JOURNAL OF COMPARATIVE RESEARCH IN ANTHROPOLOGY AND SOCIOLOGY

Copyright © The Author, 2018
Volume 9, Number 1, Summer 2018
ISSN 2068 – 0317
http://compaso.eu

Cathy O'Neil (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy,* New York, St. Martin's Press and Virginia Eubanks (2018) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*, New York, Broadway Books

Book review

Alexandra Carter[1]

Providing personal data is something that we do on a regular basis, from filling out governmental forms and documents to indicating preferences and habitual patterns on social media and consumer websites. However, there are many ways in which this data is used that are beyond our control or knowledge. For middle and upper-class individuals, this data collection can make certain things easier, like getting recommendations for different products on Amazon. Some are irritated or mildly fearful as new information comes to light about the erosion of privacy on social media, or shocked by advertisements related to things they may have spoken about on the phone or searched for, but had not explicitly shared on those platforms. However, this form of data mining, along with more traditional forms, such as filling out governmental forms, job applications, and request for public services, is also used for more pernicious and pervasive ends. Data collection, algorithms used to interpret that data, and the effects that these mathematical models have on all strata of American society are the topics of two different but related books: *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (2016) by Cathy O'Neil and *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor* (2018) by Virginia Eubanks. In both, the authors argue that data collection and automation increasingly flatten individuals into groups that drive the wedge between the rich and poor, white and people

[1] Independent researcher, alexcarter@ucla.edu

of color, further apart, both because of the way in which systems are constructed and the impacts that they have on marginalized groups in America.

Cathy O'Neil, a Harvard-trained mathematician, was prompted to write *Weapons of Math Destruction* after her experience working as a quant for the financial firm D.E. Shaw before the financial collapse of 2008. She writes, 'the crash made it all too clear that mathematics, once [her] refuge, was not only deeply entangled in the world's problems, but also fueling many of them,' including the housing crisis, unemployment, and the recession, 'aided and abetted by mathematicians wielding magic formulas.' (p. 2) This experience led her to examine the ways in which mathematical modeling, data collection, and predictive algorithms are used and abused in a wide range of contexts in the United States. As these tools are relied upon more and more to make decisions that have significant impacts on the lives of citizens, the ways in which data are collected, manipulated, and used are increasingly important. O'Neil examines different ways in which data is collected and analyzed and how this further exacerbates racial, class, and political divides in the United States. Her book employs a wide range of examples, including public-school teacher evaluations; criminal sentencing judgements; predatory for-profit college and loan targeting; college rankings and applications; predictive crime mapping and policing; restrictive hiring, scheduling, and retention practices; car insurance policies; and microtargeting of political campaigns. In all of these contexts, O'Neil argues that the data collected and algorithms used for analysis and decision making are flawed because they lack statistical rigor, are untrustworthy models, and use proxies for data that are inherently biased. These flaws, along with the universal way in which they are implemented, have earned them the punny moniker "Weapons of Math Destruction," or WMDs (p. 3).

According to O'Neil, what all of these weapons have in common is their opacity, damage, and scale (p. 31). Models are considered opaque if the participants or subjects are not aware of either the purpose or intent for data collection, or are unaware of the data being collected. Even if the participant is aware that the data are being gathered and for what purpose, the model *itself* may be obfuscated (p. 28). Often, the agency collecting the data will argue that this is either to prevent the system from being "gamed," such as with the IMPACT value-added teacher evaluation model (p. 4), or because the algorithms themselves are considered intellectual property, and are thus protected from public scrutiny (p. 4). In the first example, the algorithm measuring teacher performance is hidden with the express intent to prevent teachers from trying to inflate their scores unfairly. This is because, O'Neil asserts, the purpose of the model is to fire teachers. Therefore, the mathematical model is designed to persecute and punish teachers for "poor performance," which fulfills the next criterion of O'Neil's WMD: it is a model that works against the subject's interest; in this case, maintaining one's job. This is further complicated by the difficulty of measuring and quantifying teacher effectiveness. Here, O'Neil identifies several issues with WMDs that touch on both the mathematical aspects of these models and the ways in which they build bias and injustice into seemingly neutral mathematical algorithms. First, teacher effectiveness is not as easily quantifiable as baseball performance. In baseball, statisticians have data that is rigorous and directly related to performance (p. 17). There are decades of data that have been collected, and are highly relevant to the outcomes they seek to predict, such as batting averages, pitch statistics, and past injuries (p. 17). These can be used to predict future performance with some accuracy, and can accept new data in flexible ways to reflect changing performance. For teacher evaluation, such specific data points do not exist (p 5-7). Engaged, effective, meaningful teaching can be accomplished in many different ways, and can be affected by things far outside of the control of the teacher (p. 6-9). This includes, but is not limited to, a student's socioeconomic status and home environment, learning differences, and previous experience with testing. The data sets for teaching outcomes are also small, as teachers work with 25-35 students per year, creating statistically weak models. In situations where the outcome cannot be measured or predicted directly, a mathematical model has to rely on proxy data for quantifying performance (p. 5-7). One measure is thus student test performance on yearly standardized tests, compared against predictions for yearly gains based on past student performance. The underlying argument for the measure is that a highly effective teacher should enable students to meet or exceed these predicted outcomes. O'Neil did not address the other measures used, but the discussion of use of testing as data is telling. There are multiple weaknesses with this model. First, there is an upper limit to test performance, so high performing students will not demonstrate growth over time if they consistently achieve high scores. In addition, the sample size is too small and changes from year to year as students move through the system. This can result in classes of higher-performing students alternating with groups of lower-performing students, leading to flux in teacher evaluation. For example, O'Neil cites a teacher who received a score of 6 out of 100 for the first year, and a 96 out of 100 for the next, although his teaching methods had not changed from the first to second year (p. 136-138). Although teachers can be highly rated as effective, engaging, caring educators who are responsible for changing the lives of students by parents, other educators, and the students themselves, the algorithm that measures the same teachers will tell a different story, resulting in the removal of these teachers for "poor performance," without providing reasons why or ways they can improve. This example of an algorithm also demonstrates O'Neil's last criterion for a WMD: scale. It both lacks appropriate statistical scale to be an effective model, as it relies in very small and varying data sets from year to year, while also affecting people on a large scale; in this case, teachers across America. As with other WMDs O'Neil discusses, if a model is successful in its stated aims (in this case, removing low-performing teachers), then it can jump into other fields as well, regardless of the errors in the model (p. 17).

This model-by-proxy becomes even more troubling as data that code inherent and historic bias and inequality are used to measure people and deny or provide access to services, jobs, justice, insurance, and other needs. O'Neil discusses the use of credit scores to determine employability and access to car insurance, even though these outcomes are not directly related to the data used to determine eligibility. Search histories and social media tracking not only target people for advertisement, including by predatory loan companies and for-profit institutions, but are also used as proxies for determining credit eligibility and civic and political engagement. Data collected from prisoners to determine recidivism includes information specific to crimes, but also asks for information like neighborhood and run-ins with police before being put in jail that reinforce segregation, as people of color experience higher run-ins with police than white people, reinforcing stereotypes and keeping people of color in jail for longer sentences. In all situations, O'Neil argues that the amount of data collected is vast and is being used to group people into large categories that fall along racial, class, gender, and even mental health lines, providing access and privilege to some while denying it to others.

This is the point at which Virginia Eubanks' book, *Automating Inequality*: *How High-Tech Tools Profile, Police, and Punish the Poor* takes off. Eubanks, who is a professor of political science at University of Albany, SUNY, was prompted to write her book after her own experience navigating the labyrinthine bureaucracy that is the healthcare industry after her husband suffered massive injury and trauma due to an assault. After being denied coverage because the claims were submitted shortly after starting a new job and health plan, Eubanks discovered that her claims patterns were consistent with insurance fraud, and so her coverage was denied (p. 3-5). Eubanks writes that she had both the time and resources to deal with this situation, which could have easily ended up in bankruptcy. However, this is not the case for many poor people who face similar situations (p. 6). Eubanks asserts that lack of knowledge about the algorithms that rule our lives, along with being grouped and treated like data, rather than as individuals, targets all of us, but none more so unfairly than the poor (p. 4). Eubanks' work examines three case studies in which the increasing automation of systems negatively affect the lives of the poor and people of color in this country. In Eubanks' view, these "scientific measurements," designed to track and assess need or predicted risk efficiently, are unchallenged by the wealthy, distancing them from the need to help and from the unethical decisions made about who receives, and who is denied, services.

Eubanks takes three different high-tech tools in three different parts of the country for her study: the coordinated-entry system in Los Angeles, California to help people experiencing homelessness access services; the health-care system in the state of Indiana shifting to an automated model; and a predictive-risk algorithm used by child welfare and services call centers in Allegheny County, Pennsylvania. In each situation, Eubanks speaks with people involved in all aspects of the system: caseworkers, individuals using or being abused by the system, and in the last case, the people responsible for building the predictive risk algorithm. Eubanks asserts that each of these systems, while built for seemingly noble and useful ends, perpetuate the systems of inequality and abuse that the poor and people of color have suffered under in American society since their inception, precisely because they build off the biases inherent in how American society conceives of, and neglects to care for, its poor and marginalized populations. In the beginning of her book, Eubanks traces the development of the ways in which America has dealt with its poor since the construction of poorhouses in the nineteenth century, leading to the creation of a "digital poorhouse" (pp. 21-13) that builds on the same moralistic and punitive views of poverty held by the wealthy and powerful of the 1800s. Specifically, access to services provided by both the coordinated-entry system to combat homelessness in Los Angeles and the predictive risk model of child abuse in Allegheny come with a steep price: a loss of privacy and increased surveillance. This is in line with the 'scientific charity' movements of the nineteenth century and the highly restrictive surveillance of welfare recipients in the 1940s-1970s (pp. 22-24), which treated poverty as a moral failing, and those who were in poverty as people who needed strict surveillance and oversight if they were receiving benefits. In the coordinated-entry system, the questions asked by interviewers for the purposes of determining need and eligibility for housing assistance are incredibly intimate questions, including whether or not someone has used a crisis center service in the last six months or whether or not the person has engaged in activity considered risky, such as exchanging sex for money or running drugs, along with collecting personal protected data, such as social security numbers and birthdates (p. 93). While this information helps social workers and others assess the need for housing and the level of service an individual needs, the individual also signs a waiver that the information will be shared with as many as 168 different organizations, including police departments and governmental agencies (p. 94). For some homeless people interviewed, Eubanks found that the cost of exchanging privacy for surveillance, which occurs as people get placed in houses and are checked up on by different agencies, to be too steep (pp. 101-103). Many people experiencing homelessness also do not benefit from the system, and do not get placed in housing, despite repeatedly completing the Vulnerability Index-Service Prioritization Decision Assistance Tool (pp. 93-95). Thus, these individuals are opening themselves up to increased surveillance and scrutiny without acquiring secure housing.

Eubanks builds on this use of data collection and use of algorithmic scoring to examine the impact of predictive-risk models have on child welfare and family integrity in Pennsylvania. Consistent with O'Neil's assertions and findings, the predictive-risk model unfairly targets the poor and people of color. By gathering data on public services accessed, previous interactions with child welfare organizations, including in the parents' childhoods, and previous phone calls about the family to the call center, the Allegheny Family Screening Tool (AFST) assigns a number to a family to determine risk of abuse or neglect for a child (pp. 141-153). Unfortunately, the algorithm does not take into account that "parenting while poor" looks a lot like poor parenting, and that parents are being unfairly punished for living in poverty and run the risk of having their children taken away from them (pp. 161-162). As with the data collection of people experiencing homelessness in Los Angeles, the poor and people of color in Allegheny County are subject to increased surveillance, scrutiny, policing, and punishment because of their socioeconomic status and the ways in which their data are used in an algorithm that has built in biases against impoverished individuals.

In both books, O'Neil and Eubanks develop arguments about the dangers of pervasive, pernicious data collection and manipulation through biased algorithms that negatively impact the poor and people of color in the United States. Further, both authors argue that these algorithms and systems inherently group people together, rather than treating them as individuals. Both assert that this sort of individual treatment is the privilege of those with wealth and power. However, both also argue that it is just a matter of time before these systems have an impact on the wealthy and powerful as well.

While both of these books have much to recommend them in the ways in which we consider the use and abuse of data collection and mathematical modeling for determining benefits and risks of populations, there are some areas in which both could be improved. Despite being written by a mathematician and data scientist, *Weapons of Math Destruction* does little in the way of illuminating the ways in which algorithms are actually constructed. O'Neil provides useful questions to guide the development of these mathematical models, but does not demonstrate how they may be employed, what better proxies could be used, or how an algorithm fails in being transparent, scalable, or constructive in a real way. *Weapons of Math Destruction* also is so far-ranging in its examples that it tends to lose its focus, especially in the later chapters. By focusing more narrowly on her topic, O'Neil could have developed a stronger argument as to how these weapons are developed and employed, as well as what we could do to combat them more usefully. Conversely, Eubanks does an excellent job of focusing narrowly on the outcomes of these different models, situating them within the historical context of American poverty and the welfare movement. She provides thoughtful responses to the challenges of American poverty in the final section of her book. However, the focus of her book is much more on the history of poverty and inequality. Data collection, the trade off of privacy for benefits, and the algorithms used to deny access to services are not addressed in as much depth. Like O'Neil's work, Eubanks discusses the types of data collected but not the algorithms themselves. Overall, both books complement each other in examining the ways in which data collection, algorithms, technological tools, and predictive modelling are real and present dangers to our world, and ones which demand consistent questioning, challenging, and human oversight.

REFERENCES

Eubanks, V. (2018) *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York, NY, St. Martin's Press.

O'Neil, C. (2016) *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, NY, Broadway Books.