DISMANTLING THE LOGICS OF EUGENICS VIA EMANCIPATORY DATA SCIENCE EDUCATION

Thema Monroe-White
George Mason University
tmonroew@gmu.edu

C. Malik Boykin
Brown University
c_boykin@brown.edu

JaNiya Daniels
Berry College
janiya.daniels@vikings.berry.edu

Statistics, including Data Science, originates from eugenicist roots. Galton, Pearson, and Fisher—known as the 'founding fathers' of statistics—were also eugenicists. The rise of eugenics coincided with legalized mass sterilization efforts across Europe and 27 U.S. states between 1907 and 1931, targeting criminals, individuals with epilepsy, and the "feebleminded." Statistics, driven by the political agenda of its founders, was the primary tool used to advance the eugenics discipline. Although some scholars in statistics and data science education have worked to identify and expose these eugenics origins, less is known about those who resisted this ideology. Using counter-racial archival analysis, we uncover the work of those who utilized data and statistics to counter these dehumanizing ideologies, by uplifting those targeted.

BACKGROUND

Statistics and by extension data science, originate from eugenicist roots. Eugenics is an ideology and set of practices which aim to increase 'desirable' human traits while reducing those perceived as 'undesirable.' Galton, Pearson, and Fisher—known as the 'founding fathers' of statistics—were also eugenicists. Galton, in 1892's Hereditary Genius, stated,

> There is nothing in the history of domestic animals or in that of evolution to make us doubt that a race of sane men may be formed, who shall be as much superior mentally and morally to the modern European, as the modern European is to the lowest of the Negro races. (p. x)

The establishment of the U.S. Eugenics Records Office in 1910 and the American Eugenics Society in 1922 coincided with legalized mass sterilization efforts across Europe (including Nazi-era Germany) and 27 U.S. states between 1907 and 1931, targeting criminals, individuals with epilepsy, and the "feebleminded." Statistics, driven by the political agenda of its founders, was the primary tool used to advance the eugenics discipline. The most salient expressions of this this ideology have manifest themselves in a variety of biological and social scientific fields including, and perhaps most notably, the fields of education, psychology, and sociology via the use and misuse of standardized testing.

Questioning the Infallibility of Standardized Testing

In 1967, the newly formed Association of Black Psychologists (A.B. Psi) submitted a list of demands, which included an investigation into the misuse of standardized tests they asserted were differentially valid for Black and White test takers and thus biased against Black students in the graduate student selection process (Williams, 2008). Unheard in their demands, in the following year, 1968, a group of Black psychology Ph.D. students and members of A.B. Psi, including Charles Thomas and Robert Williams, stormed the presidential address at the annual meeting of the American Psychology Association (Williams, 2008). They took to the microphone to lodge the same complaint. It is likely not a coincidence then that the next iteration of the published professional standards for educational and psychological testing (published in collaboration between the American Education Researchers Association, American Psychological Association, National Council for Measurement in Education), for the very first time, communicated that test developers must consider whether their standardized tests demonstrated biased against minoritized groups (Sireci & Randall, 2021).

In: Kaplan, J. & Luebke, K. (Ed.) (2024). *Connecting data and people for inclusive statistics and data science education. Proceedings of the Roundtable conference of the International Association for Statistics Education (IASE), July 2024, Auckland, New Zealand.*

This catalyzed debates regarding multiple definitions of bias and fairness in standardized testing in relation to scores and predictions across members of EEOC-protected groups and informed policies of how testing is governed and implemented. As well, it begat a statistical retrofitting of standardized tests designed and statistically normed to demonstrate White superiority and to select White people for an opportunity structure aimed to serve and advance White students. Barocas, Hardt, and Narayanan (2019) surveyed the number of the bias definitions regarding education testing derived from these education testing debates in the 1960s and 1970s are statistically parallel to the kinds of linear model-derived definitions of bias and fairness metrics used to detect bias in algorithmic decision systems (see Clearly, 1968; ____ 1966; Corbett-Davies et al., 1971). Frustratingly, the professional standards for educational and psychological testing's 1974 acknowledgment of bias in these decision systems comes 50 years after Horace Man Bond, in a 1924 article published in the Crisis (a news magazine published by the NAACP), lodged similar complaints at the standardized testing industry (Thomas, 1982). In his article, Bond (1924) leveraged a critical disaggregated analysis of the data from the Army's Alpha Intelligence test to show that White people in the North scored higher than White people in the South, that Black people in the North scored higher than Black people in the South, and that Black people in the North scored as high or higher on average than White people in the South (Bond, 1924; Thomas, 1982). In what amounts to auditing the results of a selection system that had deleterious effects on Black people and on White Southerners, Bond found both racial and regional bias in a test that designed to measure the amount of intelligence latent within an individual. He asserted that knowledge of the mechanics of standardized testing should be of great interest to every Black ("Negro" at the time) intellectual. His rationale was that as long as Black people remain unaware of how these systems work and how they are inflicting structural harm on Black communities we will remain defenseless in our fight for liberation (Bond, 1924). Bond recognized and named these statistical practices to devalue Black people and their concerns as dehumanizing, and nearly a century later, Black people are entrenched in the same fight for humanity. Bond's call to action regarding learning the mathematical mechanics and theoretical assumptions of standardized testing's statistical eugenic arguments is similar to the need for Black people to interrogate algorithms today for emancipation. This perspective is even more important to consider given that the likelihood of elevating *perceptions of the legitimacy of standardized testing* is contingent in part upon how well members of one's own racial group perform on the test itself (Unzuetta & Lowery, 2010).

Questioning the Infallibility of Algorithms

The perceived legitimacy of an algorithm, whether to deploy it, and whether it is fair is contingent on how the algorithm's fairness metrics perform for the race of the people who oversee the algorithm's development and implementation. Historically, algorithm developers have largely been White and male with little training in historical and structural racism, sexism, or other biases (Monroe-White, 2021); such that the people developing algorithms disproportionately account for the wants and needs of their own communities. The case of the gender shades algorithms is a clear example of failure to consider how the algorithm works for darker-skinned females, while prioritizing algorithmic performance for lighter-skinned males (Boulamwini & Gebru, 2018). Waytz and Shroeder (2014) described the process of failing to consider the wants and needs of people who are perceived as irrelevant to one's own goals and outcomes as dehumanization by omission. We agree. The demonstrated acts of failing to disaggregate data across racial groups to ensure fairness in performance or failing to include enough minoritized groups in the training data to help improve algorithms' capacities to perform optimally when making decisions about them is dehumanization by omission. Unzuetta and Lowery's (2010) work demonstrated that people's perceptions of the legitimacy of a standardized test are partly contingent upon how well members of their own racial group perform on the test. We argue that the perceived legitimacy of an algorithm, whether to deploy it, and whether it is fair is contingent on how the algorithm's fairness metrics perform for on the race of the people who oversee the algorithm's development and implementation. This form of dehumanization is especially pernicious in instances where algorithms can do serious harm to minoritized populations and exacerbate structural inequalities in the form of failing to detect health risks and follow-on treatments (Oberymeyer et. al., 2019).

METHODOLOGY & RESULTS

Scholars in statistics and data science have worked to identify and expose the eugenics origins of the field; less is known, however, about those who resisted these ideologies. Using counter-racial archival analysis, we uncovered the work of organizations and scholars who utilized their knowledge of data and statistics to counter these dehumanizing ideologies, choosing instead to humanize data and data systems for empowerment. Specifically, this study provides descriptive profiles—answering the questions of who, what, when, where, why, and how—of historical and contemporary scholars who have used their quantitative and computational expertise to combat eugenics ideologies. The list below is incomplete; however, it highlights the works for a handful of prominent thought leaders who fall under the umbrella term "emancipatory data scientists" who use their quantitative and computational skills to uplift members of minoritized communities (Monroe-White, 2021).

Horace Mann Bond

Bond, a prominent social scientist, educator, and champion of racial equality, pioneered the use of data to dismantle racist narratives prevalent in early 20th-century America (Urban, 1989). He was particularly focused on the insidious ways biased educational practices reinforced and perpetuated inequality. Bond meticulously collected and analyzed data on educational disparities between Black and white students. He exposed the vast inequalities in school funding, teacher qualifications, and access to resources that severely limited the educational opportunities for Black Americans. This data-driven approach systematically dismantled the myth of inherent Black intellectual inferiority, a harmful ideology used to justify segregation and systemic disadvantage. Furthermore, Bond's work with intelligence testing was groundbreaking. He challenged the misuse of IQ tests frequently administered in ways that produced lower scores for Black students. He argued these tests were culturally biased and did not accurately measure the potential of Black students. His insights played a key role in highlighting the flaws within this type of testing and its potential to reinforce racist assumptions and limit opportunities (Norton 1984).

Alicia Martin

Martin's work exposed how the underrepresentation of diverse populations in scientific datasets can distort results. The underprediction of West African height using White-derived genomic models illustrates the dangers of ignoring variations across populations and applying standards built on one group as universal. Franz Boas's pioneering arguments for cultural relativism challenged this type of ethnocentrism; his insistence that people's values and behaviors can only be understood within their societal context resonates with the fight to dismantle systemic bias within data science (Verdon 2006).

Joy Buolamwini

A computer scientist and founder of the Algorithmic Justice League, Buolamwini, has been a leading voice in exposing racial and gender bias within artificial intelligence (AI), particularly facial recognition technology. Her groundbreaking work co-authored with Timnit Gebru revealed that commercial facial recognition systems had significantly higher error rates for dark-skinned women compared to light-skinned men. This glaring disparity highlighted how the lack of diversity in training datasets used to build these systems perpetuated dehumanizing biases with very real consequences. Buolamwini's research didn't stop at revealing the problem. She used her data analysis to ignite global conversations about algorithmic bias and its potential harms. Her powerful TED Talk, "How I'm Fighting Bias in Algorithms," brought this issue to mainstream attention, illustrating how flawed facial recognition could disproportionately lead to misidentification and false arrests for people of color. Through the Algorithmic Justice League, she advocates for greater transparency, accountability, and the inclusion of diverse perspectives in the development of AI technologies. Her data-driven approach has been influential in pushing technology companies and policymakers to address the biases embedded within their algorithms.

Tukufu Zuberi

Zuberi authored "Thicker Than Blood: How Racial Statistics Lie" in 2003 to reveal the complexities and inaccuracies surrounding racial statistics. Zuberi argued that racial categories are socially constructed and often fail to capture the nuanced realities of human identity and experience. He highlighted how these statistics can perpetuate harmful stereotypes and reinforce inequalities by oversimplifying diverse populations into homogenous groups. Zuberi illustrated how racial statistics can be misleading, citing examples where data fails to accurately represent the lived experiences of individuals within racial groups. He challenged readers to examine critically the ways in which race is conceptualized and measured, advocating for a more nuanced understanding that acknowledges the intersections of race, class, gender, and other social factors. Ultimately, Zuberi contended that relying solely on racial statistics perpetuates a flawed narrative of racial essentialism and undermines efforts to address systemic inequalities.

The Association of Black Psychologists

ABPsi has demanded the American Psychological Association (APA) establish a committee to investigate the misuse of standardized psychological tests that result in the denial of educational and economic opportunities for Black students. The ABPsi argued these tests lack validity and should not be used until a thorough review and reassessment are conducted. They expressed deep concern about the use of standardized testing as a tool of discrimination against Black students, asserting that the tests are biased and fail to accurately reflect their true abilities. The ABPsi has insisted on a moratorium on testing until new, fair, and culturally sensitive tests can be developed.

Data for Black Lives

Data for Black Lives (D4BL; https://d4bl.org/) stands as a powerful force using data and statistics to challenge dehumanizing ideologies that perpetuate systemic racism. At the core of their work lies the collection and rigorous analysis of data related to arrests, incarceration, and police use of force. Through visualizations and reports, D4BL has exposed the undeniable racial disparities within the criminal justice system. Statistics showing significantly higher incarceration rates for Black Americans compared to white Americans, even for similar offenses, directly contradict harmful narratives that justify mass incarceration or portray Black communities as inherently criminal. Beyond revealing injustice, D4BL actively combats the dissemination of racially biased. They have countered incomplete datasets and cherry-picked statistics often used to reinforce negative stereotypes against Black communities. D4BL's analysis has provided a more comprehensive understanding by including socioeconomic factors alongside crime data. This approach challenges the simplistic equation of increased policing with increased safety and pushes for addressing the underlying causes of crime. D4BL's impact extends from awareness to advocacy. Their data analysis forms the foundation for collaboration with policymakers and grassroots organizations, leading to evidence-based recommendations for systemic reform. They have pushed for policies that promote fairer policing, reduce harmful incarceration rates, and advocate for targeted investment in community-based programs addressing poverty and inequality. Data for Black Lives has demonstrated how data science, when used with integrity and purpose, can dismantle dehumanizing ideologies and pave the way for a more just and equitable society (Benjamin 2019).

Native Land Digital

Native Land Digital (https://native-land.ca) is an Indigenous-led, Canadian not-for-profit organization that "stives to create and foster conversations about the history of colonialism, Indigenous ways of knowing, and settler-Indigenous relations, through educational resources such as our map and Territory Acknowledgement Guide." Their interactive map allows the online users to see and investigate the territories, languages, and treaties accumulated over decades to display a fuller and richer representation of Indigenous people across the globe.

EMANCIPATING DATA SCIENCE EDUCATION

Emancipating data science means coming to terms with and atoning for the whitewashing of the discipline, recognizing that the historical horrors of eugenics continue to influence subtly scientific studies and practices. For example, Shakespeare (2006) argued the lingering influence of eugenics can be seen in the continued emphasis on finding genetic cures for disabilities. This focus, he suggested, reinforces the notion that a disability is an inherent flaw or disease in need of eradication, rather than a social construct influenced by societal barriers. He pointed to prenatal testing that flags potential disabilities as a concerning outcome, subtly pushing expectant parents toward termination. This focus on genetic solutions also diverts attention away from addressing the real challenges faced by disabled people. Instead of investing in inclusive education, accessible infrastructure, and employment opportunities, resources are directed towards finding a 'fix' for the disabled individual, perpetuating a narrative of disability as a personal tragedy in need of a medical solution (Shakespeare 2006). Likewise, in her book, Fatal Invention, Dorothy Roberts (2012) warned that the resurgence of scientific focus on race, fueled by advances in genetics, risks reviving eugenic ideologies under the guise of objectivity. She noted how the search for biological explanations for racial health disparities or the development of race-specific drugs unintentionally reinforces the idea of race as a fixed, biological category (Roberts 2012). This trend, if allowed to continue would validate preexisting biases, placing undue responsibility on individuals and minimizing the immense impact of structural inequalities on health outcomes. Furthermore, Roberts critiqued the potentially harmful implications of race-based medical practices. She argued that even if well-intentioned, the use of race-specific treatments could create new forms of medical discrimination and obscure the need to address the root social and economic causes of health disparities. Her work exposes how seemingly neutral scientific advancements can perpetuate eugenic thinking if researchers and healthcare providers

Again, we draw on the work of Unzuetta and Lowery (2010) in which they found that the perceived legitimacy of standardized tests often depends on how well one's own group performs. Furthermore, it is harder for people to care about bias even when it leads to disparate health outcomes, incarceration, and death for minoritized groups (Richeson, 2020). There is a concerning parallel with algorithms: if the dominant demographic within algorithm development prioritizes models that work well for them (i.e., White males), there is a risk of endorsing biased systems that underperform when applied to diverse populations. This systemic undervaluing of diverse perspectives can go unnoticed due to insufficiently powered models that fail to highlight differences. The fight for unbiased data science is paramount. We must combat these issues by actively including a wide range of voices and perspectives throughout the model development process. Ensuring algorithms are tested rigorously across diverse populations and held accountable for fairness is crucial to avoid the harmful dehumanization that results from unchecked bias (Monroe-White & Lecy, 2023).

Consider the case of simple linear regression. The goal of the model is to determine one equation that represents the majority of data points on a two-dimensional plane. These data points can represent various things, including people. Outside the digital sphere, however, all “data points” are not equal. As people, we are not the same, and achieving sameness is not necessarily a desirable goal. Discussing the disparate impacts on data points positioned close to and far from the regression line is a step toward liberating our understanding of data and the algorithms that use them. Although linear regression instruction may cover its limitations and appropriate use cases, a fuller analysis of those limitations (including how they cascade into broader use cases) is often left for the data professional to figure out. This siloing of experiences, combined with unilateral decision-making authority, leads to inconsistent data practices. It perpetuates unaccountable bad behaviors and codifies harm to people in vulnerable communities.

The ways in which we present and explain concepts, algorithms, and models associated with data science need to be contextualized for the multi-racial, multi-gender, multi-class, multi-abled data workforce we want to create and retain. For topics that span the data lifecycle, data visualization can serve as a powerful demonstration of context and nuances of data understanding. For example, a class project focused on reproducing a portion of the Native Land Digital map may unlock the heavy decision-making choices data practitioners make. Data instruction that centers this ongoing project can cover data collection, cleanup, storage, analysis and visualization methods, and tech stack choices. Additionally, data governance, stewardship, and ethics protocols and implications arise as learners grapple with what they’ve been previously taught and how this data broadens their perspectives of Indigenous people’s contributions to American life.

CONCLUSION

In the journey towards equity in data science education, it becomes evident that the historical biases ingrained in algorithmic development echo the systemic discrimination witnessed in standardized testing and other tools used to identify “undesirable” people. The struggle for recognition and rectification of these biases has been spearheaded by marginalized voices throughout history, from the groundbreaking efforts of Black psychologists challenging the misuse of standardized tests to contemporary champions like Buolamwini and Gebru (2018), Noble (2018), and more revealing the racial and gender disparities within artificial intelligence, facial recognition, and search engine technology. As we navigate the complexities of data science, it becomes imperative to confront the whitewashing of the discipline and actively include diverse perspectives throughout the model development process. Failure to do so perpetuates dehumanizing biases, as evidenced by the disparities in algorithmic performance for marginalized communities. Just as Horace Mann Bond pioneered the use of data to dismantle racist narratives in the early 20th century, we must continue to challenge the reliance on racial statistics and prioritize addressing the systemic inequities that underpin our data practices. Only through a concerted effort to contextualize, diversify, and critically examine our approaches can we truly emancipate data science from its historically biased foundations and pave the way for a more just and equitable society.

REFERENCES

——— (1966). Test Bias: Validity of the Scholastic Aptitude Test for Negro and White students in integrated colleges. *ETS Research Bulletin Series 1966*(2), i–23.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. John Wiley & Sons.

Bond, H. M. (1924). Intelligence tests and propaganda. *The Crisis, 28*(2), 61-64.

Buolamwini, J. & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*, in *Proceedings of Machine Learning Research*, *81*, 77-91. https://proceedings.mlr.press/v81/buolamwini18a.html.

Cleary, T A. (1968). Test bias: Prediction of grades of Negro and White students in integrated colleges, *Journal of Educational Measurement 5*(2), 115–24.

Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. *Algorithmic decision making and the cost of fairness.* arXiv Preprint arXiv:1701.08230, 2017.

Darlington, R. B. (1971). Another look at 'cultural fairness'. *Journal of Educational Measurement, 8*(2), 71–82.

Galton, F. (1892). *Hereditary genius: An inquiry into its laws and consequences*. MacMillan and Co.

Monroe-White, T. (2021). Emancipatory data science: A liberatory framework for mitigating data harms and fostering social transformation. In *Proceedings of the 2021 on Computers and people research conference*, Association for Computing Machinery, 23-30. https://doi.org/10.1145/3458026.3462161

Monroe-White, T., & Lecy, J. (2023). The Wells-Du Bois Protocol for machine learning bias: building critical quantitative foundations for third sector scholarship. *VOLUNTAS: International Journal of Voluntary and Nonprofit Organizations*, *34*(1), 170-184.

Norton, R. (1984). The Horace Mann Bond papers: A biography of change. *The Journal of Negro Education*, *53*(1), 29-40.

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, *366*(6464), 447–453.

Richeson, J. 2020. The Mythology of Racial Progress [Conference Presentation]. Society for Personality and Social Psychology. Presidential Plenary: Bias in the Age of AI and Big Data.. https://www.youtube.com/watch?v=S2knhRMZRuI

Roberts, D. E. (2012). *Fatal invention: How science, politics, and big business re-create race in the twenty-first century*. The New Press.

Shakespeare, T. (2006). The social model of disability. *The Disability Studies Reader, 2*(3), 197-204.

Sireci, S. G., & Randall, J. (2021). Evolving notions of fairness in testing in the United States. In B.E. Clauser & M.B. Bunch (Eds.) *The history of educational measurement: Key advancements in theory, policy, and practice* (pp. 111-135). Routledge.

Thomas, W. B. (1982). Black intellectuals' critique of early mental testing: A little-known saga of the 1920s. *American Journal of Education, 90*(3), 258-292.

Unzueta, M. M., & Lowery, B. S. (2010). The impact of race-based performance differences on perceptions of test legitimacy. *Journal of Applied Social Psychology, 40*(8), 1948-1968.

Urban, W.J. (1989), The black scholar and intelligence testing: The case of Horace Mann Bond. *Journal of the History of Behavioral Science, 25*, 323-334

Verdon, M. (2006). The world upside down: Boas, history, evolutionism, and science. *History and Anthropology*, *17*(3), 171-187.
Waytz, A., & Schroeder, J. (2014). Overlooking others: Dehumanization by comission and omission. *Testing, Psychometrics, Methodology in Applied Psychology, 21*(3).
Williams, R. L. (2008). *History of the Association of Black Psychologists: Profiles of outstanding Black psychologists*. AuthorHouse.
Zuberi, T. (2003). *Thicker than blood: How racial statistics lie*. (1st Ed.). University of Minnesota Press.