
 NOTE: Black includes African American. Race categories exclude Hispanic origin. Score gaps are calculated based on differences between unrounded average scores.

ACT Score by Year, 1999-2010 and 2014-2018



Black-white SAT math achievement gap over time



Source: Compiled from Annual Total Group Profile Reports produced by the College Board.

Score and GPA Differences Over Generations

Figure 27. Trend in NAEP mathematics average scores and score gaps for White and Black 17-year-old students

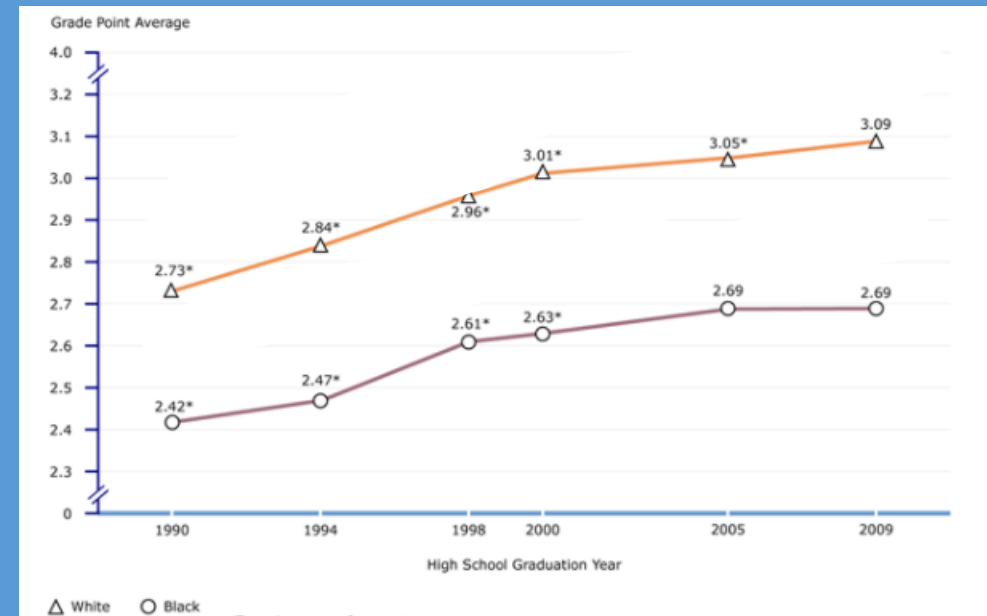
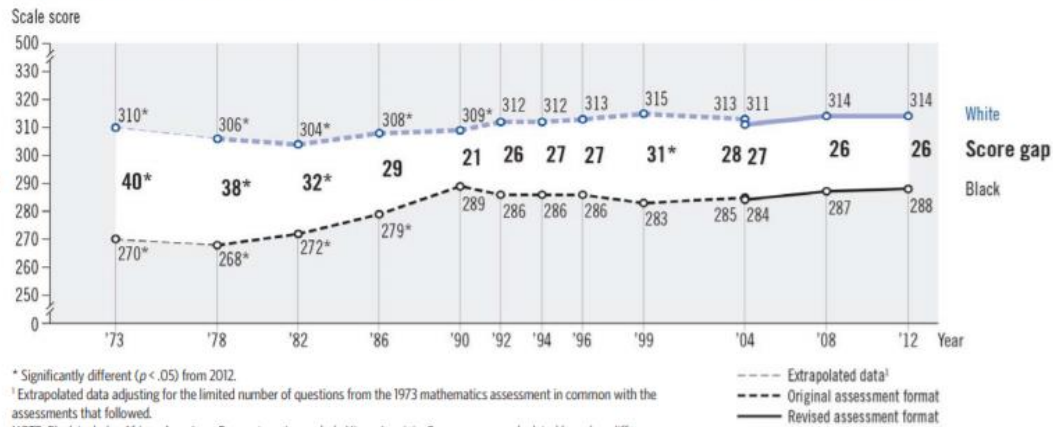
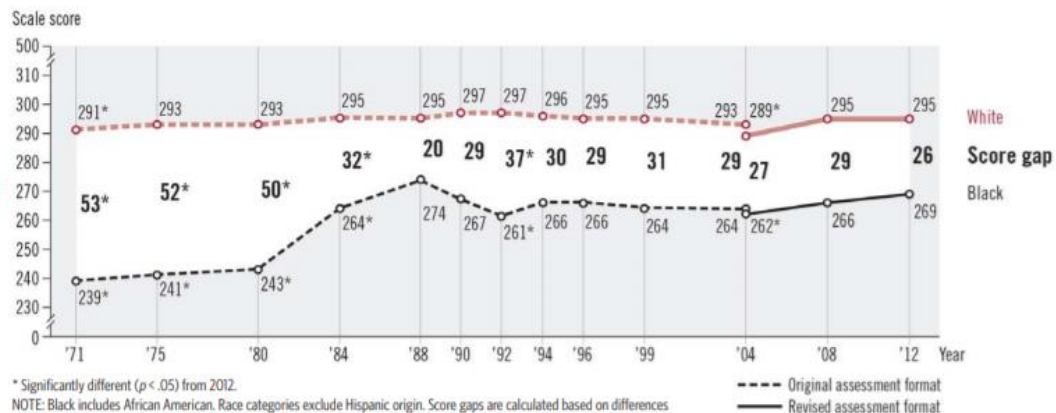


Figure 11. Trend in NAEP reading average scores and score gaps for White and Black 17-year-old students



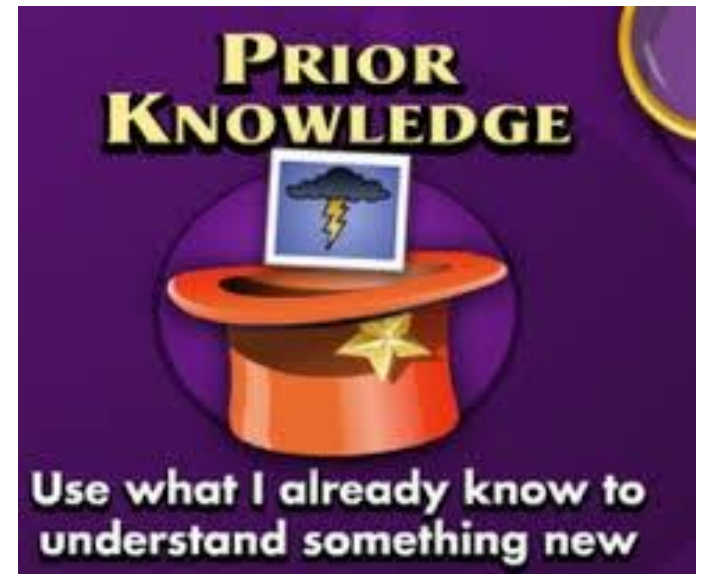
Prior Knowledge

- Includes:
 - Academic knowledge
 - Meta-cognitive skills, domain-general and domain-specific strategies, domain procedures and facts
 - General knowledge
 - Social understanding
 - Cultural knowledge and identity
- Built from experience
 - School
 - Home
 - Community



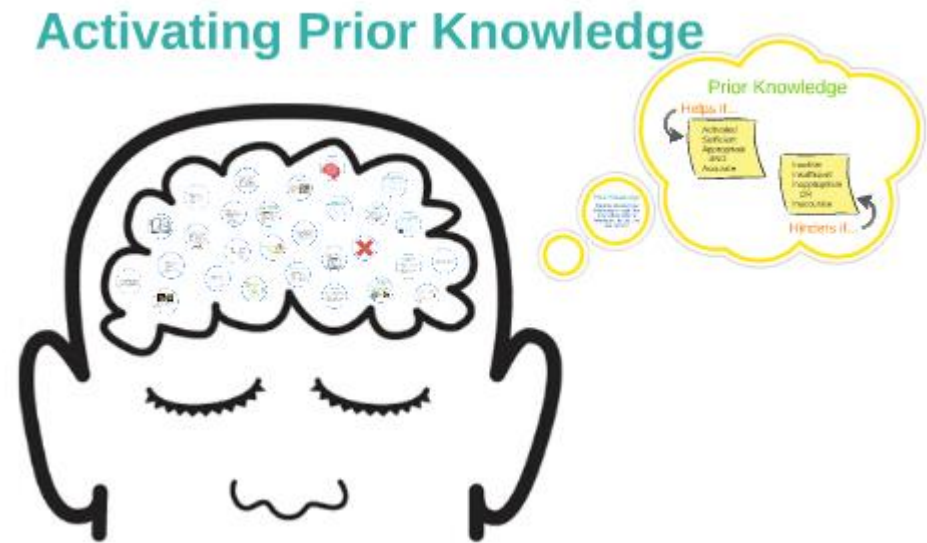
Prior Knowledge

- May be facilitative for academic learning to the extent:
 - It is consonant with knowledge, ways of knowing, practices, and identities valued in school
 - Teachers explicitly connect instruction to the prior knowledge of students
 - Especially if prior knowledge differs from that valued by the school



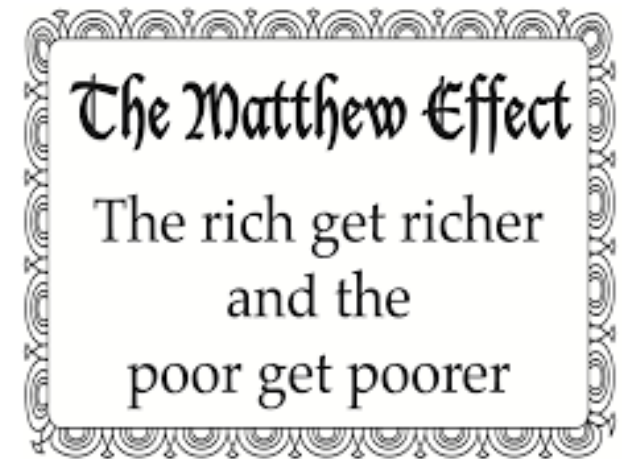
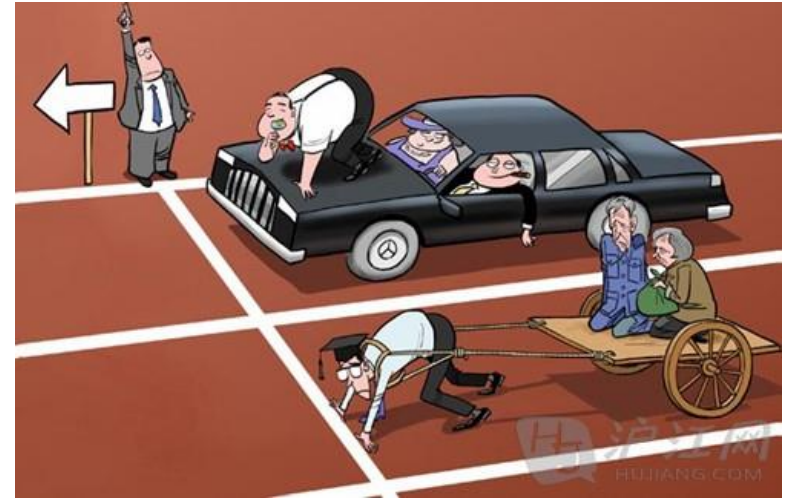
Prior Knowledge

- Key to learning (NASEM, 2018; NRC, 2000)
 - New knowledge is constructed
 - Associated with, elaborated on, or reformulation of existing knowledge



Prior Knowledge

- “Matthew Effects” (Walberg & Tsai, 1983; Stanovich, 1986)
 - The more knowledge one has, the easier acquiring new knowledge becomes
- Limitations in experience provided by school, home, or community restrict development of students’ facilitative prior knowledge
- Limitations in all those sources put a child at *extreme* disadvantage



The “Achievement Gap”

- The existence and persistence of the achievement gap:
 - A consequence of severe restrictions on development of students’ prior knowledge
 - Accumulated over generations within the home and community
 - Reinforced through social, political, legal, educational, and business structures
- Better conceptualized as, in some significant part, an opportunities gap



Test Scores and OTL

- A test score reflects what a student has learned, given the opportunities they have had to learn
 - Interaction between opportunity and what they have been able to make of that opportunity
 - Not just OTL in school
 - Not just OTL a student has received this year but back in time
- OTL as a mechanism through which structural inequity operates on achievement as manifested in test scores and grades


Testing and Structural Inequity: Higher Education


- If admissions tests are in some significant part a measure of OTL, is it fair to privilege them in decisions about awarding further OTL?
 - Does that use make such tests part of the structure of inequity?
 - UC system eligibility: test score and HSGPA *only*
 - Specific campus admission: test score, HSGPA, and 12 additional factors





Testing and Structural Inequity: Higher Education

- What might be alternative approaches and what are their advantages and liabilities?
 - Test optional/test blind policies
 - Greater use of HSGPA
 - Performance in college-level courses taken in high school
 - Measures of personal characteristics
 - Lotteries
 - Holistic admissions


Landscape™ Consistent high school and neighborhood information for colleges 

 Landscape™ helps admissions officers fully consider every student, no matter where they live.


 Landscape is only one element of an application.

 Landscape does not include any personal information beyond an applicant's test score.

Three Categories of Information

**Basic High School Data**

- Locale (e.g., Rural)
- Senior class size
- Percent of students eligible for free and reduced-price lunch
- Average SAT® scores at colleges attended
- AP® participation and performance


**Test Score Comparison**

50th Percentile

25th Percentile

75th Percentile

Applicant's test score compared to others from the same high school

**High School and Neighborhood Indicators**

- College attendance
- Household structure
- Median family income
- Housing stability
- Education levels
- Crime

Research shows these indicators are related to students' education outcomes. Two averages are created based on these six indicators. Values are shown on a scale from 1 to 100 relative to the U.S. average.

© 2019 College Board

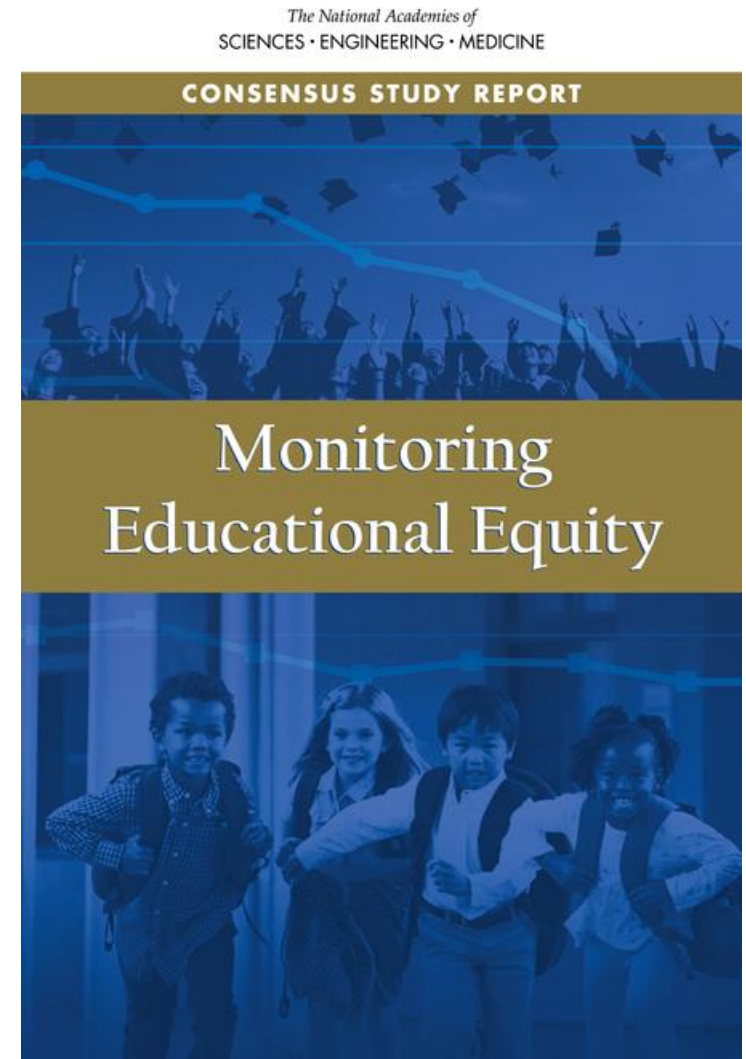
Retaining Diverse University Cohorts

- Expand student academic support services
 - UC STTF recommendation
- Can't just be academic support traditionally defined
 - Peer and faculty mentoring
 - Financial counseling
 - “Summer Bridge”
 - Housing during vacations and breaks
 - Multicultural clubs
 - Courses, majors, and minors that signal respect for diversity and inclusion
 - Diverse faculty



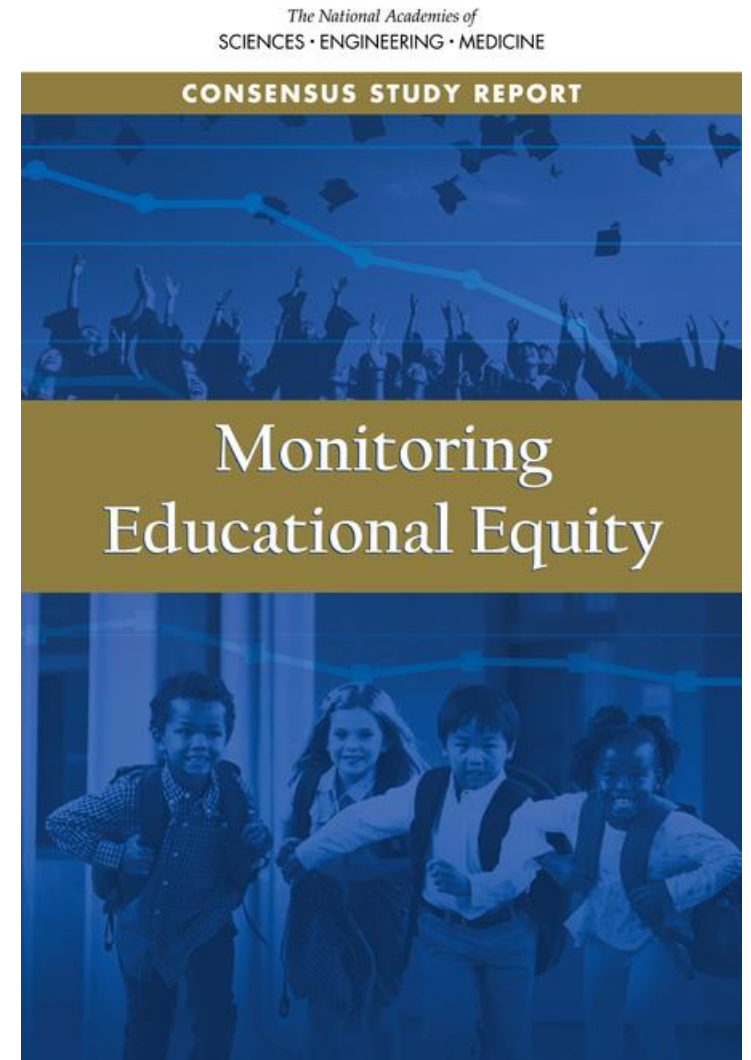
Testing and Structural Inequity: K-12

- If state assessments (and NAEP) in significant part reflect OTL, shouldn't we be tracking other outcome measures that provide relevant information to complement test scores?
 - National Academies *Indicators of Educational Equity* (NASEM, 2019)
 - Indicators of Disparities in Student Outcomes
 - Kindergarten academic readiness
 - Self-regulation and attention skills
 - Engagement in schooling
 - Performance in coursework
 - *Performance on tests*
 - On-time graduation
 - Postsecondary readiness



Access to Educational Opportunities

- National Academies *Indicators of Educational Equity* (NASEM, 2019)
 - Indicators of Disparities in Access to Educational Opportunities
 - Students' exposure to racial, ethnic, and economic segregation
 - Access to and participation in high-quality pre-K programs
 - Access to effective teaching
 - Access to and enrollment in rigorous coursework
 - Curricular breadth
 - Access to high-quality academic supports
 - School climate
 - Non-exclusionary discipline practices
 - Nonacademic supports for student success



Equalizing Opportunity: Changing Educational Practice

- NCLB (2002)
 - State responsibilities
 - Establish content and performance standards
 - Annually test students against standards
 - Report results disaggregated by demographic group at the school-building level
 - Designate schools according to performance and levy sanctions
 - Bring all students to proficiency
 - Theory of action
 - Identifying poorly performing schools and holding their educators accountable would:
 - Spur teaching and learning aligned to standards for all students
 - Reduce gaps
- ESSA (2015)
 - Relaxed some accountability requirements
 - Retained a focus on annual testing



Equalizing Opportunity: Changing Educational Practice

- Culturally responsive teaching (Gay, 2000, 2018):
 - Using the cultural characteristics, experiences, and perspectives of ethnically diverse students as conduits
 - Situating academic knowledge and skills within the lived experiences and reference frames of students makes learning more personally meaningful
 - Comprehensively including ethnic and cultural diversity content in the curriculum
 - Developing a knowledge base about cultural diversity
 - Teaching in culturally congruous ways



Support for Improving Equitable Teaching and Learning

- Describe and use a wide range of accountability measures to initiate improvement actions that help teachers instruct and students learn
 - Deep-dive, development-oriented audit and action plan done collaboratively with school staff
 - Whole-school intervention
 - Coaching in leadership skills
 - Mentoring in culturally responsive teaching
 - Demographically matching identified schools with more successful schools
 - Demonstration proof
 - Transplant effective practices
 - Making connections with philanthropic partners, local corporations, and other knowledge or financial resources to build an improvement ecosystem



Standardized Tests

- Bring the sociocultural influences of creators, raters, and those who interpret results
- Those influences and perspectives may not align with the sociocultural background of all students
 - Misalignment means that we must change the:
 - *Test* by creating an experience more attuned to students' sociocultural backgrounds
 - *Students* by familiarizing them with the tools, knowledge representations, and sociocultural characteristics of the test
 - *Interpretation* of results for the individuals or group(s) for which misalignment exists



US Public-School Student Race/Ethnicity and SES (2018)

- Race/Ethnicity
 - 47% White
 - 27% Hispanic
 - 15% Black
 - 5% Asian
 - 6% Other
- SES
 - 52% eligible for free or reduced-price lunch



Source: NCES, 2020, Tables 203.70 and 204.10

Design Assessment to Promote Equity

- Can we construct measures that value this diversity?
 - Build on the cultural knowledge and identity that diverse students bring to school
 - We know how to make assessments “born accessible”
 - Can we make assessments suited to the variety of sociocultural contexts from which students come?
 - “Born socioculturally responsive”



Socioculturally Responsive Assessment

- Draw from multiple research literatures
 - Culturally responsive teaching
 - Socioculturally responsive classroom assessment
 - Multicultural assessment
 - Second-language testing
 - International survey assessment
 - Universal design
 - Learning sciences
 - Measurement
- Draw from assessment systems in other pluralistic countries

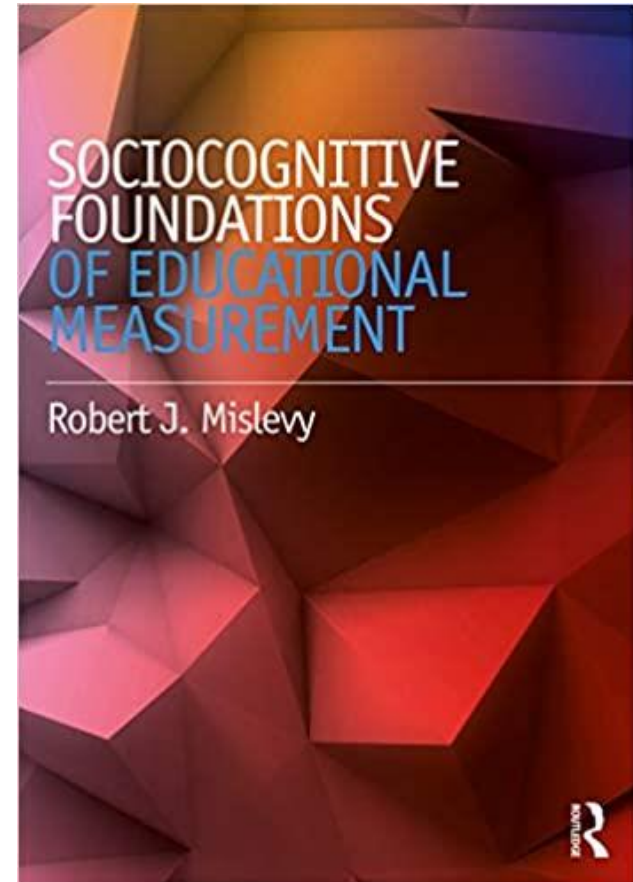
Socioculturally Responsive Assessment

- Present reasonably realistic problem situations that connect to learner experience, culture, and identity
- Allow for multiple forms of representation/expression
 - In problem stimuli
 - In responses
- Promote instruction for deeper learning by design
 - Calls for performances
 - Allows some degree of student agency
 - Provides consultative resources
 - Reinforces habits of mind typical of proficient performers



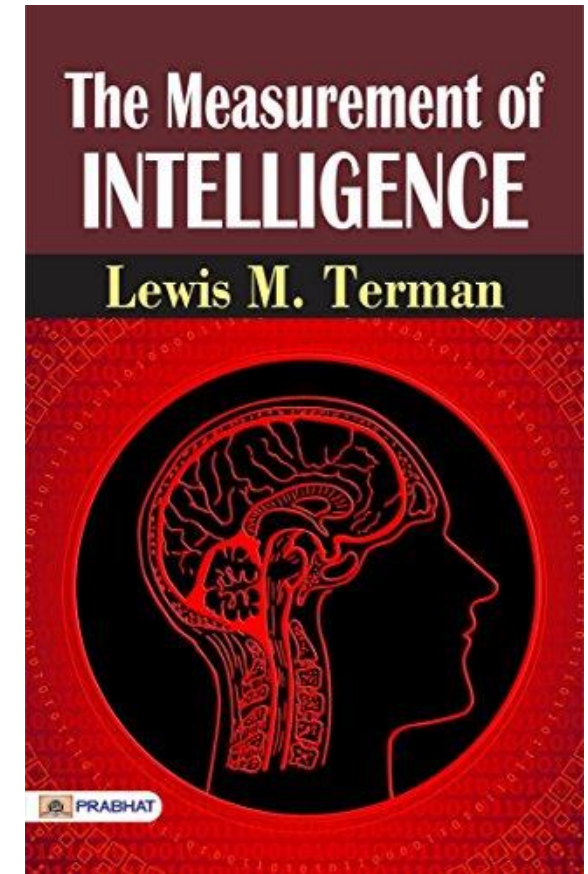
Socioculturally Responsive Assessment

- Adapt to student characteristics
 - Competency level
 - Needs associated with disability
 - Interests and prior knowledge
- Conditional fairness (Mislevy, 2018)
 - A more principled basis for adapting to the examinee
 - Tasks less comparable across examinees
 - Evidence more comparable



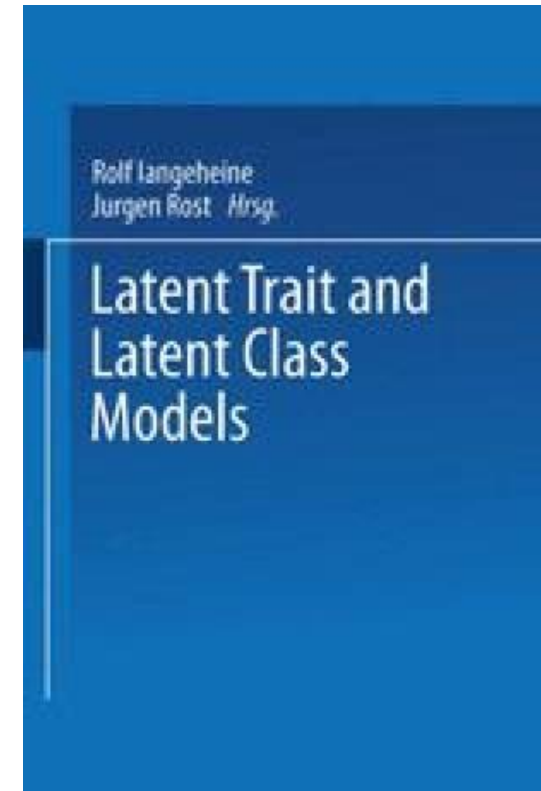
Reframing How We Think About Test Scores

- Long history of test scores as indicators of personal characteristics
 - Founding Eugenicist interpretations (c1900 - 1920s)
 - Terman, Goddard, Brigham, Yerkes
 - Intelligence as an immutable, genetic characteristic unequally distributed across demographic groups
 - Test scores as an indicator



Reframing How We Think About Test Scores

- Trait interpretations (c1900 – present)
 - A pattern of behavior, relatively stable over tasks, time and situations, taken to indicate an hypothesized personal characteristic
 - Intelligence, aptitude, knowledge, skill, personality
 - No Eugenicist assumptions
 - IRT
 - Formerly Latent Trait Theory
 - “Ability”



Reframing How We Think About Test Scores

- Construct interpretations (c1960s – present)
 - Added a theoretical account of the trait's relationships to other variables
 - Validation: an evaluation of both the measure and the underlying theory *of the trait*
- Common to trait and construct interpretations is an underlying characteristic resident *in the person*



Reframing How We Think About Test Scores

- Residue of the Founders lingers because:
 - Test design and content don't *value* diversity
 - Substantial racial/ethnic and SES score differences persist
 - Differences affect life chances
 - We use language that personalizes performance gaps



Reframing How We Think About Test Scores

- Sociocognitive interpretations (Mislevy, 2018)
 - Suggest performance is an *interaction* among person, their history, tasks, and the environments in which they perform
 - Performance is “situated”
 - Not *in* the person
 - In the shared space inhabited by person, their history, tasks, and the environments
 - Performance as a conditional propensity



Reframing How We Think About Test Scores

- Sociocultural interpretations
 - Suggest considering the identities, knowledge, ways of knowing, and the practices valued in students' families and communities
 - Poor performance may reflect a disconnect between the values of the test and student, family, and community
 - Test as a less-familiar cultural artifact
 - How does the student perform in classes taught in culturally responsive ways?
 - How does the student perform in home and community contexts that call for similar outcomes?



Reframing How We Think About Test Scores

- Should we be reframing our thinking and communications about test scores:
 - From
 - Indicators of constructs resident in people
 - To something
 - Centered more around the interaction of person (or group), history, tasks, and environments
 - Taking into account:
 - OTL
 - Other outcome indicators

Summary

- COVID-19 accelerated trends already underway
 - Reconceptualizing K-12 accountability assessment
 - Adopting test-optional postsecondary admissions policies
- Structural inequity is longstanding and pervasive
- OTL as generational and broad-based
 - School, home, community
 - A mechanism through which structural inequity affects achievement, helping create large group differences in test scores, grades, and life chances

Summary

- Implications for postsecondary admissions
 - Trying and evaluating alternative admissions approaches
 - Meet diversity goals
 - Expanding student services
 - Maintain diverse admitted classes



Summary



- Implications for K-12 assessment
 - Collecting broad-based indicators of student outcomes
 - Gathering various indicators of OTL
 - Investing in culturally responsive teaching
 - Conceptualizing and enacting accountability as support for improving equitable teaching and learning
 - Creating assessments “born socioculturally responsive”
 - Reframing how we think about test scores



Equity and Assessment in the Post-COVID-19 Era

Randy Bennett

Educational Testing Service
Princeton, NJ 08541
rbennett@ets.org

E.F. Lindquist Award presentation at the virtual annual meeting of the American Educational Research Association, April 2021