Can AI Solve the Diversity Problem in the Tech Industry? Mitigating Noise and Bias in Employment Decision-Making

Kimberly A. Houser*

22 STAN. TECH. L. REV. 290 (2019)

ABSTRACT

After the first diversity report was issued in 2014 revealing the dearth of women in the tech industry, companies rushed to hire consultants to provide unconscious bias training to their employees. Unfortunately, recent diversity reports show no significant improvement, and, in fact, women lost ground during some of the years. According to a Human Capital Institute survey, nearly 80% of leaders were still using gut feeling and personal opinion to make decisions that affected talent-management practices. By incorporating AI into employment decisions, we can mitigate unconscious bias and variability (noise) in human decision-making. While some scholars have warned that using artificial intelligence (AI) in decision-making creates discriminatory results, they downplay the reason for such occurrences—humans. The main concerns noted relate to the risk of reproducing bias in an algorithmic outcome ("garbage in, garbage out") and the inability to detect bias due to the lack of understanding of the reason for the algorithmic outcome ("black box" problem). In this paper, I argue that responsible AI will abate the problems caused by unconscious biases and noise in human decision-making, and in doing so increase the hiring, promotion, and retention of women in the tech industry. The new solutions to the garbage in, garbage out and black box concerns will be explored. The question is not whether AI should

* Kimberly A. Houser is an assistant professor at Oklahoma State University. The author would like to thank the participants at the 2018 Law and Ethics of Big Data Colloquium in Wellesley, Massachusetts, sponsored by Babson College, Virginia Tech, Center for Business Intelligence and Analytics, Pamplin College of Business, and Indiana University, Department of Legal Studies, for their helpful remarks. The author would also like to additionally thank Angie Raymond, Ramesh Sharda, Griffin Pivateau, and Laurie Swanson Oberhelman for their insightful comments and to Haley Amster, Justin Bryant, Abigail Pace, Katherine Worden, Collin Hong, and Caroline Lebel for their thoughtful and thorough editing, proof-reading, and cite-checking.

be incorporated into decisions impacting employment, but rather why in 2019 are we still relying on faulty human decision-making.

TABLE OF CONTENTS

I. INTRODUCTION

Although 1.4 million computer science jobs in the United States will be available by 2020, only 29% of those positions are expected to be filled, and less than 3% of those jobs will be filled by women.[1] The New Yorker has reported that Silicon Valley loses more than $16 billion annually from the turnover of half of the women who enter the tech field.[2] This mass exodus

1. Swati Mylavarapu, *The Lack of Women in Tech Is More Than a Pipeline Problem*, TECHCRUNCH (May 10, 2016), perma.cc/E2KY-PHW2.

2. Alex Hickey, *Systemic Gender Discrimination Costing Tech Billions*, CIODIVE (Dec. 7, 2017), perma.cc/9BXS-YMEQ; *see also* Jennifer L. Glass et al., *What's So Special about STEM? A Comparison of Women's Retention in STEM and Professional Occupations*, 92 SOC. FORCES 723 (2013) (demonstrating that women in STEM are significantly more likely to leave their field than women in other professions).

signals a significant problem in the industry and represents a substantial obstacle to the U.S. tech industry remaining at the forefront of the world economy. While tech companies in recent years have spoken about reducing the gender gap,[3] little progress has been made.[4]

Traditional methods of hiring include on-campus interviews, online job postings and referrals.[5] Applicants who come from referrals are considered to be better risks.[6] This type of preference can lead to the exclusion of qualified candidates and reinforces the homogenization of an organization. Research has shown that unconscious biases are rife in the tech industry and one of the main factors negatively impacting women in this field.[7] According to a Human Capital Institute survey, nearly 80% of leaders were still using gut feeling and personal opinions to make employment decisions.[8] Not only are human decision-makers unaware of their biases, they are also unaware of the inconsistency of their decisions (known as noise). As Nobel Prize winner Daniel Kahneman points out, human decision-making is fraught with bias and unjustifiable variability.[9] These types of unconscious biases are linked to discriminatory behavior.[10]

---

3. Thomas Ricker, *How Do Tech's Biggest Companies Compare on Diversity?: The Tech Diversity Scorecard,* THE VERGE (Aug. 20, 2015), perma.cc/5XPF-KJJ4. In 2014, Google began releasing its diversity report to the public along with Facebook, Amazon and Apple. In 2015, Microsoft released its first diversity report.

4. *See* David McCandless et al., *Diversity in Tech: Employee Breakdown of Key Technology Companies*, INFORMATION IS BEAUTIFUL (2017), perma.cc/KHJ2-RUUZ; *see also* Visier Insights' Equal Pay Day Brief Finds Younger Female Workers Lost Ground in 2017, VISIER (Apr. 10, 2018), perma.cc/92RW-672N.

5. Rosie Quinn, *Why Traditional Recruitment Methods Are No Longer Enough To Acquire Top Talent*, CIIVSOFT (May 5, 2018), perma.cc/PCB3-ZRC6; Tey Scott, *How Scrapping the Traditional College Recruitment Model Helped LinkedIn Find More Diverse Talent*, LINKEDIN TALENT BLOG (Feb. 6, 2017), perma.cc/E59Q-M5KJ.

6. Stephanie Denning, *The New Hiring Practices at McKinsey and Goldman Sachs*, FORBES (Apr. 27, 2019), perma.cc/HPC9-5LRC. In addition, 60% of recruiters consider culture fit one of the most important factors in hiring, which also results in homogenization. JOBVITE, JOBVITE RECRUITER NATIONAL REPORT 2016: THE ANNUAL SOCIAL RECRUITING SURVEY 1 (2016), perma.cc/476P-P9A9.

7. Luna An et al., *Gender Diversity in Tech: Tackling Unconscious Bias*, MEDIUM (Aug. 14, 2017), perma.cc/2795-UEHC ("Unconscious biases are deep-seated ideas and impressions about certain groups that we carry with us and cause us to draw unfounded conclusions about people in those groups.").

8. HUM. CAP. INST., INSIGHTFUL HR: INTEGRATING QUALITY DATA FOR BETTER TALENT DECISIONS (2015), https://perma.cc/EYQ5-W3DV.

9. *See infra* Part II.

10. *See* Anthony G. Greenwald & Linda Hamilton Krieger, *Implicit Bias: Scientific Foundations*, 94 CALIF. L. REV. 945, 961 (2006) ("[E]vidence that implicit attitudes produce discriminatory behavior is already substantial.").

The responsible use of artificial intelligence,[11] however, can mitigate unconscious bias by reducing the impact of human decision-makers on the process, and create better employment decisions which are based on skills, traits and behaviors rather than factors (such as sex, race, or pedigree) that do not correlate with merit or success. A Harris Poll revealed that 75% of employers reported making a bad hire in the last year.[12] The responsible use of artificial intelligence in employment decision-making not only increases the diversity of candidates and employees, but actually results in more successful employment outcomes.[13] AI is the ability of a machine to perform functions that humans engage in[14] through the use of a programmed series of steps known as algorithms. Although there are many domains of AI, as used herein it refers to algorithms processing data to produce an outcome.[15]

AI can be used to anonymize resumes as well as interviewees, identify the skills, traits, and behaviors needed to succeed in a certain job, match applicants with open positions, and predict when an employee is likely to leave, thereby giving the organization time to remediate the situation and improve retention.[16] These measures can attenuate the inherent bias and

---

11. Although others use the terms "people analytics," "talent management," "machine learning," or "predictive analytics" interchangeably, those terms refer to very specific processes. "AI" as used herein is intended to reflect the broad category of the use of computers to perform tasks ranging from removing names from resumes to data mining performance reviews.

12. Ladan Nikravan Hayes, *Nearly Three in Four Employers Affected by a Bad Hire, According to a Recent CareerBuilder Survey*, CAREERBUILDER (Dec. 7, 2017), perma.cc/5BJC-2YUF.

13. Charles A. Sullivan, *Employing AI,* 63 VILL. L REV. 395, (2018). While this paper focuses on reducing the gender disparity in the tech industry, the author acknowledges that different issues are encountered by underrepresented minority groups, LGBT+ individuals, and those with disabilities. The author also acknowledges that black women and others who fall into more than one category face more complex issues around discrimination than do white women. While this paper is meant to improve conditions for women overall, further research does need to study the impact of the recommended methods discussed in this paper on other groups and combinations of groups, but initial reports confirm that the reduction of unconscious bias and noise in employment decisions also improves hiring rates for URMs. Guòrun I. Jákupsstova, *AI Is Better Than You at Hiring Diversely*, NEXT WEB (May 31, 2018), perma.cc/NQN8-TMFC.

14. Lauri Donahue, Commentary, *A Primer on Using Artificial Intelligence in the Legal Profession*, HARV. J. L. & TECH. DIG., (Jan. 3, 2018) ("'Artificial Intelligence' is the term used to describe how computers can perform tasks normally viewed as requiring human intelligence, such as recognizing speech and objects, making decisions based on data, and translating languages.").

15. As an example, Spotify reviews both your previous music selections and the selections of other users who have chosen similar music in the past. These music selections are the data, and the recommended songs are the outcome.

16. Rohit Punnoose & Pankaj Ajit, *Prediction of Employee Turnover in Organizations*

noise present in human decision-making, which are a pervasive problem in the tech industry.[17] Additionally, AI can be used to moderate the problem of human bias baked into the algorithmic process ("algorithmic bias") by detecting and correcting problems in biased data sets.[18] These fixes result in better accuracy, consistency and fairness in employment decisions.[19] Most importantly, the use of AI in employment has been shown to increase the hiring, promotion and retention of women.[20] As one example, Pymetrics, which recently received the Technology Pioneer Award from the World Economic Forum, relies on gamified solutions[21] which have resulted in a significant increase in the hiring of women by their clients.[22] While the term AI is used throughout, this paper does not suggest that human decision-making be completely replaced with machines. It proposes that algorithmic-based decisions are the key to increasing diversity in the tech industry and

---

Using Machine Learning Algorithms: A Case for Extreme Gradient Boosting, 5 INT'L J. ADVANCED RES. ARTIFICIAL INTELLIGENCE 22, 26 (2016). For instructions on how to create your own prediction model, see Marian Dragt, *Human Resources Analytics—Predict Employee Leave*, MD2C (Apr. 11, 2018), perma.cc/F2E9-QSRN.

17. *See infra* Part VI.

18. *See infra* Part VII.

19. Infor Talent Science, a company that employs AI to collect behavioral information using a survey, reported a 26% increase in URMs in a sample of 50,000 hires. Bouree Lam, *Recruiters Are Using 'Algorithmic Hiring' to Solve One of the Workplaces' Biggest Problems*, BUSINESS INSIDER (June 28, 2015), perma.cc/YF36-UYPL. In an analysis of seventeen studies comparing human and machine predictions of performance, the authors concluded that machines were 25% better at evaluating candidates than human, even when humans had access to more information. Nathan Kuncel et al., *Mechanical Versus Clinical Data Combination in Selection and Admissions Decisions: A Meta-Analysis*, 98 J. APPLIED PSYCHOL. 1060, 1060-72 (2013).

20. *See infra* Part VI.

21. Gamification is the incorporation of game elements into non-game contexts. Miriam A. Cherry, *The Gamification of Work*, 40 HOFSTRA L. REV. 851, 852 (2012). Gamification in human resources (HR) includes coding challenges, virtual hotel and restaurant simulations, earning points and badges for completing activities, and a virtual escape room for assessing collaboration skills. Chiradeep BasuMallick, *Gamification in Recruitment: All You Need to Know*, HR TECHNOLOGIST (Nov. 30, 2018), perma.cc/K2P3-XA5Y; Sara Coene, *9 Examples of Gamification in HR*, HR TREND INST. (Feb. 25, 2019), perma.cc/5PWB-Z5LC.

22. *Pymetrics Awarded as Technology Pioneer by World Economic Forum*, BUSINESS WIRE (June 21, 2018), perma.cc/EY7J-38ZT. Pymetrics was founded by Harvard and MIT-trained Ph.Ds, and uses neuroscience to create games which applicants play in order to be matched with positions. Companies utilizing their services have reported that the diversity of candidates has increased by 20% and retention by 65%. PYMETRICS, SUBMISSION TO THE AUSTRALIAN HUMAN RIGHTS COMMISSION: HUMAN RIGHTS AND TECHNOLOGY 2 (Oct. 2, 2018), perma.cc/TMY2-GKL6.

explores solutions for the potential risks noted by various scholars in the adoption of such programs.

This paper makes three contributions to the intersection of the law, social science and technology. First, it suggests a way to increase gender diversity in the tech industry, which is not only the socially responsible thing to do, but is also the smart thing to do. Second, it provides a solution to the problem of unconscious bias and noise in human-decision making without changing or advocating new laws. Third, it explains how AI can improve employment decision-making and root out and correct discriminatory outcomes in AI applications. Part I of this paper describes the environment in the tech industry that women experience and counters the common reasons expressed for the lack of women in this industry. Part II explores unconscious bias in gender discrimination law and why it is an insufficient remedy for the disparities noted in this paper. Part III makes the business case for women in tech, explaining why it is more than an equity issue. Part IV examines the failure of current methods to increase diversity. Part V explains the research on unconscious bias and noise inherent in human decision-making. Part VI describes why AI is superior to human decision-making and how it can be implemented to reduce the impact of unconscious bias and noise. Part VII explains how AI can also be used to discover and correct the risks of algorithmic bias itself. Part VIII addresses the legal concerns of using AI in employment decisions, followed by the conclusion.

II. WOMEN IN THE TECH INDUSTRY

The walk-out of 20,000 Google employees to protest Google's sexist culture in November 2018 demonstrates the frustration with the tech industry's failure to fulfill its promises.[23] Although tech companies began publishing diversity reports in 2014, little has changed, and sexism and discrimination continue to occur.[24] A survey of women working in Silicon

23. Maria Fernandez, *Go Deeper: Google's Restlessness for Better Company Culture*, AXIOS (Nov. 3, 2018), perma.cc/LG4K-NM5D.

24. There has been no meaningful improvement since 2015 when the ELEPHANT IN THE VALLEY report as referenced *infra* in note 25 came out. Quentin Fottrell, *Woman Leaders Are Still Getting Screwed by Tech Companies,* N.Y. POST (Feb. 8, 2018), perma.cc/HTK8-8GGM ("Female managers are not only under-represented in technology companies, they're also paid significantly less than men. In the Bay Area, they're paid $172,585 per year, 10 percent less than men. In Seattle, female managers are paid an average of $158,858 per year, also 10 percent less than men."); *id.*

Valley, *The Elephant in the Valley*, revealed the lack of diversity and extreme sexism faced by women at these tech firms, with 88% reporting evidence of unconscious biases.[25] The following is a description of some of the problems women encounter in the tech industry.[26]

A. Issues Faced by Women

One of the issues affecting gender bias in the field is the lack of female role models and leaders in the industry.[27] Women make up barely 11% of tech industry executives[28] and only 9% of senior IT leadership roles such as CIOs.[29] Amazon and Microsoft , headquartered in Washington, reveal a stunning lack of diversity, especially at senior levels.[30] Of Amazon's eighteen most powerful executives, seventeen are men.[31] At the recent Consumer Electronics Show (CES), all of the keynote speakers were male.[32] An especially discouraging fact is that a recent LivePerson survey of 1,000

25. *Elephant in the Valley,* ELEPHANT IN THE VALLEY, perma.cc/97EH-PDB8 (archived Apr. 19, 2019). Respondents had at least ten years with tech companies and were very familiar with the myriad gender equity issues. The creators of the survey, from Stanford and Klein Perkins, wanted to put numbers to the experiences of women in the tech field.

26. Kristin Houser, *The Tech Industry's Gender Problem Isn't Just Hurting Women*, FUTURE SOC'Y (Jan. 31, 2018), perma.cc/94W6-AAKY.

27. This lack was noted by 90% of the respondents to a survey by booking.com. Nick Ismail, *Gender Bias in the Tech Industry Is All Encompassing,* INFO. AGE (Nov. 8, 2017), perma.cc/GE8D-6WEB.

28. Sheelah Kolhatkar, *The Tech Industry's Gender-Discrimination Problem*, NEW YORKER (Nov. 13, 2017), perma.cc/N5P9-AU78. Women leaders also earn about 10% less than their male counterparts. Fottrell, *supra* note 24.

29. NAVIGATING UNCERTAINTY: THE HARVEY NASH / KPMG CIO SURVEY 2017 20 (2017), perma.cc/K3FK-JG7X; *see also* Luke Graham, *Women Take Up Just 9 Percent of Senior IT Leadership Roles, Survey Finds*, CNBC (May 22, 2017), perma.cc/6P3L-7UU3 (finding virtually no increase in women in IT leadership roles from the previous year in survey of 4,498 CIOs and tech leaders).

30. As Bloomberg Businessweek noted, "The search for a second home gives Amazon something else: an unprecedented opportunity to deal with a problem besetting all of big tech—a stunning lack of diversity. And Amazon is one of the bigger sinners. Men make up 73 percent of its professional employees and 78 percent of senior executives and managers, according to data the company reports to the government. Of the 10 people who report directly to Chief Executive Officer Jeff Bezos, all are white, and only one—Beth Galetti, the head of human resources—is a woman. The board of directors is also resisting shareholder pressure to improve gender balance." Emily Chang, Jeff Green & Janet Paskin, *Amazon Has a Rare Chance to Get More Diverse Fast*, BLOOMBERG BUSINESSWEEK (May 10, 2018), perma.cc/HZ9Q-XJ9P.

31. Jason Del Ray, *It's 2017 and Amazon Only Has One Woman Among Its 18 Most Powerful Executives*, RECODE (Oct. 21, 2017), perma.cc/29MQ-DGC6.

32. Andrew Mosteller, *Female Tech Leaders Take on Equality Issues,* LENDIO (Mar. 3, 2018), perma.cc/S7M5-38AU.

people showed that while half of the respondents could name a famous male tech leader, only 4% could name a female tech leader and one-quarter of them named Siri and Alexa—who are virtual assistants, not actual people.[33] The lack of women in leadership roles stems from the inability for most women to move up in these companies. This glass ceiling, as well as widespread disrespect, harassment and exclusion, results in about half of the women[34] entering the field to leave it (compared with 17% of men).[35] *The Elephant in the Valley* survey indicated that 87% of the women reported receiving demeaning comments from male colleagues, 47% said they had been asked to do lower-level tasks that male colleagues were not asked to do, and 66% said they had been excluded from important social or networking events.[36] Comments on the survey indicated that women were disrespected in numerous ways, such as being asked to take notes at meetings or order food, and being ignored in favor of male subordinates during meetings.[37]

While this figure does not surprise most women, men seem flabbergasted[38] that 90% of the women surveyed reported witnessing sexist

---

33. Monica Torres, *Survey Reveals People Think Siri and Alexa Are Female Tech Leaders*, THE LADDERS (Mar. 23, 2018), perma.cc/4TDM-BB5C.

34. Hickey, *supra* note 2; *see also* Glass et al., *supra* note 2.

35. Houser, *supra* note 2625.

36. *Id*. For some of the more outrageous experiences, see EMILY CHANG, BROTOPIA: BREAKING UP THE BOY'S CLUB OF SILICON VALLEY (2018).

37. *Elephant in the Valley, supra* note 25. In another interesting experiment, two co-workers, one male and one female, exchanged email signatures to see if they were being treated differently because of their sex. The male using the female email signature said the treatment was like night and day. As "Nicole," he found clients to be rude, dismissive, and condescending. Nicole indicated that she had the most productive week of her life *because she did not have to convince clients to respect her*. Nicole Hallberg, *Working While Female*, MEDIUM (Mar. 9, 2017), perma.cc/2C3M-NPYW; *see also* @lindworm, *Working as a Woman Can #Suck*, TWITTER (Mar. 9, 2017), perma.cc/27TK-4WV8 (displaying actual string of tweets). This sexism extends to women in positions of leadership as well. Jenny Campbell, who sold her business for £50 million, was repeatedly misidentified as the wife of her male subordinate when he was standing next to her. Jamie Johnson, *I Sold My Business for £50 Million, But It's Still Assumed I Am the Wife of the Man Standing Next to Me in Meetings*, TELEGRAPH (Mar. 23, 2018), perma.cc/A4ZF-M6XP. In order to avoid this exact problem, Penelope Gazin and Kate Dwyer, the owners of a successful online marketplace, invented an imaginary co-founder named Keith. While some of the obstacles they faced were overt, such as the web developer who deleted their site after one of the owners indicated a lack of interest in dating him, most of their experiences were the result of subtle sexism. While it took days for the women to receive input from collaborators, "Keith" received responses immediately. John Paul Titlow, *These Women Entrepreneurs Created A Fake Male Cofounder to Dodge Startup Sexism*, FAST COMPANY (Aug. 29, 2017), perma.cc/HE9Z-T6V3.

38. Alice Berg, *Gender Discrimination Problem: Silicon Valley VS Women*,

behavior at their company and industry conferences.[39] Sixty percent had been harassed themselves and 33% feared for their safety due to work-related circumstances.[40] However, most of these incidents are not reported due to fear of retaliation.[41] In one well-publicized case, AJ Vandermeyden, who sued Tesla for discrimination, retaliation, and other workplace violations, was fired after filing suit.[42] Therese Lawless, Vandermeyden's attorney, confirmed that firing whistleblowers is a common form of retaliation against women who complain of discrimination. "That's the message the company sends if you speak up. That's why people are fearful," Lawless said.[43]

In addition to sexism, there is also the issue of stereotyping. Characteristics that tend to be valued in men, often resulting in the advancement in their careers, have the opposite effect when exhibited by women. Eighty-four percent of those surveyed reported that they had been told they were "too aggressive," and 66% reported being excluded from key social/networking opportunities because of their gender.[44] While peers prefer successful men to unsuccessful ones, successful women are

---

MUNPLANET (Jan. 10, 2018), perma.cc/Y98L-4WPS.

39. *Elephant in the Valley, supra* note 25.

40. Houser, *supra* note 26.

41. Kolhatkar, *supra* note 28. Most companies, upon hearing a report from a woman, will ask her to "prove" the event occurred if the man denies it happened. CHAI R. FELDBLUM & VICTORIA A. LIPNIC, U.S. EQUAL EMP'T OPPORTUNITY COMM'N, REPORT OF THE CO-CHAIRS OF THE EEOC SELECT TASK FORCE ON THE STUDY OF HARASSMENT IN THE WORKPLACE (2016). Because there is no audio or video proof, nor witnesses, no action is taken and the woman eventually leaves the company because of being shunned, demoted, or transferred after making the report, or chooses to leave because of the firm's disrespect shown by choosing to believe the perpetrator rather than the woman harassed. According to Melinda Gates, the co-chair of the Bill & Melinda Gates Foundation, "Men who demean, degrade or disrespect women have been able to operate with such impunity—not just in Hollywood, but in tech, venture capital, and other spaces where their influence and investment can make or break a career[.] The asymmetry of power is ripe for abuse." Kolhatkar, *supra* note 28.

42. Kolhatkar, *supra* note 28. Ellen Pao, whose discrimination case was widely publicized, was one of the women who found herself the target of harassment by a male colleague she briefly dated. After filing suit, Pao was terminated and her complaint was amended to include a count of retaliation. Although Pao lost her case in 2015, everyone got a glimpse of the vitriolic response of the venture capital firm she worked for, Kleiner Perkins, which only confirmed in the public's eye the likelihood that Pao was accurately describing what she had encountered. Eric Johnson, *Why Did Ellen Pao Lose Her Gender Discrimination Lawsuit? 'People Were Not Ready'*, VOX (Oct. 2, 2017), https://perma.cc/YS4T-QAM4.

43. *Id.*

44. *Elephant in the Valley*, *supra* note 25.

penalized.[45] In employee reviews, women are referred to as "abrasive," "aggressive," "bossy," and "strident."[46] Women who attempt to negotiate salaries are viewed as being "difficult to work with" even when using the same language as men.[47] This is known as the likeability penalty.[48]

One of the most interesting findings I came across was that many men do not acknowledge that a gender diversity problem even exists.[49] If men, who hold 80% of the leadership positions in tech companies, do not even believe that the low levels of women in their company is a problem, it is unlikely it will get resolved. According to a 2017 study, 50% of men reported that it is sufficient when just one in ten senior leaders in their company is a woman.[50] What is especially telling is a study showing when 17% of the people in a room are women, men report that they think the room is 50-50 men and women.[51] When 33% of the people in the room are women, men believe they are outnumbered.[52] This skewed view of reality may explain why men in tech express a belief that there are enough women in leadership positions in their companies already[53] despite women only comprising 11% of executives in tech as indicated above.

---

45. Successful women are viewed as being less likeable. While terms like "confident" and "strong" are used to describe successful men, women are called "bossy" and "aggressive." *7 Tips for Men Who Want to Support Equality*, LEANIN.ORG, perma.cc/7KMG-5M67. In tech, these types of terms appear in 85% of female high performers' evaluations compared to only 2% of the men's evaluations. Kieran Snyder, *The Abrasiveness Trap: High-Achieving Men and Women Are Described Differently in Reviews*, FORTUNE (Aug. 26, 2014) [hereinafter Snyder, *Abrasiveness*], perma.cc/VB7V-LX7W.

46. Snyder, *Abrasiveness*, *supra* note 45.

47. Hannah Riley Bowles et al., *It Depends Who is Asking and Who You Ask: Social Incentives for Sex Differences in the Propensity to Initiate Negotiation—Sometimes It Does Hurt to Ask*, 103 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 84 (2007).

48. Marianne Cooper, *For Women Leaders, Likability and Success Hardly Go Hand-in-Hand*, HARV. BUS. REV. (Apr. 30, 2013), perma.cc/MBD9-HHPK.

49. Lauren Williams, *Facebook's Gender Bias Goes So Deep It's in the Code,* THINK PROGRESS (May 2, 2017), perma.cc/9M9Z-GZJH.

50. ALEXIS KRIVKOVICH ET AL., MCKINSEY & CO., WOMEN IN THE WORKPLACE 2017 (Oct. 2017), perma.cc/J2TG-UD2R.

51. Linnea Dunne, *So You Think You Were Hired on Merit? Gender Quotas and the Perception Gap*, LINNEA DUNNE BLOG (Aug. 21, 2017), perma.cc/WLW3-LUKS.

52. *Id*.

53. Emma Hinchliffe, *58% of Men in Tech Say There Are Enough Women in Leadership Roles, But Women Don't Agree,* MASHABLE (Sept. 20, 2017), https://perma.cc/3BPG-2MWN. When a study was shown to male faculty members demonstrating the unjustified preference for male lab managers (where simply changing names from female to male on a resume made the lab manager more likely to be hired), they still assessed bias against women as being low. Alison Coil, *Why Men Don't Believe*

B. Reasons Attributed to the Lack of Women in the Tech Industry

Although women hold 57% of professional occupations in the U.S. workforce, they occupy only 26% of professional computing jobs.[54] Reasons alleged to explain this gender gap in the tech field include: lack of pipeline, lack of interest, and the confident assertion of meritocracy.[55] Although women with STEM degrees are available,[56] companies hire men with science and engineering degrees at twice the rate of women.[57] One study shows that when a tech job was available, 53% of the time, companies interviewed no female candidates at all.[58] Women also receive lower salaries for the same job at the same company 60% of the time.[59]

Although women are interested in these jobs, many are alienated during the recruiting process itself.[60] Researchers from Stanford revealed that

---

The Data On Gender Bias In Science, WIRED (Aug. 25, 2017), https://perma.cc/RNQ6-PWSX. To review the actual study, see Corinne A. Moss-Racusin et al., *Science Faculty's Subtle Gender Biases Favor Male Students*, 109 PROCS. NAT'L ACAD. SCI. 16,474 (2012).

54. NAT'L CENTER FOR WOMEN & INFO. TECH., BY THE NUMBERS (2019), https://perma.cc/J4Y5-7B7R.

55. Banu Ozkazanc-Pan, *Women in Tech Suffer Because of the American Myth of Meritocracy*, THE LADDERS (May 4, 2018), https://perma.cc/XK4F-5PZC; Reshma Saujani & Ayah Bdeir, *Opinion: You Can't Solve the Gender Gap in Tech If You Don't Understand Why It Exists*, BUZZFEED NEWS (Mar. 6, 2019), https://perma.cc/84BT-VD8C.

56. At Cornell, 55% of the incoming freshmen in the fall of 2018 in the engineering school who indicated an interest in computer science were women. The year before, 38% of declared computer science majors were women. *55 Percent of Incoming Eng Students Interested in Computer Science Are Women*, CORNELL CIS, https://perma.cc/BA7Z-DULH (archived Apr. 13, 2019) [hereinafter *55 Percent*]. Dartmouth College graduates more women in computer science than men, at 54%. Thea Oliver, *An In-Depth Look at the Gender Gap in the Tech Industry,* TECHNICALLY COMPATIBLE (May 12, 2017), https://perma.cc/E7G4-XR6N. Harvey Mudd graduated 56% women in computer science in 2018. *Harvey Mudd Graduates Highest-Ever Percentage of Women Physics and Computer Science Majors*, HARVEY MUDD COLL. NEWS (May 15, 2018), https://perma.cc/Y5TW-QX9Z; *see also* Kristen V. Brown, *TechShift: More Women in Computer Science Classes*, SFGATE (Feb. 18, 2014), https://perma.cc/P2EN-24KJ ("Berkeley, Stanford and a handful of other universities have experienced a marked uptick in the numbers of female computer science students."). In addition, outside of the United States, women in STEM are the rule, not the exception. In Iran, 70% of STEM graduates are female, with over 60% in Oman, Saudi Arabia, and the UAE. Annalisa Merelli, *The West Is Way Behind Iran and Saudi Arabia When It Comes to Women in Science*, QUARTZ (Mar. 8, 2018), https://perma.cc/QD5Z-WY8N.

57. Liana Christin Landivar, *Disparities in STEM Employment by Sex, Race, and Hispanic Origin*, AM. COMM. SURV. REP. (Sept. 2013), https://perma.cc/TB4M-RQ62.

58. NANCY M. CARTER & CHRISTINE SILVA, CATALYST, THE MYTH OF THE IDEAL WORKER: DOES DOING ALL THE RIGHT THINGS REALLY GET WOMEN AHEAD? (2011).

59. *The State of Wage Inequality in the Workplace*, HIRED (2019), https://perma.cc/HS75-8E6C.

60. Alison T. Wynn & Shelley J. Correll, *Puncturing the Pipeline: Do Technology*

"through gender-imbalanced presenter roles, geek culture references, overt use of gender stereotypes and other gendered speech and actions, representatives may puncture the pipeline, deflating the interest of women at the point of recruitment into technology careers."[61] In addition, the wording of job ads can dissuade women from applying. Researchers discovered that women were less likely to apply to jobs described with masculine words such as "competitive" and "dominant."[62] Gendered job ads resulted in women believing 1) the company employed more men than women, 2) they did not belong, and 3) the job would not be appealing.[63]

Even when women are hired, the tech companies are unable to retain them.[64] Women leave the tech field at a rate 45% higher than men do.[65] Of the very low levels of women who are hired, half leave because of the work environment.[66] A longitudinal study regarding retention reveals the real problem, and it is not the pipeline.[67] In a survey of 716 women in tech who left the field, all of them said they enjoyed the work, but not the workplace environment.[68] Women increasingly are speaking out about what they see as a hostile culture due to buddy networks.[69] These informal networks, which benefit men, tend to exclude otherwise qualified women. [70]

Companies Alienate Women in Recruiting Sessions? 48 SOC. STUD. SCI. 149 (2018), https://perma.cc/JK3B-B4JD.

61. For an explanation of how recruiters from tech companies alienate female applicants, see Wynn et al., *supra* note 60 (84% of presenters were male, women representatives were used to hand out swag, presenters promoted the fraternity-like environment, and in one case, a presenter made multiple references to pornography and prostitution).

62. Danielle Gaucher et al., *Evidence That Gendered Wording in Job Advertisements Exists and Sustains Gender Inequality*, 101.1 J. PERSONALITY & SOC. PSYCHOL. 109 (2011).

63. *Id.*

64. Kim Williams, *Women in Tech: How to Attract and Retain Top Talent*, INDEED BLOG (Nov. 6, 2018), perma.cc/7AY8-D8ZU.

65. Williams, *supra* note 64.

66. Kieran Snyder, *Why Women Leave Tech: It's the Culture, Not Because 'Math Is Hard'*, FORTUNE (Oct. 2, 2014), perma.cc/F5XF-KZXC [hereinafter Snyder, *Women Leave Tech*].

67. Glass, *supra* note 2.

68. Snyder, *Women Leave Tech*, *supra* note 66. *See also* Kolhatkar, *supra* note 28, for a discussion of what women experience in the workplace environment.

69. Chang, *supra* note 36.

70. Laura Colby, *Women and Tech,* BLOOMBERG: QUICKTAKE (Aug. 8, 2017), https://perma.cc/D974-DTSM.

"Undermining behavior from managers" was also reported as a major factor,[71] as well as the inability to move up in the company.[72]

While women initially are attracted to the tech industry, structural barriers to advancement and workplace issues force them out.[73] A 2018 report from McKinsey found that women made up 48% of entry-level roles in tech with only 23% of those advancing to senior management roles.[74] Men are not only promoted more frequently, even when a woman is given credit for her contributions to the growth of the company, she does not receive the promotion.[75] Men also tend to receive recognition more often than women.[76] While men are hired based on their potential, women are hired based on their proven experience.[77] This is known as the "prove it again syndrome."[78]

---

71. Liza Mundy, *Why Is Silicon Valley So Awful to Women?*, THE ATLANTIC, Apr. 2017.

72. Tekla S. Perry, *Women Leave Tech Jobs Because They Can't Climb the Ladder*, IEEE SPECTRUM (Nov. 6, 2018), https://perma.cc/LFS3-8M4L.

73. *See* Allison Schnidman, *Why Women Are Leaving Their Jobs (Your First Guess Is Wrong)*, LINKEDIN TALENT BLOG (Nov. 5, 2015), https://perma.cc/3XW9-NE2F (revealing that, according to a survey of 4,000 women in the tech industry, the top three reason women left their tech jobs were concern for the lack of opportunities for advancement, dissatisfaction with leadership and the work environment). *See also* CAROLINE SIMARD ET AL., ANITA BORG INST. FOR WOMEN & TECH., CLIMBING THE TECHNICAL LADDER: OBSTACLES AND SOLUTIONS FOR MID-LEVEL WOMEN IN TECHNOLOGY (2008) (providing results from a survey of female mid-level managers at Silicon Valley high-tech firms regarding barriers to advancement).

74. ALEXIS KRIVKOVICH ET AL., MCKINSEY & CO., WOMEN IN THE WORKPLACE 2018 (2018).

75. Berg, *supra* note 38.

76. According to a survey of 1,000 professionals, 50% of men indicated that they received recognition at work at least a few times per month compared with only 43% of women. Bryson Kearly, *Is Gender Equality in the Workplace Still an Issue? Studies Say Yes!*, HR INSIGHTS (Apr. 20, 2016), https://perma.cc/K3EA-XLRW.

77. Joan C. William et. al., *Tools for Change: Boosting the Retention of Women in the STEM Pipeline*, 6 J. RES. GENDER STUD. 11 (2016).

78. Eileen Pollack covered this in detail in a New York Times article describing the "Prove-It-Again! bias" which requires women to provide more evidence of competence than men in order to be seen as equally competent. Pollack points out a study conducted at Yale proving that a young male scientist will be viewed more favorably than a woman with the same qualifications. When professors at six major research institutions were presented with identical summaries of the accomplishments of two imaginary applicants, they were significantly more willing to offer the man a job. When they did choose a woman, they paid her $4,000 less than what they paid the men hired. In keeping with Pollack's findings, a peer-reviewed study of top U.S. graduate programs in the sciences funded by the National Academy of Sciences demonstrated that both female and male professors rated male applicants for a lab manager position as "significantly more competent and hirable than (identical) female applicants." Eileen Pollack, *Why Are There Still So Few Women in Science?,* N.Y. TIMES MAG. (Oct. 3, 2013), https://perma.cc/6FRA-74VB.

While few come right out and say that women lack ability, the explanation most often used to disguise the prejudice against women is that the tech industry is a "meritocracy" implying that the men are simply more qualified.[79] This argument does not stand up to scrutiny. A study of 1.6 million students showed that the top 10% of STEM classes contain an equal number of men and women.[80] In terms of actual skills related to the position, women may in fact be better at coding than men. A study reflecting data from the largest open source community (GitHub) with 12 million collaborators across 31 million software repositories showed that while women's codes were rated more harshly than men's when gender was visible, when gender was hidden, the women's codes were found to be rated consistently better.[81] This study refutes the argument that women are somehow less qualified or capable than men and demonstrates how the meritocracy argument is largely a reflection of gender bias rather than actual verifiable fact.

Because decision-makers are unaware of their own biases, they explain their decision as being "on the merits" without factoring in their preference for a candidate based on factors that have nothing to do with job skills.[82] In addition, decision-makers may focus their attention on information that confirms their existing belief system and disregard potentially relevant information that would tend to contradict it.[83] "Most interviews are a waste

---

79. In a now-deleted article on Forbes, tech writer Brian S. Hall argued that Silicon Valley was in fact a meritocracy. He stated, "If you aren't able to make it here, it's almost certainly not because of any bias," and argued anyone claiming bias should blame their own "refusal to put in the hard work." Dexter Thomas, *Forbes Deleted a White Tech Writer's Article That Called Silicon Valley a 'Meritocracy'*, L.A. TIMES (Oct. 8, 2015), https://perma.cc/E74T-3HNQ. *See also Is Tech a Meritocracy?*, https://perma.cc/F23A-RQ3A (archived May 8, 2019) (providing numerous criticisms of the allegation that tech is a meritocracy).

80. R. E. O'Dea et al., *Gender Differences in Individual Variation in Academic Grades Fail to Fit Expected Patterns for STEM*, 9 NATURE COMM. 3777 (2018).

81. Josh Terrell et al., *Gender Differences and Bias in Open Source: Pull Request Acceptance of Women Versus Men*, 3 PEER J. COMP. SCI. e111 (2017). *See also* Julia Carriew, *Women Considered Better Coders—But Only If They Hide Their Gender*, THE GUARDIAN (Feb. 12, 2016), https://perma.cc/KEN7-VZNT (describing GitHub research study).

82. For example, someone who graduated from Harvard may exhibit a preference for a candidate who also attended Harvard. Julia Mendez, *The Impact of Biases and How to Prevent Their Interference in the Workplace*, INSIGHT INTO DIVERSITY (Apr. 27, 2017), https://perma.cc/JAD3-9R4J.

83. Kathleen Nalty, *Strategies for Confronting Unconscious Bias*, 45 COLO. LAW. 45 (2016). "Another type of unconscious cognitive bias—attribution bias—causes people to make more favorable assessments of behaviors and circumstances for those in their 'in groups' (by giving second chances and the benefit of the doubt) and to judge people in

of time because 99.4% of the time is spent trying to confirm whatever impression the interviewer formed in the first 10 seconds," according to Laszlo Bock, the author of Work Rules! [84] Similarly, companies who tout meritocracy actually demonstrate more bias against women than those who do not.[85] The solution discussed in Part VII below is the use of AI to moderate this type of human error from the hiring process.

Unconscious biases and noise not only influence employment decisions, but also how the workplace culture evolves.[86] The effect of unconscious biases is well correlated with discriminatory employment decisions.[87] Although studies bear this out, the courts have a difficult time reconciling these subtler forms of discrimination with the law.

III. GENDER DISCRIMINATION LAW

Title VII of the Civil Rights Act prohibits the discrimination by an employer against any individual with respect to compensation, terms, conditions, or privileges of employment, because of such individual's race, color, religion, sex, or national origin.[88] Although overt forms of discrimination have been reduced due to antidiscrimination law and changes in societal norms,[89] cases involving more covert forms of

---

their 'out groups' by less favorable group stereotypes." *Id*. "The adverse effects of many of these cognitive biases can be compounded by affinity bias, which is the tendency to gravitate toward and develop relationships with people who are more like ourselves and share similar interests and backgrounds. This leads people to invest more energy and resources in those who are in their affinity group while unintentionally leaving others out." *Id*.

84. Jennifer Alsever *How AI Is Changing Your Job Hunt*, FORTUNE (May 19, 2017), https://perma.cc/24FG-T8G2.

85. Emilio J. Castilla & Stephen Benard, *The Paradox of Meritocracy in Organizations*, 55 ADMIN. SCI. Q. 543 (2010). A study at Cornell revealed that when the participants were asked to award bonuses to men and women with similar profiles, telling them that their company valued merit-based decisions actually increased the likelihood of higher bonuses to the men. *Id*.

86. An, *supra* note 7.

87. *See* Melissa Hart, *Subjective Decisionmaking and Unconscious Discrimination*, 56 ALA. L. REV. 741, 744-745 (2004) ("There is little doubt that unconscious discrimination plays a significant role in decisions about hiring, promoting, firing, and other benefits and tribulations of the workplace"); Audrey Lee, *Unconscious Bias Theory in Employment Discrimination Litigation*, 40 HARV. C.R.-C.L. L. REV. 481, 483-87 (2005) ("Courts have recognized the existence of unconscious discrimination since the earliest Title VII decisions"); Greenwald & Krieger, *supra* note 10.

88. 42 U.S.C. § 2000e-2(a) (2012).

89. Lee, *supra* note 87at 488.

discrimination have been less successful. As such, current application of law does not provide an adequate remedy for those harmed by non-obvious non-intentional discrimination.[90] Class action suits for these types of matters are seldom certified,[91] and most tech companies have arbitration or confidentiality requirements that prevent women from getting their day in court.[92]

Although social science has greatly advanced our understanding of how unconscious biases influence the workplace and can lead to discrimination,[93] courts have been inconsistent in their treatment of this evidence.[94] Because courts have required proof of "intent" in disparate treatment cases,[95] most actions relying on unconscious bias as the cause of an adverse action, assert a disparate impact claim.[96] However, cases relying

90. Intentional discrimination theory would not cover harms due to subjective human decision-making because "intent" requires some outward showing of prejudice resulting in a protected group being subjected to an adverse employment action due to their membership in the group. Stephanie Bornstein, *Reckless Discrimination*, 105 CALIF. L. REV. 1055 (2017).

91. See discussion in Subpart II.A.

92. IMRE S. SZALAI, THE EMP. RIGHTS ADVOC. INST. FOR LAW & POL'Y, THE WIDESPREAD USE OF WORKPLACE ARBITRATION AMONG AMERICA'S TOP 100 COMPANIES 6 (2018).

93. Linda Hamilton Krieger, *The Content of Our Categories: A Cognitive Bias Approach to Discrimination and Equal Employment Opportunity*, 47 STAN. L. REV. 1161 (1995) (suggesting that biased discriminatory decisions may result not from intentional action, but rather "unintentional categorization-related judgment errors characterizing normal human cognitive function.").

94. *See* Anthony Kakoyannis, *Assessing the Viability of Implicit Bias Evidence in Discrimination Cases: An Analysis of the Most Significant Federal Cases*, 69 FLA. L. REV. 1181 (2018) (discussing the differences in the treatment of implicit bias evidence by the courts); Melissa Hart and Paul M. Secunda, *A Matter of Context: Social Framework Evidence in Employment Discrimination Class Actions*, 78 FORDHAM L. REV. 37, 50 (2009) (providing cites to courts which have permitted certification and those which have not).

95. Disparate treatment occurs when employees can show that they were treated differently than those who are not members of the same protected class. To assert this cause of action, courts require that plaintiffs show that their employer engaged in "intentional" discrimination by taking an adverse employment action on the basis of membership in the protected class. However, the defendant employer is able to avoid liability by demonstrating a nondiscriminatory justification for the action. The plaintiff employees would still be able to prevail if they can show that the justification was simply a pretext. Bornstein, *supra* note 90.

96. In 1971, the Supreme Court in Griggs v. Duke Power Co., 401 U.S. 424 (1971) first enunciated the disparate impact theory of discrimination. Under disparate impact theory, an employment practice that is neutral on its face, but in application has a disproportionately negative effect on a statutorily protected group is unlawful, unless the employer can prove that the practice is job-related and a business necessity. *Id*. at 431. However, liability can still attach if the plaintiff can show an alternative less discriminatory practice. Wards Cove Packing Co. v. Atonio, 490 U.S. 642, 644 (1989). *See* Stewart v. City of St. Louis, 2007 U.S. Dist. LEXIS 38473, at *22 n.4 (E.D. Mo. 2007); 42

on unconscious bias evidence to certify class action lawsuits have not been uniformly successful due to inconsistencies in how lower courts have interpreted *Wal-Mart v. Dukes*.[97]

A. Unconscious Bias in Case Law

In *Wal-Mart v. Dukes*, some 1.5 million female Wal-Mart employees alleged in a class action complaint that the company discriminated against them by denying women equal pay and promotions.[98] Wal-Mart did not have any testing procedures in place for evaluating employees and used discretionary local decision-making with respect to employment matters. The plaintiffs alleged this store-level discretion violated Title VII. The Supreme Court refused to allow the certification of the class explaining that there was "no common question" among the 1.5 million plaintiffs[99] despite social science evidence explaining how local subjective decision-making resulted in the lower pay and lack of promotions of its female employees due to the unconscious biases of the decision-makers.[100]

After the *Wal-Mart* case, it was uncertain whether unconscious bias evidence would be allowed with respect to class action certification relying on statistical analysis.[101] However, the court in *Ellis v. Costco Wholesale*

USCS § 2000e-2(k) (2012). With a disparate impact case, the plaintiff is not required to show intent, but rather the impact of the employment decision can prove discrimination. *Griggs,* 401 U.S. at 432. This standard is codified at 42 U.S.C.S. § 2000e-2(k) (2012).

97. 564 U.S. 338 (2011). *See* Annika L. Jones, Implicit Bias as a Social-Framework Evidence in Employment Discrimination, 165 U. PA. L. REV. 1221, 1231 (2017) (discussing inconsistent treatment of the use of social science evidence in support of the commonality requirement in class action certification after *Wal-Mart*).

98. 564 U.S. at 338.

99. Rules 23(a) and (b) of the Federal Rules of Civil Procedure set forth the requirements for class certification. Rule 23(a) requires: (1) the class is so numerous that joinder of class members is impracticable; (2) there are questions of law or fact common to the class; (3) the claims or defenses of the class representatives are typical of those of the class; and (4) the class representatives will fairly and adequately protect the interests of the class.

100. *See* Camille A. Olson et al., *Implicit Bias Theory in Employment Litigation*, 63 PRAC. LAW. 37 (2017) (explaining the *Wal-Mart* decision and contrasting it with cases where implicit bias theory was accepted).

101. *See* Tanya K. Hernandez, *One Path for Post-Racial Employment Discrimination Cases—The Implicit Association Test Research as Social Framework Evidence*, 32 LAW & INEQ. 309, 332-35 (2014) (noting that the failure of courts to recognize implicit bias evidence leaves plaintiffs without a remedy for discrimination due to unconscious thoughts or feelings that influence their decisions).

Corp.[102] granted the employees' motion for class certification based in part upon unconscious bias expert testimony that the employer's corporate culture created and reinforced stereotyped thinking, which allowed gender bias to infect the promotion process from leadership down. Although in *Wal-Mart v. Dukes*, the court concluded that the plaintiffs were unable to show that the statistical analysis evidence of unconscious bias related specifically to Wal-Mart's employment practices, and thus was insufficient to prove the existence of questions of law or fact common to the particular proposed class per Fed. R. Civ. P. 23(a),[103] the court did find it in *Ellis* at least for class certification purposes.[104] Costco ultimately settled the case for $8 million after the case was remanded.[105]

B. Unconscious Bias in Cases That Were Settled

Although case law has been inconsistent in the treatment of unconscious bias testimony in the court room since *Wal-Mart*, such claims have had enough traction to result in significant out-of-court settlements. The Court's receptivity to unconscious bias arguments in the Home Depot and FedEx class action suits resulted in those cases settling for considerable amounts, $87.5 million and $53.5 million respectively.[106] Unconscious bias has been raised in a number of other class actions against Fortune 500

---

102. 285 F.R.D. 492 (N.D. Cal. 2012).

103. *Wal-Mart,* 564 U.S. at 356; Christine A. Amalfe, *The Limitations of Implicit Bias Testimony—Post Dukes*, GIBBONS P.C. (Mar. 2013), https://perma.cc/5PH3-38LF.

104. Order Granting Plaintiff's Motion for Class Certification; and Denying Defendant's Motion to Eliminate Class Claims, Ellis v. Costco Wholesale Corp., 285 F.R.D. 492, (N.D. Cal. 2012) (No. C-04-3341 EMC).

105. Order Granting Motion for (1) Preliminary Approval of Class Action Settlement; (2) Approval of Class Notice and Notice Plan; and (3) Setting of Schedule for Final Approval at Exhibit 1 § 3; Ellis v. Costco Wholesale Corp., 285 F.R.D. 492 (N.D. Cal. 2012) (No. C04-3341 EMC).

106. Dana Wilkie, *Tips for Rooting Out Hidden Bias*, SOC'Y FOR HUM. RES. MGMT (Dec. 1, 2014), https://perma.cc/DD8M-JVEP.

companies such as American Express,[107] Morgan Stanley,[108] General Electric,[109] Best Buy,[110] Bank of America,[111] and Cargill.[112]

The EEOC has also taken a special interest in the tech industry[113] and in rooting out unconscious bias in the workplace. "As the EEOC looks ahead to the next 50 years, there are two main issues to solve. First is helping companies create a more robust talent pipeline for women and minorities with greater representation at every level of management. This includes identifying and ***eliminating unconscious and conscious bias in the workplace***. This could also ***include more advanced analytics to assess systematic discrimination and patterns of practice both at the company and the industry level***. [emphasis added]."[114] Due to increased awareness of unconscious bias and the failure of the tech industry to meaningfully

107. In 2002, American Express Financial Advisors agreed to pay $31 million to 4,000 female financial advisers who had filed a class action suit for discrimination against them. Jonathan D. Glater, *American Express Unit Agrees to Settle Sex Discrimination Case*, N.Y. TIMES (Feb. 21, 2002), https://perma.cc/H284-3JYB.

108. In 2004, Morgan Stanley agreed to pay $54 million to 340 women to settle a sex discrimination case rather than stand trial on the EEOC's suit that alleged it denied equal pay and promotions to women in a division of its investment bank. Dan Ackerman, *Morgan Stanley: Big Bucks for Bias*, FORBES (Jul. 13, 2004), https://perma.cc/3AWE-WHPG.

109. In 2009, a former high-ranking attorney for General Electric Co. settled her class action gender discrimination lawsuit on behalf of 1,500-1,700 female employees alleging system company-wide gender discrimination against the company for an undisclosed amount. John Christoffersen, *Former GC Settles Gender Discrimination Suit Against General Electric*, LAW.COM (Jan. 28, 2009), https://perma.cc/KNL9-YAM5.

110. In 2011, Best Buy agreed to settle a class action lawsuit accusing the largest U.S. electronics retailer of job discrimination, paying a total of $200,000 to the nine named plaintiffs plus as much as $10 million for legal fees and costs. REUTERS, *Best Buy Settles Class-Action Racial Job Discrimination Lawsuit*, HUFFPOST (June 6, 2011), https://perma.cc/F57U-TMXX.

111. In 2013, Bank of America agreed to pay $39 million to 4,800 women who worked in its Merrill Lynch brokerage operation. Patrick McGeehan, *Bank of America to Pay $39 Million in Gender Bias Case*, N.Y. TIMES (Sept. 6, 2013), https://perma.cc/7QCG-2WUG.

112. In 2014, Cargill Meat Solutions agreed to pay $2.2 million to settle discrimination charges brought by the Department of Labor's Office of Federal Contract Compliance Programs (OFCCP). Daniel McCoy, *Cargill to Pay $2.2 Million to Settle Discrimination Claims*, WICHITA BUS. J. (Jan. 23, 2014), https://perma.cc/A4HS-J53A.

113. U.S. GOV'T ACCOUNTABILITY OFFICE, DIVERSITY IN THE TECHNOLOGY SECTOR: FEDERAL AGENCIES COULD IMPROVE OVERSIGHT OF EQUAL EMPLOYMENT OPPORTUNITY REQUIREMENTS (2017).

114. *Meeting of May 18, 2016—Promoting Diverse and Inclusive Workplaces in the Tech Sector*, U.S. EQUAL EMP. OPPORTUNITY COMMISSION (May 18, 2016), https://perma.cc/2P6W-M3N4.

increase diversity, these organizations should be very concerned with the potential for unconscious bias discrimination suits.[115]

IV. THE BUSINESS CASE FOR WOMEN IN TECH

There are a number of reasons why the tech industry should be disturbed with the lack of women and under-represented minorities (URMs) in their ranks. In addition to potential lawsuits, EEOC investigations, and the inability to hire all of the employees they will need in the near future, there is also great value in creating a diverse workforce. Studies show that women in leadership roles produce significant benefits to companies, including better decision-making, improved working environment, a more collegial atmosphere, and increased innovation.[116] It is well-established that "diverse teams outperform homogeneous teams."[117] Decision-making research shows that diverse teams can avoid group-think.[118] Women are better able to examine different points of view and consider different perspectives than men are.[119] It also does not make sense fiscally to treat women in this way. Research consistently shows that companies with high gender diversity are more profitable and less volatile those with low gender diversity.[120] Based on the available statistics about the benefits of women in tech, it seems counterintuitive for the tech industry to have such a gender imbalance.[121] Tech companies seem to be on the cutting edge, anticipating what the public wants even before the public

---

115. In 2015, Katie Moussouris initiated a class action suit against Microsoft alleging gender discrimination which is still pending. The complaint noted that she and other women earned less than their male counterparts and that men were given preferential treatment in promotions resulting of the unconscious biases of decision-makers. Plaintiff's Motion for Class Certification at 1, Moussouris v. Microsoft Corp., 311 F. Supp. 3d 1223 (W.D. Wa. 2018) (No. C15-1483JLR).

116. *Why It Pays to Invest in Gender Diversity,* MORGAN STANLEY (May 11, 2016), https://perma.cc/5CCA-XKQL.

117. Katherine Phillips, *How Diversity Makes Us Smarter,* SCI. AM. (Oct. 1, 2014), https://perma.cc/U4VZ-XXA7.

118. Anna Johansson, *Why Workplace Diversity Diminishes Groupthink and How Millennials Are Helping*, FORBES (Jul. 20, 2017), https://perma.cc/RQ4J-PWUP.

119. Nancy Wang, *Diversity Is the Key to Startup Success—What Can Early-Stage Founders Do About It?*, FORBES (Nov. 12, 2018), https://perma.cc/DK3X-T3J4.

120. Phillips, *supra* note 117.

121. *See* Erin Griffith, *There's a Simple Way to Improve Tech's Gender Imbalance*, FORTUNE (June 1, 2016), https://perma.cc/4254-T3CY (stating that women influence household spending and are more likely than men to adopt technology trends early).

knows it.[122] Given the reputation loss the tech industry has suffered over the past few years;[123] it does not make sense for them to fail to improve gender equity in their organizations. As Part VII discusses diversity can be improved with the responsible development and use of AI.

A. Financial Benefit

A number of studies have demonstrated a link between gender diversity and corporate performance.[124] For example, Morgan Stanley's Sustainable + Responsible Investment (SRI) and Global Quantitative Research teams have collected and analyzed data on this issue from around the world, and created a proprietary gender-diversity framework for ranking more than 1,600 stocks globally.[125] The results indicate that a company's percentage of female employees is positively correlated with its return on equity.[126] A recent study by Quantopian showed that women-led Fortune 1000 companies posted greater cumulative returns than those of the S&P 500, with an even more pronounced difference after the financial crisis of 2008.[127] The equity returns of the female-led companies were 226% higher than the S&P 500's.[128] The opposite holds true at the other end of the spectrum. Companies in the bottom 25% in terms of gender and ethnic diversity were the least likely to record profits higher than the national industrial average.[129]

---

122. *See* Sameen, *25 Cutting Edge Technology Examples You Won't Believe Exist*, LIST 25 (Sept. 13, 2018), https://perma.cc/XDJ2-4MCN (describing innovative technologies).

123. *See* Jon Swartz, *Regaining Trust Is the No. 1 Issue for Tech in 2019*, BARRON'S (Dec. 28, 2019), https://perma.cc/2GR6-VGPU (stating that Big Tech stock prices dropped significantly in 2018 due to reports of hate speech, fake news, election interference, perceived anti-conservative bias, privacy violations, business dealings in China, and a general loss of trust).

124. MCKINSEY & CO., WOMEN MATTER: TEN YEARS OF INSIGHTS ON GENDER DIVERSITY (2017), https://perma.cc/XA74-8GJ7.

125. *Why It Pays to Invest in Gender Diversity, supra* note 116.

126. *Id.* According to the Credit Suisse Research Institute, companies with at least one woman on their board of directors outperform those without any women by 26%. Press Release, Credit Suisse, Large-Cap Companies with At Least One Woman on the Board Have Outperformed Their Peer Group with No Women on the Board by 26% Over the Last Six Years, According to a Report by Credit Suisse Research Institute (Jul. 31, 2012), https://perma.cc/M5ZN-264U.

127. Pat Wechsler, *Women-Led Companies Perform Three Times Better Than the S&P 500*, FORTUNE (Mar. 1, 2013), https://perma.cc/8BU6-X49P.

128. *Id.*

129. Vivian Hunt et al., *Why Diversity Matters*, MCKINSEY & CO. (Jan. 2015), https://perma.cc/FY9E-4R7C.

It should be especially important to tech companies that gender diversity also leads to greater innovation.[130] A study conducted at the University of Arizona of Fortune 500 companies found that "companies that have women in top management roles experience . . . 'innovation intensity' and produce more patents—by an average of 20% more than teams with male leaders."[131] According to Vivek Wadhwa, a Distinguished Fellow at Carnegie Mellon University's College of Engineering, "We are effectively leaving out half of our population by excluding women from the innovation economy. Given where technology is headed, with technologies advancing exponentially and converging, the skills needed to solve the larger problems require a broad understanding of different fields and disciplines."[132]

B. Increased Numbers of Women in Tech

More women in the ranks leads to more women in leadership roles. More women in leadership roles leads to more women in the ranks. In addition to the financial benefit, women leaders are associated with the hiring of more women throughout the company.[133] Women leaders also tend to lessen the pay gap between men and women.[134] Women CEOs pay high-level women employees more than male CEOs do.[135] Because women tend to support diversity and social responsibility, they also implement more favorable HR policies[136] which also attract more women to the industry.[137] There are two ways to get more women into leadership roles: hire them into the role or promote from within the company. Creating a larger pool of female employees from which to promote greatly increases

---

130. *Why It Pays to Invest in Gender Diversity, supra* note 116; Cristina Diaz-Garcia et al., *Gender Diversity Within R&D Teams: Its Impact on Radicalness of Innovation,* 15 INNOVATION: MGMT., POL'Y & PRAC. 149 (2013).

131. Yoni Blumberg, *Companies With More Female Executives Make More Money—Here's Why*, CNBC (Mar. 2, 2018), https://perma.cc/BQ56-5YT8.

132. Houser, *supra* note 26.

133. Sue Duke, *The Key to Closing the Gender Gap? Putting More Women in Charge*, WORLD ECON. F. (Nov. 2, 2017), https://perma.cc/L8XL-GDY7.

134. Geoffrey Tate & Liu Yang, *Female Leadership and Gender Equity: Evidence from Plant Closure*, 117 J. FIN. ECON. 771 (2015).

135. *Id*.

136. Alison Cook & Christy Glass, *Do Women Advance Equity? The Effect of Gender Leadership Composition on LGBT-Friendly Policies in American Firms*, 69 HUM. REL. 1431, 1435 (2016).

137. Marta Riggins, *5 Meaningful Things Companies Are Doing to Recruit and Retain Women*, LINKEDIN TALENT BLOG (Mar. 8, 2018), https://perma.cc/88R2-FNW6.

the chances of women moving into leadership positions, which will in turn increase the number of women and workplace conditions for women overall. Women leaders are also associated with the hiring of more URMs.[138]

C. Benefits to Women in Leadership

According to Shivaram Rajgopal, the Vice Dean of Research at Columbia Business School, "Women in leadership positions serve as a significant deterrent against a permissive culture towards sexual harassment," he told *The Christian Science Monitor*. "You rarely hear of such issues at Yahoo! where Marissa Mayer was the CEO . . . . [Facebook's Mark] Zuckerberg has [chief operating officer] Sheryl Sandberg to temper the . . . culture."[139] There are numerous studies showing that women do tend to make better managers for a variety of reasons.[140] According to one survey of 7,280 corporate leaders by Zenger and Folkman, women demonstrated higher competencies than men in 12 of the 16 leadership categories surveyed.[141] The two areas in which women outscored men by the highest percentage were taking initiative and driving for results, both important to

138. CENTER FOR EMP. EQUITY, IS SILICON VALLEY TECH DIVERSITY POSSIBLE NOW? 15 (2018), https://perma.cc/5US2-9FXX. A study by PwC showed that 42% of female board directors considered racial diversity to be important compared to only 24% of male directors. Anne Fisher, *Would Your Company Be Any Different If It Had More Women on the Board?*, FORTUNE (May 27, 2015), https://perma.cc/LA3Q-6RBD (citing PWC, 2014 ANNUAL CORPORATE DIRECTORS SURVEY—THE GENDER EDITION (2015)).

139. Houser, *supra* note 26.

140. A meta-analysis of 45 studies comparing the leadership skills of men and women concluded that women tended to be transformational leaders, while men tended to be transactional or laissez-faire leaders. Transformational leaders "establish themselves as role models by gaining followers' trust and confidence. They state future goals, develop plans to achieve those goals, and innovate, even when their organizations are generally successful. Such leaders mentor and empower followers, encouraging them to develop their full potential and thus to contribute more effectively to their organizations. By contrast, transactional leaders establish give-and-take relationships that appeal to subordinates' self-interest. Such leaders manage in the conventional manner of clarifying subordinates' responsibilities, rewarding them for meeting objectives, and correcting them for failing to meet objectives." The conclusion of the meta-analysis was that women are generally more effective leaders, while men are only somewhat effective or hinder effectiveness. Alice Eagly & Linda L. Carli, *Women and the Labyrinth of Leadership*, HARV. BUS. REV., Sep. 2007, https://perma.cc/6ARB-YGQW (referring to Alice H. Eagly et al., *Transformational, Transactional, and Laissez-Faire Leadership Styles: A Meta-Analysis Comparing Women and Men*, PSYCHOL. BULL. (2003)).

141. Jack Zenger & Joseph Folkman, *Are Women Better Leaders than Men?,* HARV. BUS. REV. (Mar. 15, 2012), https://perma.cc/XJF2-STQC.

the tech industry.[142] By promoting women to prominent positions of leadership, companies may be able to prevent some of the more outrageous conduct which leads to sexual harassment claims.[143]

D. The Need to Fill Tech Jobs in 2020

Perhaps the most important financial reason is to maintain U.S. tech industry's dominant position in the world. Currently, there are 11.8 million people employed in the tech field in the U.S.[144] It is estimated that by 2020 the U.S. will not be able to fill the additional 1 million open positions in tech.[145] The tech industry needs to start focusing on solutions that work rather than creating D&I webpages and paying billions to consultants for training programs that do not work.[146] If tech companies keep limiting their hiring pool, it will not be possible to fill all of these needed positions. The top five tech companies in the world are located in the U.S. and account for 45% of the S&P 500's year-to-date gain.[147] Being unable to hire enough employees to perform the work required could be catastrophic not just to these organizations, but to the U.S. economy itself. While the U.S. economy has grown at a rate of 1.5% per year from 2006-2016, the average annual growth rate for the digital economy was 5.6% over the same time period.[148] Because the tech industry has been unable to increase the numbers of women and URMs in any significant way since they began releasing diversity reports, immediate action needs to take place. The tech industry seems mystified by the lack of success, but it is not too difficult to figure it out. What they are doing simply does not work.

---

142. *Id*.

143. *See* Frank Dobbin & Alexandra Kalev, *Training Programs and Reporting Systems Won't End Sexual Harassment. Promoting More Women Will*, HARV. BUS. REV. (Nov. 15, 2017), https://perma.cc/RA43-DF9V (stating that "[h]arassment flourishes in workplaces where men dominate in management and women have little power").

144. THE COMPUTING TECH. INDUS. ASSOC., CYBERSTATES 2019 6 (2019), https://perma.cc/YN4K-QASB.

145. Alison DeNisco Rayome, *CIO Jury: 83% of CIOs Struggle to Find Tech Talent*, TECHREPUBLIC (June 16, 2017), https://perma.cc/SA9M-JNEW.

146. *See* Part IV *infra*.

147. Jamillah Williams, *Diversity as a Trade Secret*, 107 GEO. L.J. 1, 5 (2018).

148. U.S. DEP'T OF COMMERCE, BUREAU OF ECON. ANALYSIS, MEASURING THE DIGITAL ECONOMY: AN UPDATE INCORPORATING DATA FROM THE 2018 COMPREHENSIVE UPDATE OF THE INDUSTRY ECONOMIC ACCOUNTS (2019).

V. CURRENT DIVERSITY AND INCLUSION METHODS DO NOT WORK

After Google started issuing its diversity report in 2014, other tech companies followed suit. In order to address the problem, these companies spent billions of dollars on consultants and training programs but failed to increase the numbers of women and minorities as demonstrated in more recent reports.[149] The most common fix was to institute unconscious bias workshops and mentoring programs.

A. Training Does Not Work

The primary method put forth for tackling the lack of diversity by tech companies has been diversity training. The concept behind this idea is that by explaining to employees and managers their biases, they can actively avoid them. Although billions have been spent on diversity training, studies show that it has had no effect in decreasing bias or in increasing diversity in the companies in which training occurred.[150] Although those taking the training may be able to complete a quiz, they often quickly forget what they learned.[151] Unfortunately, not only does diversity training not work,[152] it actually can cause more harm.[153] Multiple studies have shown that it has the

---

149. According to a Statista report, in 2017 women held 23% of technical positions at Apple, 17% at Facebook, and 19% at Google. Clare Byrne, *4 Charts That Show Tech is Still a Man's World*, WORLD ECON. F. (Aug. 7, 2017), https://perma.cc/GPX2-USBZ. This shows only a slight increase over the 2014 figures of 20%, 15% and 17% respectively. Felix Richter, *Women Vastly Underrepresented in Silicon Valley Tech Jobs*, STATISTA (Aug. 14, 2014), https://perma.cc/CA9B-JQRX. According to data from the Government Accountability Office, between 2007 and 2015, women in tech positions stood still at 22% on average. Joe Davidson, *Mostly White Male Tech Sector Needs Government Help on Diversity*, WASH. POST (Dec. 4, 2017), https://perma.cc/6JEL-P5C8.

150. Frank Dobbin & Alexandra Kalev, *Why Diversity Programs Fail*, HARV. BUS. REV., July-Aug. 2016, at 52, https://perma.cc/FAA3-6HQH [hereinafter Dobbin & Kalev, *Why Diversity*]; Cheryl R. Kaiser et al., *Presumed Fair: Ironic Effects of Organizational Diversity Structures*, 104 J. PERSONALITY & SOC. PSYCHOL. 504 (2013); Joanne Lipman, *How Diversity Training Infuriates Men and Fails Women*, TIME (Jan. 25, 2018), https://perma.cc/XD5D-XNS2; Mike Noon, *Pointless Diversity Training: Unconscious Bias, New Racism and Agency*, 32 WORK, EMP'T AND SOC'Y 198, 203 (2017).

151. Dobbin & Kalev, *Why Diversity, supra* note 150, at 54.

152. An EEOC Task Force found in 2016 that thirty years of social science research has failed to show that training reduces sexual harassment. Christina Folz, *No Evidence that Training Prevents Harassment, Finds EEOC Task Force*, SOC'Y FOR HUM. RES. MGMT. (June 19, 2016), https://perma.cc/US2L-LKSE; Jena McGregor, *To Improve Diversity, Don't Make People Go to Diversity Training. Really*, WASH. POST (July 1, 2016), https://perma.cc/7YF8-PXC4.

153. Susan Bisom-Rapp, *Sex Harassment Training Must Change: The Case for Legal Incentives for Transformative Education and Prevention,* 71 STAN. L. REV. ONLINE 62 (2018);

potential to actually increase bias.[154] In fact, men tend to interpret required diversity training as an assignment of blame. [155] Instead of encouraging the equitable treatment of women and minorities, men interpreted the message as requiring them to provide special treatment to women or minorities or demonstrated a fear of losing their jobs to women and minorities.[156] Since 2014, some 70,000 of its employees participated in Google's diversity training program.[157] However, their 2018 report indicates that the composition of women barely budged over this timeframe.[158] In fact, this diversity training led one man to sue for discrimination.[159]

Training can cause unanticipated harm to the workplace culture. Studies shows that training can actually cause women and URMs to believe that their co-workers are more biased than they actually are.[160] In addition, research reveals that when employees are told that biases are "unconscious," they feel as though they cannot do anything to change their behavior as bias is just "human nature."[161] Managers who were told that stereotypes against women are common, felt more comfortable indicating

---

Dobbin & Kalev, *supra* note 150, at 54; Kaiser, *supra* note 150, at 504. In one study, white men who participated in a hiring simulation where participants received either a pro-diversity or neutral message experienced "cardiovascular reactivity" (a negative physiological response) in response to the diversity message. Tessa L. Dover et al., *Members of High-Status Groups Are Threatened by Pro-Diversity Organizational Messages*, 62 J. EXPERIMENTAL SOC. PSYCHOL. 58 (2016).

154. Dobbin & Kalev, *Why Diversity, supra* note 150; Kaiser, *supra* note 150; Lipman, *supra* note 150; Noon, *supra* note 150.

155. Lipman, *supra* note 150.

156. *Id.*

157. David Pierson & Tracey Lien, *Diversity Training Was Supposed to Reduce Bias at Google. In Case of Fired Engineer, It Backfired*, L.A. TIMES (Aug. 9, 2017), https://perma.cc/K4A4-MYNE.

158. The Google Diversity Report 2014-2018 shows that in 2014 30.6% of its workforce was female. In 2018, that figure is 30.9% demonstrating no significant improvement. GOOGLE, GOOGLE WORKFORCE COMPOSITION 2014-2018, https://perma.cc/P7JH-2Y4S.

159. "I went to a diversity program at Google and . . . I heard things that I definitely disagreed with," said Damore, a former Google employee. "[T]here was a lot of, just, shaming—'No, you can't say that, that's sexist'; 'You can't do this.'" Pierson & Lien, *supra* note 157.

160. Lipman, *supra* note 150.

161. Michelle M. Duguid & Melissa C. Thomas-Hunt, *Condoning Stereotyping? How Awareness of Stereotyping Prevalence Impacts Expression of Stereotypes*, 100 J. APPLIED PSYCHOL. 343 (2015); *see also* Jeffrey Halter, *Unconscious Bias: The Fatal Flaw in Diversity Training*, NETWORK EXEC. WOMEN (Jan. 9, 2018), https://perma.cc/YF2L-SN7S (describing how men leave unconscious bias training thinking that it is a "Get out of Jail Free" card because they cannot do anything about it).

that they did not want to hire women because the women were "unlikable" thus giving license to flaunt their biases.[162] One study showed that five years after diversity training there were 6% fewer black women in management positions.[163] In addition, the mere existence of a diversity training program can result in men being less likely to believe that gender discrimination exists at their company despite evidence to the contrary.[164]

B. Mentoring Programs Do Not Work

Tech companies have also tried mentoring programs. Mentoring is simply a way to shift the burden to women to fix the discrimination they encounter in the workplace. Teaching women how to negotiate or "lean in" does not work. In fact, women who try to negotiate salary are viewed as difficult to work with.[165] Women who are taught how to be leaders are judged more harshly than their male counterparts exhibiting the same behaviors.[166]

Mentoring is essentially asking those who are marginalized to advocate for themselves after receiving advice. When that advice is given by a man, it may not be entirely applicable as men do not face the same issues or have the same experiences as women do in the workplace. Women mentors may provide valuable advice, but often, have little influence in making promotion decisions.[167] In addition, there are simply not enough women in managerial or leadership positions in these companies to take on the mentoring of early career women and URMs and doing so puts an additional burden on those female mentors potentially harming their own careers. Some studies have shown that women who advocate for diversity are actually penalized in

162. Lipman, *supra* note 150; *see also* Kaiser, *supra* note 150 at 505 ("[A]sking people to suppress their stereotypes can inadvertently increase stereotyping and prejudice").

163. Dobbin & Kalev, *Why Diversity, supra* note 150, at 59.

164. Kaiser, *supra* note 150, at 506; Rachel Thomas, *How Diversity Branding Hurts Diversity*, MEDIUM (Dec. 7, 2015), https://perma.cc/BFN6-QQQ2.

165. Bowles, *supra* note 47.

166. Snyder, *Abrasiveness*, *supra* note 45.

167. Nancy Novak, *Gender Diversity Doesn't Happen Without Advocacy*, CIO (Apr. 16, 2018), https://perma.cc/E6ZA-5L5S.

their performance evaluations.[168] While mentoring can help women feel included, it has not been shown to advance a woman's career.[169]

These measures have failed because they do not address the underlying reason for the lack of diversity—subjective decision-making by humans, which is the rule in 80% of companies.[170] Humans simply cannot make objective decisions, and this failure harms women and URMs more significantly than others with respect to employment decisions. Addressing the issue of bias and noise not only helps to increase diversity, it has been proven to result in more qualified applicants, better hires, better promotions, and better retention rates.[171] The following Part explains the reasons underlying poor decision-making by humans.

VI. UNCONSCIOUS BIAS AND NOISE

Social scientists have discovered that unconscious errors of reasoning distort human judgments (unconscious bias) and random chance variability in decisions (noise) occurs more often than people realize.[172] Unconscious biases occurs when people are unaware of the mental shortcuts they use to process information.[173] Noise refers to variability in human decision-making due to chance or irrelevant factors.[174] Daniel Kahneman gives the example of a faulty scale. If your scale consistently reads 10 pounds heavier than you know yourself to be, it is biased. If every time you step on the scale

---

168. Stefanie K. Johnson & David R. Hekman, *Women and Minorities Are Penalized for Promoting Diversity*, HARV. BUS. REV. (Mar. 23, 2016), https://perma.cc/HBK2-M47L.

169. SYLVIA ANN HEWLITT, FORGET A MENTOR, FIND A SPONSOR THE NEW WAY TO FAST-TRACK YOUR CAREER (2013).

170. *Supra* note 8, at 11.

171. *See* Part VI *infra*.

172. *See* Jim Holt, *Two Brains Running*, N.Y. TIMES (Nov. 25, 2011), https://perma.cc/UAS7-JK2L (discussing Kahneman and Tversky's series of experiments that "revealed twenty or so 'cognitive biases'—unconscious errors of reasoning that distort our judgment of the world."); J. Nathan Matias, *Bias and Noise: Daniel Kahneman on Errors in Decision-Making*, MEDIUM (Oct. 17, 2017), https://perma.cc/BSX7-YYF3 (discussing Kahneman's series of experiments at an insurance company that revealed unexpected variance in decisions of different underwriters as well as within the individual underwriter him- or herself).

173. PARADIGM, MANAGING UNCONSCIOUS BIAS—STRATEGIES TO ADDRESS BIAS & BUILD MORE DIVERSE, INCLUSIVE ORGANIZATIONS 2 (2016), https://perma.cc/6STW-UUF9.

174. Daniel Kahneman et al., *Noise: How to Overcome the High, Hidden Cost of Inconsistent Decision Making*, HARV. BUS. REV., Oct. 2016, https://perma.cc/CLF8-WJEM [hereinafter Kahneman et al., *Noise: How to Overcome*].

you get a different number, it is noisy.[175] Although people believe they are objective, noise and bias result in inaccurate and inconsistent decisions.[176] Prior to discussing how the responsible development and use of AI can mitigate these human errors impacting employment decisions, the behavioral science background on the many flaws in human judgment will be explored.

A. Unconscious Bias

In Thinking, Fast and Slow, Kahneman explains how our brains operate using two systems: System 1 (Fast Thinking) operates automatically (such as driving a car), whereas System 2 (Slow Thinking) involves reasoning (such as solving a math problem).[177] Both systems have deficiencies impacting our ability to make objective decisions. The main problem with System 1 thinking is it is not prone to doubt or careful reflection because it involves mental short-cuts to reach a conclusion. This means we believe the answers we come up with using System 1 thinking are accurate and we do not spend time analyzing how we got there, resulting in unjustified confidence in our System 1 decisions. This in turn leads to judgments which are neither accurate nor logical but become endorsed by System 2 and turn into deep-rooted beliefs.[178] While Kahneman goes on to explain the potential errors in judgment when people rely on System 1 thinking, the most important takeaway is that not only are our judgments untrustworthy, but the more decisions we make based on these biases, the stronger the neural connections become that reinforce our belief in the biases' conclusion. These cognitive biases affect opinions on social issues (as we can see with prejudice) and hence affect social institutions (such as the workplace) that rely on individuals to make rational (unbiased) judgments.

Social science literature regarding unconscious bias has established that we are generally unaware of our own prejudices. We routinely make decisions and judgments based on incomplete information/knowledge and a lack of objective truth. We use "mental shortcuts," known as heuristics, to fill in gaps in our knowledge with similar data from past experiences and

---

175. *Id.*
176. MAHZARIN BANAJI & ANTHONY G. GREENWALD, BLINDSPOTS: HIDDEN BIASES 152 (2016).
177. DANIEL KAHNEMAN, THINKING, FAST AND SLOW 20-22 (2011) [hereinafter KAHNEMAN, THINKING].
178. Ameet Ranadive, *What I Learned from "Thinking, Fast and Slow,"* MEDIUM (Feb. 20, 2017), https://perma.cc/32KW-X7HJ.

cultural norms. The notion of cognitive biases was introduced to the modern era by Amos Tversky and Daniel Kahneman in 1972.[179] Essentially, Tversky and Kahneman exposed that while heuristics are useful, they also lead to errors in thinking.[180] Their theories also explain why the tech industry seems unable to correct its diversity problem. Representative bias, for example, occurs when people conclude that someone works in a given profession due to how similar that person appears to a stereotype.[181] When a hiring agent internally views their ideal applicant as a young white male from an Ivy League school, this stereotype limits their ability to value other types of candidates.

The reason employment interviews are still handled by humans with biases is due to the validity illusion. As Kahneman and Tversky explain, people tend to overrate their own ability to make accurate predictions.[182] This validity illusion exists because of confirmation bias—focusing on information that fits our prediction and discarding that which does not.[183] Therefore, during an interview, the interviewer, without realizing it, may notice only the information that confirms their pre-existing belief rather than the information relevant to the position.[184]

---

179. Daniel Kahneman & Shane Frederick, *Representativeness Revisited: Attribute Substitution in Intuitive Judgment*, *in* HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT 49, 51-52 (Thomas Gilovich, Dale Griffin, & Daniel Kahneman eds., 2002).

180. Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 SCI. 1124, 1124 (1974) [hereinafter Tversky & Kahneman, *Judgment*].

181. *Id.* at 1124. For example, if someone's experience as a child was to see male doctors and female nurses during office visits, when they hear the word "doctor" an image of a male will immediately come to mind. Recall the riddle: "[A] father and son are in a horrible car crash that kills the dad. The son is rushed to the hospital; just as he's about to go under the knife, the surgeon says, 'I can't operate—that boy is my son!' Explain." A study at BU found that most of the participants could not answer because they did not consider the possibility that the surgeon was the boy's mother. Rich Barlow, *BU Research: A Riddle Reveals Depth of Gender Bias*, BU TODAY (Jan. 16, 2014), https://perma.cc/8WL2-WL2S.

182. Daniel Kahneman & Amos Tversky, *On the Psychology of Prediction*, 80 PSYCHOL. REV. 237, 249 (1973).

183. Raymond S. Nickerson, *Confirmation Bias: A Ubiquitous Phenomenon in Many Guises*, 22 REV. GEN. PSYCHOL. 175, 175-76 (1998) (explaining scholarly definitions of confirmation bias, each line of study emphasizing different features of behavior).

184. Laszlo Block, *Here's Google's Secret to Hiring the Best People*, WIRED (Apr. 7, 2015), https://perma.cc/Q5Z8-E8F4 (stating that a University of Toledo study concluded "most interviews are a waste of time because 99.4 percent of the time is spent trying to confirm whatever impression the interviewer formed in the first ten seconds").

We tend to prefer people who look like us, think like us and come from backgrounds similar to ours.[185] This is known as the affinity bias[186] or "in-group favoritism."[187] One Silicon Valley investor explained that Silicon Valley's diversity problem comes from mistaking certain traits, such as personality type or alma mater, for actual skills: "You don't necessarily have to be biased against somebody. You can be biased in favor of somebody." [188] This favoritism is "responsible for many discriminatory outcomes, from relying on homogenous employee referral networks for new hires to giving more favorable performance reviews to employees with similar personality traits as you."[189] In fact, part of the problem is that not only are these biases unknown to those who possess them, but they are particularly difficult for individuals and organizations to correct.[190]

While we believe ourselves to be open-minded and objective, research shows that the beliefs and values acquired from family, culture and a lifetime of experiences heavily influence how we view and evaluate others. Stereotyping involves making assumptions based on the group an individual belongs to, which has enormous implications for women working in male dominated fields. One study reported how potential employers systematically underestimated the mathematical performance of women, which resulted in hiring less qualified males over higher-performing women.[191] A recent study of college students also discloses that males tend to believe they are smarter than women (they were about two times as

---

185. *See* Caren B. Goldberg, *Relational Demography and Similarity-Attraction in Interview Assessments and Subsequent Offer Decisions: Are We Missing Something?*, 30 GRP. & ORG. MGMT. 597, 598 (2005) (explaining, and then complicating, the prevailing view that recruiters favor applicants similar to themselves).

186. ELIZABETH LINOS & JOANNE REINHARD, CHARTERED INST. OF PERSONNEL & DEV., A HEAD FOR HIRING: THE BEHAVIOURAL SCIENCE OF RECRUITMENT AND SELECTION, at 7, 24 (2015).

187. Marilynn B. Brewer, *The Importance of Being We: Human Nature and Intergroup Relations*, 62 AM. PSYCHOL. 728, 729 (2007).

188. Tracy Ross, *The Unsettling Truth About the Tech Sector's Meritocracy Myth*, WASH. POST (Apr. 13, 2016), https://perma.cc/8EDF-6WDX.

189. *Id.*

190. Valerie Martinelli, *The Truth About Unconscious Bias in the Workplace*, TALENTCULTURE (Mar. 31, 2017), https://perma.cc/J7KS-NZ5B.

191. CHRISTIANNE CORBETT & CATHERINE HILL, AM. ASS'N OF U. WOMEN, SOLVING THE EQUATION: THE VARIABLES FOR WOMEN'S SUCCESS IN ENGINEERING AND COMPUTING 3 (2015). Even the President of Harvard University publicly attributed the lack of females in science and math at top universities to innate differences between the sexes; although Larry Summers apologized, he resigned shortly thereafter. Alan Finder et al., *President of Harvard Resigns, Ending Stormy 5-Year Tenure*, N.Y. TIMES (Feb. 22, 2006), https://perma.cc/AJ48-38ND.

likely to report this).[192] This is especially ironic given that for a number of years colleges have been excluding well-qualified female applicants from college admissions, preferring instead the less-qualified male applicants in order to achieve gender parity.[193]

Status quo bias exists when people prefer situations to remain the same.[194] Thus, if the workplace is primarily male and has been for a long time, there is an inherent bias against introducing greater numbers of women into that environment because it will change how "things have always been done." Part of the reason for this bias is loss aversion. People weigh the potential for loss more heavily than the potential for gain. Even when the workplace will gain advantages to the increase in the number of women, status quo bias will cause people to bristle against the idea because they fear change. Another reason is the exposure effect. When you have been exposed to a workplace that is predominantly male over a long period of time, it becomes a preference. Malcolm Gladwell also describes research from psychology and behavioral economics revealing mental processes that work rapidly and automatically from relatively little information.[195] Unfortunately, snap judgments can be the result of subconscious racial or gender bias. Gladwell explains that prejudice can operate at an intuitive unconscious level, even in individuals whose conscious attitudes are not prejudiced.[196] Another issue is that humans do not make consistent decisions.

---

192. *See* Katelyn M. Cooper et al., *Who Perceives They Are Smarter? Exploring the Influence of Student Characteristics on Student Academic Self-Concept in Physiology*, 42 ADVANCES PHYSIOLOGICAL EDUC. 200, 205 fig.2 (2018).

193. Marcus B. Weaver-Hightower, *Where the Guys Are: Males in Higher Education*, 42 CHANGE: MAG. HIGHER LEARNING, July 8, 2010, at 29, 30.

194. William Samuelson & Richard Zeckhauser, *Status Quo Bias in Decision-Making*, 1 J. RISK & UNCERTAINTY 7, 8 (1988). Samuelson & Zeckhauser give the example of choosing a sandwich for lunch that you have once had before because of the perceived risk in choosing a different sandwich that you might not like. *Id.* at 10. In addition, they point to the popular example of "New Coke" being the preferred taste in a blind test but not preferred in the marketplace where consumers see and prefer the Coke with which they are familiar. *Id.* at 11.

195. *See, e.g.*, MALCOLM GLADWELL, BLINK: THE POWER OF THINKING WITHOUT THINKING 8-11 (2005).

196. *Id.* at 96-97.

B. Noise

According to Kahneman, although organizations expect consistency in decision-making, this is seldom the case. "The problem is that humans are unreliable decision makers; their judgments are strongly influenced by irrelevant factors, such as their current mood, the time since their last meal, and the weather. This variance in decisions is known as noise."[197] Research demonstrates that even looking at identical information, managers' decisions will vary from those of other managers. Noise also occurs when managers make decisions inconsistent with their prior decisions.[198] This inconsistent decision-making costs organizations billions in lost productivity, and exposes them to potential liability.[199]

In making employment decisions, not only do we want to eliminate or reduce unconscious biases, we also want to eliminate or reduce noise. Consistent decisions are more equitable and can help avoid claims of discrimination.[200] The difficulty is that not only are human decision-makers unaware of the bias and noise in their decisions,[201] these problems may not be detectable by other humans.[202] Because professionals believe they can accurately predict who will make a good hire for a particular position, they are not likely to believe that they are inconsistent in their own decision-making or that their decisions vary significantly from those of their colleagues.[203] Relying on humans to make employment decisions produces not only biased and inconsistent results, but also less accurate ones. Kahneman and others have suggested incorporating AI into the decision-making process can mitigate the impact of illogical human decisions.[204]

---

197. Kahneman et al., *Noise*: *How to Overcome, supra* note 174.

198. *Id.*

199. *Id.*

200. *See* Part VIII *infra*.

201. Mendez, *supra* note 82.

202. Kahneman et al., *Noise*: *How to Overcome, supra* note 174.

203. *Guide: Use Structured Interviewing*, GOOGLE: RE:WORK, https://perma.cc/HDB3-XLRQ (archived Apr. 10, 2019). This overconfidence bias increases with the addition of more information. Rather than increasing accuracy, human decision-makers fail to adjust for their cognitive limitations. Claire I. Tsai et al., *Effects of Amount of Information on Judgment Accuracy and Confidence*, 107 ORG. BEHAV. & HUM. DECISION PROCESSES 97, 97 (2008).

204. James Pethokoukis, *Nobel Laureate Daniel Kahneman on AI: 'It's Very Difficult to Imagine That with Sufficient Data There Will Remain Things That Only Humans Can Do'*, AEIDEAS (Jan. 11, 2018), https://perma.cc/NP4P-5GMY.

VII. USING AI TO REDUCE BIAS/NOISE IN HUMAN DECISION-MAKING

The reasons for the lack of women in the tech industry can be attributed to several factors, lower levels of hiring and promotion, unfriendly workplace-culture, and the inability of tech companies to keep the women they do hire. Ultimately, the underlying cause of the lack of diversity stems from flawed human decision-making, especially by managers and others who influence the hiring, promotion, and retention of women. As explained in the previous Part, human decisions are fraught with unjustified variability and bias. It is possible, however, to circumvent many of the problems inherent in faulty human decision-making through the responsible use of AI.[205] As mentioned in the Introduction, an algorithm is a series of rules programmed into a computer, while AI is a broader term covering the process of using a machine to perform a function formerly performed by humans, such as conducting word searches, testing for skills, and analyzing data and producing outcomes.[206] Some utilizations of AI include anonymizing resumes and interviewees, performing structured interviews through online submission or chatbots, parsing job descriptions, using neuroscience games which can identify traits, skills and behaviors which can then be used to match candidates with open positions, predicting which employees are looking to leave to improve retention, mining employee reviews for biased language, and standardizing promotion decisions.

A. Tackling Unconscious Bias

Companies that have moved from traditional recruiting methods to using AI have found success in creating a more diverse slate of candidates.[207] One method shown to increase the diversity of candidates is

---

205. EXEC. OFFICE OF THE PRESIDENT, BIG DATA: A REPORT ON ALGORITHMIC SYSTEMS, OPPORTUNITY, & CIVIL RIGHTS 14 (2016), https://perma.cc/MEK5-8EBT. The report can be read as contemplating only rule-based, algorithmic decision-making using big data—so-called good-old fashion artificial intelligence—but its basic point applies equally to the opportunity for machine learning using big data.

206. Donahue, *supra* note 14. Data mining is the process of a machine reviewing a data set to find patterns and relationships between variables, such as what type of coding skills translate to good performance in a particular position. Machine learning is the ability of a machine to improve its performance of an outcome by modifying the algorithm on its own without the intervention of the programmer.

207. Pymetrics, for example, reports that traditional resume-reviewing results in women and URMs being at a 50-67% disadvantage while companies reported an

the use of algorithms to remove race, gender, and national origin from the initial evaluation process.[208] For example, Unbias.io, removes faces and names from LinkedIn profiles to reduce the effects of unconscious bias in recruiting, while Interviewing.io eliminates unconscious bias by providing an anonymous interviewing platform.[209] Another company, Entelo, anonymizes interviewing by removing all indication of gender or race.[210] Talent Sonar writes gender-neutral job descriptions and hides applicants' names, gender, and other personal identifiers from those doing the hiring.[211] Textio, a program that rewords job ads to appeal to a broader demographic, managed to successfully increase the Australian software company Atlassian's percentage of women among new recruits from 18% to 57%.[212]

Removing names and gender identifications from resumes results in the hiring of more women.[213] One problem is that although organizations will

---

increase in diversity of 20-100% of their hires. PYMETRICS: EMPLOYERS, https://perma.cc/JL3X-NYWE (archived May 21, 2019).

208.

> Interviewing.io allows candidates to practice interviewing with former executives at tech companies. When a candidate becomes proficient at the practice interviews, they can be invited to interview anonymously at tech companies and can skip directly to the tech phase of the interview—that is, the part of the interview where the candidate is tested on solving algorithmic coding problems or some other technical problem. In other words, Interviewing.io allows people to skip the initial in-person screening that is currently one point at which bias can creep in during traditional interviews. Finally, if a candidate feels they have done well at the tech interview, they can choose to "unmask" themselves—at which point the next phase is generally an onsite interview. So, while the face-to-face interaction that sometimes triggers bias does still occur with Interviewing.io, it occurs at a much later stage. By this time many candidates will have already demonstrated facility for many of the tasks associated with the job and perhaps any bias is reduced or counteracted by awareness of their technical skill.

Nancy Leong, *The Race Neutral Workplace of the Future*, 51 U.C. DAVIS L. REV. 719, 724 (2017).

209. *Id.*

210. "The software allows recruiters to hide names, photos, school, employment gaps and markers of someone's age, as well as to replace gender-specific pronouns—all in the service of reducing various forms of discrimination." Rebecca Greenfield & Riley Griffin, *Artificial Intelligence Is Coming for Hiring, and It Might Not Be That Bad,* L.A. TIMES (Aug. 10, 2018), https://perma.cc/RK3J-BS8V.

211. Alsever, *supra* note 84.

212. Simon Chandler, *The AI Chatbox Will Hire You Now*, WIRED (Sept. 13, 2017), https://perma.cc/QN6T-WZFZ.

213. Claire Cain Miller, *Is Blind Hiring the Best Hiring?*, N.Y. TIMES MAG. (Feb. 28, 2016), https://perma.cc/T76G-ZXQX. One study presented at the National Academy of Sciences in 2012 found that "female student[s] [were judged] to be less competent and

hire and promote men based on their potential, women tend to be hired and promoted only on proven past performance.[214] Even when resumes are identical, organizations choose a man for an interview more often than the women with identical credentials.[215] By hiding the attributes of the applicants that could give rise to biased assumptions, the possibility of discrimination on this basis is removed.[216]

Formalized hiring and promotion criteria also help reduce subjectivity.[217] Slack, for example, uses "white board interviews" where candidates solve problems at home. Organizations can prevent bias by first removing a candidates' identifying information, and then evaluating the candidates' work against a comprehensive checklist.[218] Structured interviews are another proven way to eliminate or reduce bias.[219] During a

---

less worthy of being hired than identical male student[s], and also [were] offered a smaller starting salary and less career mentoring." Corinne A. Moss-Racusin et al., *Science Faculty's Subtle Gender Biases Favor Male Students,* 109 PROC. NAT'L ACAD. SCI., 16474, 16477 (2012).

214. Carter & Silva, *supra* note 58.

215. Rhea E. Steinpreis, Katie A. Anders, & Dawn Ritzke, *The Impact of Gender on the Review of the Curricula Vitae of Job Applicants and Tenure Candidates: A National Empirical Study*, 41 SEX ROLES 509, 522-24 (Oct. 1999).

216. Alsever, *supra* note 84.

217. Eric Luis Uhlmann & Geoffrey L. Cohen, *Constructed Criteria Redefining Merit to Justify Discrimination*, 16 PSYCHOL. SCI. 474, 474 (2005). One study at Yale showed that employers often justify bias after the fact. When those evaluating candidates for a police chief position saw a female name for the candidate who had a degree and a male name for the candidate who did not, the evaluators indicated that they chose the man because "street smarts" were the most important factor. When the names were reversed, the evaluators justified choosing the man because the degree was the most important factor. However, when the hiring criteria was formalized before they looked at applications (i.e. reviewers decided in advance whether formal education or street smarts was more important), bias was reduced and the candidate matching the criteria was more likely to be chosen. *Id*.

218. Lauren Romansky & Emily Strother, *Slack's Unique Diversity Strategy Offers Some Lessons for Silicon Valley and Beyond*, TALENT DAILY (May 15, 2018), https://perma.cc/J452-JCTG. Interview.io uses software that disguises voices so that the interviewer cannot determine the sex of the interviewee. Slack has successfully increased its diversity through AI measures such as the ones described in this paper. Slack employs 44.7% women and an independent third party confirmed Slack's pay equity achievement. Their workforce consists of 12.6% underrepresented minorities and 8.3% identify as LGBTQ. *Diversity at Slack: An Update on Our Data, April 2018*, SLACK BLOG (Apr. 2018), https://perma.cc/XK5T-UBDG. Another future fix suggested by Leong is to incorporate virtual reality offices where employees interact with one another via a virtual world where employees could choose any avatar and everyone would understand that it would not necessarily reflect their gender, race or national origin. Leong, *supra* note 208.

219. Michael McDaniel et al., *The Validity of Employment Interviews: A Comprehensive Review and Meta-Analysis*, 79 J. APPLIED PSYCHOL. 599, 599 (1994).

structured interview, each candidate answers questions identical to those asked of the other interviewees. According to Loren Larsen, CTO of HireVue, "[b]y using a structured interview process where the questions are carefully designed and the answers are analyzed and compared with hundreds or thousands of samples, we can begin to predict job performance more accurately than human evaluators and begin to remove bias that has always been present in the hiring process."[220] Mya Systems created a chatbot that recruits, interviews, and evaluates job candidates using performance-based questions.[221] The chatbot compares the answers with the job requirements and answers the candidates' questions about the position and company.[222] This technological capability permits the initial evaluation of the candidate to be based on predetermined criteria without human biases creeping in.

Fortune magazine has identified some 75 startups entering the field of AI hiring programs.[223] Not only can companies use AI to anonymize candidates; it can also discover desired attributes in candidates and employees.[224] Although Amazon was not successful in this regard, as discussed in Part VII *infra*, Pymetrics has succeeded in increasing gender diversity through AI. In addition to creating custom unbiased gamified assessments of candidates, they have also prevented bias from creeping in by continually auditing their own algorithms for biased outcomes.[225] The way Pymetrics works is that candidates engage in neuroscience games which allows the algorithm to match candidates with positions based on identified factors.[226] It creates an algorithm to match candidates with its client's open jobs. Unlike traditional pre-employment tests, in which candidates can create answers based on what they think the employer is looking for, games test actual behaviors, skills and traits shown to be

220. Melissa Locker, *How to Convince a Robot to Hire You*, VICE (Oct. 17, 2018), https://perma.cc/LEF8-3DV5.

221. Chandler, *supra* note 212.

222. *Id*.

223. Alsever, *supra* note 84.

224. Robert Bolton, *Artificial Intelligence: Could an Algorithm Rid Us of Unconscious Bias?*, PERSONNEL TODAY (Nov. 16, 2017), https://perma.cc/4RSN-FTB4 (explaining that an algorithm can be designed to assess "performance against a predetermined personality profile" rather than asking someone if they have the right experience).

225. Greenfield & Griffin, *supra* note 210.

226. PYMETRICS, *supra* note 207. Pymetrics' technology can also match candidates with other positions should they not have the skills needed for the job they are interested in.

associated with top performers in the company.[227] Unilever reported that since they began using Pymetrics, they have doubled the number of applicants they hire after the final round of interviews, increased revenue by hiring a better quality of employee, and increased the diversity of their applicant pool.[228]

One of the advantages of moving to online assessments and games is the ability to locate non-traditional applicants. Many people with technical skills have not gone to college or have left the workforce for a period of time. Traditional resume screening can eliminate qualified candidates by discarding those without a college degree or who have a large gap in their work history. By focusing on the skills rather than pedigree of applicants, companies can not only locate more qualified employees best suited to the position, but also a more diverse group. HackerRank notes that because 65% of programmers are at least partially self-taught,[229] if a company is conducting their own search on college campuses, they are likely choosing from a smaller, wealthier, homogenous set of candidates and ignoring a large pool of qualified, more diverse candidates. HackerRank allows anyone to go online and participate in coding and other technical assessments so that the hiring company can assess the applicant's skills, rather than pedigree.[230] GapJumpers reports that their skills-testing AI resulted in 60% of the women and URMs applying getting an initial interview, up from 20% with resume screening.[231]

Another subcomponent of AI, data mining, involves the discovery of patterns and relationships in a data set. [232] Data mining can be used to

---

227. Chris Ip, *To Find a Job, Play These Games*, ENGADGET (May 4, 2018), https://perma.cc/Z3TS-MQYD.

228. Wanda Thibadeaux, *Unilever Is Ditching Resumes in Favor of Algorithm-Based Sorting*, INC. (June 28, 2017), https://perma.cc/3XVC-4XY2. Infor reports boosting employee diversity for some of its clients by 26%, and Famous Footwear achieved a 33% lower turnover rate after adopting AI in its hiring process. Jill Strange, *Cut Out the Unconscious Bias in Your Hiring with Smart Data*, DIGINOMICA (Jul. 10, 2017), https://perma.cc/3BJE-BXXL. Another company, Stella, also matches applicant with potential positions through the use of AI. Like Pymetrics, it also audits their algorithms to detect bias. Greenfield & Griffin, *supra* note 210.

229. VIVEK RAVISANKAR, HACKERRANK, STUDENT DEVELOPER REPORT 2018 2 (2018), https://perma.cc/BVQ4-R84S.

230. I refer to specific companies in this paper as examples of the wide range of services available today. For a more inclusive list, see Kayla Kozan, *The 38 Top Recruiting Software Tools of 2019*, IDEAL BLOG (Feb. 4, 2019), https://perma.cc/DK4X-3ML7.

231. Miller, *supra* note 213.

232. *Data Mining: What It Is and Why It Matters*, SAS INSIGHTS, https://perma.cc/5C85-BSKF (archived May 15, 2019).

extract information from large data sets to identify which factors are associated with retention.[233] Data mining can discover correlations, patterns, and variances within huge data sets to predict outcomes.[234] The following examples used internal employee data for the analysis, not data mined from the internet.[235] Retention is an important part of keeping the women that are hired in tech. In an early study performed by Chien and Chen (2008), they successfully demonstrated how using "decision tree analysis to discover latent knowledge and extract the rules" to assist in personnel selection decisions and generate recruiting and human resource management strategies.[236]

AI can analyze employee data to reduce turnover, increase efficiency, improve employee engagement, and predict job performance. Sysco's AI program improved its retention rate from 65% to 85% by tracking employee satisfaction scores, which allowed them to institute immediate improvements and saved Sysco nearly $50 million in hiring and training costs for new associates.[237] AI advances in recent years along with the availability of lower cost cloud storage has vastly improved the ability of machines to analyze large data sets.[238] Deloitte's Bersin Study on People Analytics indicates that the use of AI strongly correlates with improved

---

233. EXEC. OFFICE OF THE PRESIDENT, *supra* note 205.

234. SAS INSIGHTS, *supra* note 232.

235. Data mined from the internet or social media presents significant problems that technology, at least today, cannot fully address. Companies such as IBM and Google are working on software that can flesh out and correct bias in large data sets, but for the purposes of this paper, data mining examples are limited to known datasets.

236. Chen-Fu Chien & Li-Fei Chen, *Data Mining to Improve Personnel Selection and Enhance Human Capital: A Case Study in High-Technology Industry*, 34 EXPERT SYS. WITH APPLICATIONS 280, 281 (2008). A similar study was performed in 2013 by Azar et. al with respect to the banking industry. The study used data mining and decision tree analysis to create recruiting rules to improve employee retention. After examining 26 variables, only five were found to impact job performance. This time, gender was identified and found to have no impact on "upgrade status." The research design expressly indicated that "[t]he present thesis takes this approach further *by avoiding opinion-based methods* that are traditionally used in the selection of new employees" (emphasis added). The study concluded that data mining techniques are an extremely important tool to help managers discover covert knowledge, which can assist in human resource management. Adel Azar et al., *A Model for Personnel Selection with a Data Mining Approach: A Case Study in a Commercial Bank*, 11 SA J. HUM. RES. MGMT. 449, 449 (2013).

237. Thomas H. Davenport, Jeanne Harris, & Jeremy Shapiro, *Competing on Talent Analytics*, HARV. BUS. REV., Oct. 2010, at 54, 56, https://perma.cc/BS8T-86KE.

238. AI is being adopted more frequently because of the increase in cloud-based HR systems. Dimple Agarwal et al., *People Data: How Far Is Too Far?*, *in* THE RISE OF THE SOCIAL ENTERPRISE: 2018 DELOITTE GLOBAL HUMAN CAPITAL TRENDS 89, 89 (2018).

talent outcomes and profitability.[239] The report also noted that 40% of companies utilize cloud-based HR management systems allowing for easier analytics. The availability of open source AI permits companies to customize their algorithms.[240]

AI can be used to monitor managers as well. AI can search for bias in employee reviews by scanning for certain words to describe female employees that are not used to describe males, such as "aggressive" for female, where the same traits would be described as "leadership-material" for a male.[241] Using AI can directly address the problem with women being judged more harshly than men.[242] In addition to text mining performance reviews, SAP has developed algorithms to help companies review job descriptions for biased language.[243] Overall, responsible AI has been very successful in reducing bias and increasing diversity.

B. *Reducing/Eliminating Noise*

Kahneman's most recent area of research examines how variability in decisions (known as noise) can negatively impact organizations and how AI can reduce and or eliminate this variability. AI can help managers and employers make better decisions by mitigating human biases and decreasing the inevitable variability in decision-making.[244] Not only are humans inconsistent in their own decision-making from day to day, inconsistent decisions also result from two different humans looking at the same data.[245] By contrast, an algorithm will always provide the same

239. *Deloitte's Bersin Finds Effective Use of People Analytics Is Strongly Related to Improved Talent and Business Outcomes*, PR NEWS WIRE (Nov. 13, 2017), https://perma.cc/8X75-JC4Z.

240. *Id*.

241. Courtney Seiter, *Bias at Work: Three Steps to Minimizing Bias in Performance Reviews*, BUFFER OPEN BLOG (Feb. 8, 2018), https://perma.cc/5VWK-8PMB.

242. Moss-Racusin, *supra* note 53, at 16,474-79 (stating that science faculty rated male applicants significantly more competent than female applicants with identical application materials).

243. Tami Reiss, *Big Data to Help Overcome Unconscious Bias—Thanks SAP!*, MEDIUM (May 19, 2016), https://perma.cc/B5JG-59Y3.

244. Kahneman et al., *Noise*: *How to Overcome, supra* note 174. NOISE, a new book coming out in 2020 or 2021 by Daniel Kahneman, Olivier Sibony, and Cass R. Sunstein, will explain how variability in human decisions occurs and offer solutions. I anticipate there will be a great influx of research on noise and algorithms coming out after this book is published.

245. *Id*. at 40.

decision for the same data set.[246] Creating rules that are consistently applied to data sets reduces noise and creates uniformity throughout the organization.[247] The use of AI will also achieve greater accuracy than the use of human decision-makers and avoid decisions which treat similarly situated applicants and employees differently.[248] According to Kahneman, "An algorithm could really do better than humans, because it filters out noise. If you present an algorithm the same problem twice, you'll get the same output. That's just not true of people."[249] Machine decision-making outcomes are not only more consistent, they also tend to be more accurate than those made by humans.[250]

Kahneman's study on insurance underwriters revealed that the organizations involved were unaware of the variability in their risk assessment determinations, which Kahneman found to vary on average by 48% in company A and 60% in company B.[251] Kahneman recommended the use of algorithms to reduce such noise in decision-making.[252] According to Kahneman, "It has long been known that predictions and decisions generated by simple statistical algorithms are often more accurate than those made by experts, even when the experts have access to more

246. Kahneman also provides a method for auditing decisions for noise. *Id.* at 42.

247. Kahneman explains that large data sets are not necessary, the creation of rules based on set criteria is key. *Id*. at 46.

248. Kahneman explains that decisions made by algorithms are "often more accurate than those made by experts, even when the experts have access to more information than the formulas use." *Id*. at 41.

249. MIT Initiative on Digital Econ., *Where Humans Meet Machines: Intuition, Expertise and Learning*, MEDIUM (May 18, 2018), https://perma.cc/66VT-6A84.

250. This is especially true in medical testing. *See* Yun Liu et al., *Artificial Intelligence-Based Breast Cancer Nodal Metastasis Detection: Insights Into the Black Box for Pathologists*, 143 ARCHIVES PATHOLOGY & LABORATORY MED. 859, 861-62 (2018) (noting that Google researchers' AI to detect metastatic breast cancer by evaluating lymph node slides was able to detect 99.3% of the cancers, while human pathologists were only able to detect 81%); Kazimierz O. Wrzeszczynski et al., *Comparing Sequencing Assays and Human-Machine Analyses in Actionable Genomics for Glioblastoma*, 3 NEUROLOGY GENETICS, Aug. 2017, e164 at 1, 6 (discussing IBM's Watson, known as WGA (Watson Genomic Analytics), which was able to analyze a genome of a patient in 10 minutes compared to the human experts who took 160 hours); European Lung Found., *AI Improves Doctors' Ability to Correctly Interpret Tests and Diagnose Lung Disease*, SCIENCEDAILY (Sept. 18, 2018), https://perma.cc/TJJ5-2H3J (comparing AI's 100% compliance rate for pulmonary function tests (PFTs) with human pulmonologists' 74% compliance rate, and noting further that AI provided a correct diagnosis 82% of the time compared with 45% by the pulmonologists).

251. Kahneman et al., *Noise*: *How to Overcome, supra* note 174, at 42.

252. Elisabeth Goodman, *Decision Making: Noise, Intuition and the Value of Feedback*, ELISABETH GOODMAN'S BLOG (Feb. 1, 2017), https://perma.cc/6A8H-NALU.

information than the formulas use. It is less well known that the key advantage of algorithms is that they are noise-free: Unlike humans, a formula will always return the same output for any given input. Superior consistency allows even simple and imperfect algorithms to achieve greater accuracy than human professionals."[253]

One example of an algorithmic fix for reducing noise in hiring is the use of structured interviews. By using chatbots to conduct interviews companies eliminate the variability that a human interviewer or multiple interviewers would bring into the process. Unlike humans, a chatbot asks each interviewee the same set of questions. In addition, by setting criteria for promotions in advance, using an algorithm to assess employees will reduce both the bias of managers and noise in the decision by applying rules uniformly.[254]

VIII. USING AI TO REDUCE ALGORITHMIC BIAS

A number of scholars have written about the dangers and risks of various subcomponents of AI, such as big data, machine learning, data mining and predictive analytics.[255] In their paper, *Big Data's Disparate Impact*, Barocas & Selbst point to the risks of biased data and argue that data mining can unintentionally reproduce historical discrimination resulting from widespread prejudice in society.[256] In their conclusion, they acknowledge the potential benefit to using "data mining to generate new knowledge and improve decision making that serves the interests of both decision makers and protected classes," but caution that such adoption must

---

253. Kahneman et al., *Noise: How to Overcome, supra* note 174, at 41.

254. This is one of the advantages of the use of AI in employment decision-making. Although human can ignore rules, machines cannot. See the study described in Uhlmann, *supra* note 217.

255. *See* ERIC SIEGEL, PREDICTIVE ANALYTICS: THE POWER TO PREDICT WHO WILL CLICK, BUY, LIE, OR DIE (2d ed. 2016) (describing how predictive analytics are currently being used by the government and business to identify preferences and risks and noting that the use of data about groups that have been historically discriminated against can result in discriminatory outcomes); VIKTOR MAYER-SCHÖNBERGER & KENNETH CUKIER, BIG DATA: A REVOLUTION THAT WILL TRANSFORM HOW WE LIVE, WORK, AND THINK (2013) (discussing the potential bias from the likelihood of errors in contained in big data); CATHY O'NEIL, WEAPONS OF MATH DESTRUCTION: HOW BIG DATA INCREASES INEQUALITY AND THREATENS DEMOCRACY (2015) (discussing potential risks of big data).

256. Solon Barocas & Andrew D. Selbst, *Big Data's Disparate Impact*, 104 CALIF. L. REV. 671, 677 (2016).

be done with care.[257] In Kate Crawford's article, the *Hidden Biases of Big Data*, [258] she states: "Hidden biases in both the collection and analysis stages present considerable risks, and are as important to the big-data equation as the numbers themselves."[259] Crawford points to a number of examples where the use of big data resulted in inaccurate predictions. She suggests that unquestioning reliance on these predictions would have resulted in the wrongful allocation of resources.[260] She emphasizes, "In the near term, data scientists should take a page from social scientists, who have a long history of asking where the data they're working with comes from, what methods were used to gather and analyze it, and what cognitive biases they might bring to its interpretation."[261] She correctly suggests that looking to the sources and purposes of data, rather than the resulting numbers, is key to potentially addressing these types of issues. While data mined from the internet or purchased from data brokers will reflect societal prejudices and errors, most of the AI fixes suggested in this paper do not use these types of data sets, but rather rely on data sourced from the company itself or the industry. These scholars, however, have done a great service in bringing these risks to the attention of those working in technology. As indicated in the previous Part, since these articles have been published, not only have recent developments in AI been shown to be very successful in reducing bias and noise in human decision-making, new tools are now being developed to both detect and remedy the potential for human biases to creep into data used in machine decision-making. Although most of the suggested AI techniques described above do not create significant risks of discrimination, the following Subparts will highlight some recent advances in AI specifically designed to mitigate the potential for discriminatory results.

A. "Garbage In, Garbage Out"

The primary risk of incorporating AI in employment decision-making described by almost every scholar in this field is the potential for discriminatory outcomes. As explained above, data mined from the internet, social media and data brokers is likely to be error-prone and reflect societal

257. *Id.* at 732.
258. Kate Crawford, *The Hidden Biases in Big Data*, HARV. BUS. REV. (Apr. 1, 2013), https://perma.cc/JC8Q-YT8K.
259. *Id.*
260. *Id.*
261. *Id.*

prejudices.[262] Gender and racial stereotyping in society have created an internet which reflects these prejudices resulting in, for example, some disturbing results based on the algorithms that run Google searches.[263] Data from these sources should not be used to identify traits a company should look for when hiring a new employee when the harm can be the exclusion of an entire gender, race, or community.[264] This potential problem is known as "garbage in, garbage out."[265]

An additional type of risk results from using data sets skewed in favor of a gender or race.[266] For example, if you run an algorithm on an organization's data set seeking to identify common traits in the top performers, and 80% of those top performers are male, the results will also be skewed in favor of the male gender. This appears to be the reason why Amazon had to scrap its in-house machine-learning algorithm to sort through resumes to hire the best candidates.[267] A group of programmers in

---

262. *See, e.g.*, Barocas & Selbst, *supra* note 256 at 674 (stating that "an algorithm is only as good as the data it works with" and can "reflect historic patterns of prejudice"); Pauline T. Kim, *Data-Driven Discrimination at Work*, 58 WM. & MARY L. REV. 857, 921-22 (2017) [hereinafter Kim, *Data-Driven Discrimination*] (describing that algorithms "built on inaccurate, biased, or unrepresentative data can produce outcomes biased along lines of race, sex, or other protected characteristics").

263.

> In one study, Harvard professor Latanya Sweeney looked at the Google AdSense ads that came up during searches of names associated with white babies (Geoffrey, Jill, Emma) and names associated with black babies (DeShawn, Darnell, Jermaine). She found that ads containing the word "arrest" were shown next to more than 80% of "black" name searches but fewer than 30% of "white" name searches. Sweeney worries that the ways Google's advertising technology perpetuates racial bias could undermine a black person's chances in a competition, whether it's for an award, a date, or a job.

Nanette Byrnes, *Artificial Intolerance*, MIT TECH. REV. (Mar. 28, 2016), https://perma.cc/HXR7-QTH8 (citing Latanya Sweeney, *Discrimination in Online Ad Delivery*, ACM QUEUE, Mar. 2013, at 1, 10).

264. Although there have been developments in correcting bias on internet-generated data sets (of purchased data sets sourced from the internet), there is a long way to go. One suggestion to bring attention to potential flaws in the data set is the use of data sheets. Timnit Gebru et al., Datasheets for Datasets (Apr. 16, 2019) (working paper), https://perma.cc/H75E-U3EM. For a sample data sheet, see *Datasheet for RecipeQA*, RECIPEQA, https://perma.cc/HVF3-VK3G (archived May 16, 2019).

265. This term is used, for example, by Pauline Kim in *Big Data and Artificial Intelligence: New Challenges for Workplace Equality*, 57 U. LOUISVILLE L. REV. 313 (2019); David Lehr & Paul Ohm, *Playing with the Data: What Legal Scholars Should Learn About Machine Learning*, 51 U.C. DAVIS L. REV. 653, 656 (2017); Sullivan, *supra* note 13 at 2; Barocas & Selbst, *supra* note 256 at 684-87.

266. Barocas & Selbst, *supra* note 256 at 684-687.

267. Jeffrey Dastin, *Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women*, REUTERS (Oct. 9, 2018), https://perma.cc/5UPB-NHLE.

their Edinburgh office developed an algorithm to identify candidates by teaching the program to recognize some 50,000 terms which had shown up on past resumes.[268] The data set used to identify those terms consisted of resumes submitted to it since 2014, the vast majority of which came from male applicants.[269] Because the algorithm had been fed data from primarily men it awarded higher scores to the resumes submitted by male applicants. Even when the sex of applicants was not used as a factor, it still downgraded resumes that includes the term "women's" such as women's tennis or women's chess.[270] The algorithm also taught itself to look for verbs more commonly associated with male engineers, such as "executed."[271] Not only did the program weed out female applicants, it also failed to meet its goal of designating the best candidates.[272] The results appeared to be almost random, demonstrating that the program did not even work. It was not just a discriminatory algorithm; it was a deficient one as it was unable to produce the desired result. Although some have pointed to this example as showing that AI is biased, it actually just reflects that it was a poor design which failure was compounded by using biased unbalanced data which unsurprisingly created biased results.[273] This exact type of danger can be remedied by addressing these risks in the design of the algorithm and by balancing the data and/or increasing the diversity of existing data points as discussed in the next paragraph.

Organizations are actively working on methods for creating better data sets for use in AI. IBM is one company that has published its work on creating balanced data sets. "AI holds significant power to improve the way we live and work, but only if AI systems are developed and trained responsibly, and produce outcomes we trust. Making sure that the system is trained on balanced data, and rid of biases is critical to achieving such trust" write IBM fellows Aleksandra Mojsilovic and John R. Smith.[274] In an article

---

268. *Id.*

269. *Id*.

270. *Id*.

271. *Id*.

272. *Id*.

273. Although the study does not disclose the gender of the twelve engineers who developed the program, it is likely that they were male as Amazon's software engineers are "overwhelmingly male." Rachel Goodman, *Why Amazon's Automated Hiring Tool Discriminated Against Women*, ACLU (Oct. 12, 2018), https://perma.cc/FNL8-UNDD. The need for a diverse set of programmers is discussed *infra*.

274. Aleksandra Mojsilovic & John Smith, *IBM to Release World's Largest Annotation Dataset for Studying Bias in Facial Analysis*, IBM RES. BLOG (June 27, 2018),

by IBM developer, Kenneth Jensen, he explains how balancing works.[275] To create a balanced data set, developers duplicate the results from the less frequent category, called boosting. Developers also discard the results of the more frequent category, called reduction. They can also combine boosting and reduction to obtain more balanced results.[276] The goal is to reduce the impact of a skewed data set.[277] Thus, if you are using a data set to look for the traits of successful programmers where 80% of your programmers are male, you would balance the data set to more evenly reflect both genders to improve your outcome.[278]

When your data sets contain little or no information about certain groups of people, your algorithm will not accurately evaluate people who belong to that group. In a paper presented at the Conference on Fairness, Accountability, and Transparency, the authors were able to create a test to measure the accuracy of commercial classification algorithms.[279] The test was able to demonstrate how unbalanced data inputs (77% male and 83% white) resulted in facial recognition software with a 1% error rate for identifying light-skinned men, but a 20-34% error rate for darker-skinned women.[280] According to Joy Buolamwini, author of the paper and founder of Algorithmic Justice League:

---

https://perma.cc/CE9N-K36C.

275. Kenneth Jensen, *Use Balancing to Produce More Relevant Models and Data Results*, IBM DEVELOPER (Sept. 19, 2016), https://perma.cc/38B8-5T58. The author points out that by using test data sets, you can determine whether business objectives are being met and correct for unbalanced data (80% males in a sample, for example). *Id.* While this article is specific to the use of the IBM SPSS Modeler, the methods used can be applied to other analytics programs.

276. *How to Handle Imbalanced Classification Problems in Machine Learning?*, ANALYTICS VIDHYA (Mar. 17, 2017), https://perma.cc/6YSZ-D36Y.

277. For example, MIT's facial recognition AI, created by primarily white male students, was unable to recognize black female faces because of the lack of diversity in the data set. In order to remedy this effect, IBM agreed to release an annotated data set of facial images that was balanced in terms of skin tone, gender and age. Malik Murison, *IBM Takes Steps to Tackle AI Bias*, INTERNET BUS. (June 29, 2018), https://perma.cc/MX4D-CHT2.

278. *Id*.

279. Joy Buolamwini & Timnit Gebru, *Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification*, 81 PROC. MACHINE LEARNING RES. 77 (2018).

280. One of the authors, Joy Buolamwini, decided to investigate this issue after noticing that an art installation using facial recognition did not work as well for darker skinned people. She was able to show the bias in the data by feeding 1,200 images of darker-skinned people and women into the programs to discover the actual bias. She suggests that white male programmers did not notice this flaw because the program worked quite well on white male faces. Although IBM says it is not a direct result of Buolamwini's research, it has employed a more balanced model with a half a million

> If the training sets aren't really that diverse, any face that deviates too much from the established norm will be harder to detect . . . Training sets don't just materialize out of nowhere. We actually can create them. So there's an opportunity to create full-spectrum training sets that reflect a richer portrait of humanity.[281]

By being aware of the risks involved in using data sets which do not represent society at large, and repeatedly testing for any potential bias, biased results can be mitigated.

Another potential solution to reduce the potential harm of biased data sets is to drastically increase the diversity of data points reviewed. This solution has been incorporated into credit checks by reviewing more data points than those traditionally measured.[282] For example, to help those with little or no credit history, ZestFinance created an algorithm that considers tens of thousands of pieces of information beyond what is used for a typical credit score (which uses a limited number of data points).[283] The same could be done with employment data. In addition, data analytics can be used to search for discrimination in the outcomes themselves. Companies, such as ZestFinance, frequently test the results that their automated processes produce to discover any discriminatory results that could harm applicants.[284]

---

images of a wider range of facial images. However, Ruchir Puri, chief architect of IBM's Watson artificial-intelligence system, acknowledges that her research bought up some very significant points. Larry Hardesty, *Study Finds Gender and Skin-Type Bias in Commercial Artificial Intelligence Systems*, MIT NEWS (Feb. 11, 2018), https://perma.cc/5EK5-C439.

281. Jason Bloomberg, *Bias Is AI's Achilles Heel. Here's How to Fix It*, FORBES (Aug. 13, 2018), https://perma.cc/9SM2-L6DT.

282. *ZestFinance Introduces Machine Learning Platform to Underwrite Millennials and Other Consumers with Limited Credit History*, BUSINESSWIRE (Feb. 14, 2017) [hereinafter *ZestFinance Introduces*], https://perma.cc/H3GM-7NEE.

> Affirm and ZestFinance are both founded on the idea that by looking at tens of thousands of data points, machine-learning programs can expand the number of people deemed creditworthy. In other words, algorithms fed by so much diverse data will be less prone to discrimination than traditional human-driven lending based on a much more limited number of factors. Among the insights discovered by ZestFinance's algorithms: that income is not as good a predictor of creditworthiness as the combination of income, spending, and the cost of living in any given city.

Nanette Byrnes, *Artificial Intolerance*, MIT TECH. REV. (Mar. 28, 2016), https://perma.cc/3LYT-2F4R.

283. *ZestFinance Introduces*, *supra* note 282.

284. *Id.*

Accenture is one of the organizations that has created AI that can test data sets for bias.[285] Known as the "Fairness Tool," it can examine the data set for sensitive variables, identify and remove any coordinated influence that would result in an unfair outcome, evaluate false positives and negatives, and display the impact that the fixes have on the model's accuracy.[286] MIT's Computer Science and Artificial Intelligence Laboratory has also created a method for detecting and mitigating bias resulting from under-represented segments of society in training data for machine learning.[287]

Another issue mentioned by critics of the use of AI in employment decisions is that biases in the programmers themselves may produce biased results.[288] In addition to failing to notice that the facial recognition software did not work well on black female faces because it was almost 99% accurate on white male faces, programmers may, due to their unconscious biases, measure the accuracy of an algorithms without considering races and genders other than their own.[289] For example, programmers might chose inappropriate "target variables" or "class labels."[290] Because more men tend

---

285. Rumman Chowdhury, *Tackling the Challenge of Ethics in AI*, ACCENTURE BLOG (June 6, 2018), https://perma.cc/MBJ6-XJDP.

286. *Id*.

287. Alexander Amini et al., *Uncovering and Mitigating Algorithmic Bias Through Learned Latent Structure*, PROC. 2019 AAAI/ACM CONF. ON ARTIFICIAL INTELLIGENCE, ETHICS, AND SOC'Y 289 (Jan. 27-28, 2019), https://perma.cc/G7DE-6VMV.

288. Jack Clark, *Artificial Intelligence Has a 'Sea of Dudes' Problem*, BLOOMBERG (June 23, 2016), https://perma.cc/U4QA-4BTD.

289. There are numerous examples where male software developers have made assumptions resulting in the exclusion of other groups. For example, there are games that do not provide female avatars. In almost all games, a male character is the default and in other games although the male character is free, users have to pay for a female avatar. Madeline Messer, *I'm a 12-Year-Old Girl. Why Don't the Characters in My Apps Look Like Me?*, WASH. POST (Mar. 4, 2015), https://perma.cc/Z39Q-5LEV. *See also* SARA WACHTER-BOETTCHER, TECHNICALLY WRONG: SEXIST APPS, BIASED ALGORITHMS, AND OTHER THREATS OF TOXIC TECH (2017) (discussing how the only female programmer in an app development meeting had her ideas dismissed because the males were relying on their impressions of their stay at home wives' shopping activities, stating, "Oh, 51% of the women can't be tech-savvy" and insisting that the women cared only about shopping and other leisure related activities and eventually launching a product that failed.) In another example, Slack's use of a brown hand for its "add to Slack" function would most likely have not occurred to a white male programmer. Slack employs the "highest percentage of female and black engineers of any tech company." Tobias Hardy, *How Slack Is Doing Diversity and Inclusion Right*, LAUNCHPAD, https://perma.cc/FUU4-NEHG (archived May 20, 2019).

290. Target variables are what the machine is looking for, and class labels are how the data is classified. Because defining these target variables and class labels are sometimes subject to the discretion of the programmers, the potential for unintentional

to be programmers, their own biases could affect the process.[291] Because researchers have become more aware of these less obvious types of risks, new processes are being developed to remedy these situations. Prejudicial classifications, data errors, and incorrect or missing variables, for example, can be audited for and eliminated in a number of ways.[292] Another solutions is to make the programmers' thinking process more transparent by requiring them to document what went into their algorithm prior to creating it.[293] Bruneis & Goodman suggest eight criteria that developers would need to identify for review by others.[294] By considering the implications of classifications prior to the creation of a program and auditing the outcomes, bias can be detected and mitigated.[295] The best solution to eliminate bias in programmers, however, is to hire a diverse group to create the programs meant to provide fair and consistent employment decisions. As Fei-Fei Li, Chief Scientist of Artificial Intelligence and Machine Learning at Google and Co-Director of the Stanford Institute for Human-Centered Artificial Intelligence said:

---

discrimination against "systemically disadvantaged protected classes" may occur. Barocas & Selbst, *supra* note 256 at 677-80.

291. Clark, *supra* note 288. At Google, for example, only 10% of the employees working on machine intelligence are women. Tom Simonite, *AI Is the Future—But Where Are the Women?*, WIRED (Aug. 17, 2018), https://perma.cc/VCB5-AS9N.

292. *See* Andrea Romei & Salvatore Ruggieri, *Discrimination Data Analysis: A Multi-Disciplinary Bibliography*, *in* DISCRIMINATION AND PRIVACY IN THE INFORMATION SOCIETY: DATA MINING AND PROFILING IN LARGE DATABASES 109, 122 (Bart Custers et al. eds., 2013) (acknowledging the growth in discrimination discovery and prevention in data analysis, the chapter provides an "annotated bibliography of the literature on discrimination data analysis").

293. Robert Brauneis & Ellen P. Goodman, *Algorithmic Transparency for the Smart City*, 20 YALE J.L. & TECH. 103 (2018).

294. These include: the predictive goals of the algorithm and the problem it is meant to solve, the training data considered relevant to reach the predictive goals, the training data excluded and the reasons for excluding it, the actual predictions of the algorithm as opposed to its predictive goals, the analytical techniques used to discover patterns in the data, other policy choices encoded in the algorithm besides data exclusion, validation studies or audits of the algorithm after implementation, and a plain language explanation of how the algorithm makes predictions. Henrik Chulu, *Let Us End Algorithmic Discrimination*, MEDIUM (Aug. 3, 3018), https://perma.cc/TPN4-WPZZ (summarizing Brauneis & Goodman, *supra* note 293).

295. *See, e.g.*, Marco Tulio Ribeiro et al., *"Why Should I Trust You?": Explaining the Predictions of Any Classifier*, PROC. 22ND ACM SIGKDD INT'L CONF. ON KNOWLEDGE DISCOVERY & DATA MINING 1135, (Aug. 9, 2016), https://perma.cc/AXS3-FM7M (explaining a software system that helps make sense of algorithmic decisions called LIME); Edward Ma, *Anchor Your Model Interpretations by Anchors*, MEDIUM (Sept. 8, 2018), https://perma.cc/7X49-ADDY (expanding on LIME with the use of anchors to explain algorithmic decisions).

> If we don't get women and people of color at the table—real technologists doing the real work—we will bias systems . . . This is the time to get women and diverse voices in so that we build it properly, right? And it can be great. It's going to be ubiquitous. It's going to be awesome. But we have to have people at the table.[296]

Not only will hiring more women and URMs as programmers reduce the potential for bias, it will have a significant effect by addressing concerns ***prior*** to the creation of an algorithm.

B. "Black Box"

Another criticism of the use of AI in employment decisions is known as the "black box" problem, which results from the difficulty in pin-pointing why a machine produced a particular outcome.[297] The concern raised is that if AI outcomes cannot be explained, they may contain unknown biases.[298] While some researchers are looking for ways to provide an understanding of algorithmic outcomes without opening the black box,[299] others have proposed various ways to audit for fairness and debias algorithms.[300] One

---

296. Kate Brodock, *Why We Desperately Need Women to Design AI*, MEDIUM (Aug. 4, 2017), https://perma.cc/3S43-RHUX.

297. *See, e.g.*, FRANK PASQUALE, THE BLACK BOX SOCIETY 3 (2015) (describing how a "black box" can mean "a system whose workings are mysterious; we can observe its inputs and outputs, but we cannot tell how one becomes the other."); Danielle Keats Citron & Frank Pasquale, *The Scored Society: Due Process for Automated Predictions*, 89 WASH. L. REV. 1, 6 (2014) (discussing that the "black box" problem with AI is that the algorithm may convert "inputs to outputs without revealing how it does so."); Kim, *Data-Driven Discrimination*, *supra* note 262 at 921-22 ("An algorithm may be a 'black box' that sorts applicants or employees and predicts who is most promising, without specifying what characteristics or qualities it is looking for.").

298. Citron & Pasquale, *supra* note 297.

299. *See* Riccardo Guidotti et al., *A Survey of Methods for Explaining Black Box Models*, 51 ACM COMPUTING SURVEYS 1 (2018) (providing a survey of literature on addressing the black box problem); Andrew D. Selbst & Solon Barocas, *The Intuitive Appeal of Explainable Machines*, 87 FORDHAM L. REV. 1085, 1129 (requiring developers to document the reasoning behind the choices made in developing a machine learning model); Sandra Wachter et al., *Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR*, 32 HARV. J.L. & TECH. (2018) (explaining why a particular decision was made so that the subject can make a change to obtain a different result in the future is more valuable than opening the "black box").

300. Technological ways to debias algorithms can be accomplished pre-processing, in-processing, and post-processing. For a collection of research on these methods, see Abhishek Tiwari, *Bias and Fairness in Machine Learning* (July 4, 2017) (unpublished manuscript), https://perma.cc/U3XE-XPGZ. *See also* Joshua A. Kroll et al., *Accountable Algorithms*, 165 U. PA. L. REV. 633, 637 (2017) (describing "how technical tools for

example of a tool developed to address this challenge and provide greater transparency into the algorithmic process is known as Quantitative Input Influence.[301] This method can help identify the reason for an algorithmic outcome by measuring and displaying the influence of inputs on outputs. In other words, the more prominent an input, the greater impact it had on the algorithmic decision. Methods such as these can provide an understanding of why an outcome was produced without having to peer into the "black box."[302]

In addition, AI can be used to prevent and detect bias in the algorithmic outcomes.[303] The following are examples of some of the methods that have been developed to deal with potential discriminatory results in machine learning.[304] At the 2016 Neural Information Processing Systems (NIPS) conference, researchers demonstrated their method known as hard de-biasing for reviewing and removing gendered stereotypes resulting from biased training data.[305] At the 2018 International Conference in Machine Learning counterfactual fairness testing was also shown to be effective in rooting out bias.[306] With counterfactual testing, instead of ignoring

verifying the correctness of computer systems can be used to ensure that appropriate evidence exists for later oversight"); Pauline Kim, *Auditing Algorithms for Discrimination*, 166 U. PA. L. REV. ONLINE 189, 190-91 (2017) [hereinafter Kim, *Auditing Algorithms*] (describing how auditing the outcomes of "decisionmaking [sic] algorithms similarly offers a method of detecting when they may be biased against particular groups").

301. Anupam Datta et al., *Algorithmic Transparency via Quantitative Input Influence*, 2016 IEEE SYMP. ON SECURITY & PRIVACY, 598 (2016).

302. Lehr & Ohm, *supra* note 265, at 710 (cautioning that such influence tests are not available for every type of machine learning algorithm). For a survey of literature reviewing explainability and transparency models, see Amina Adadi & Mohammed Berrada, *Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)*, 6 IEEE ACCESS 52138 (2018) (summarizing literature on the growing field of explainable AI to mitigate the black box problem and future research trajectories).

303. *See* Rumman Chowdhury & Narendra Mulani, *Auditing Algorithms for Bias*, HARV. BUS. REV. (Oct. 24, 2018), https://perma.cc/YX2H-7MWT (discussing a fairness tool to audit outcomes developed by Accenture).

304. For a full and very helpful description of the machine learning process, see Lehr & Ohm, *supra* note 265, at 670-702.

305. Tolga Bolukbasi et al., *Man Is to Computer Programmer as Woman Is to Homemaker?: Debiasing Word Embeddings*, PROC. 30TH INT'L CONF. ON NEURAL INFO. PROCESSING SYS. (Dec. 5-10, 2016), https://perma.cc/9A8W-JCD4. In this experiment, gender stereotyping was learned when the algorithm was trained on Google News as a data set. The programmers were able to "fix" the bias by removing the gendered stereotypes such as associating receptionist with female but leaving the association between queen and female essentially debiasing the algorithm. *Id.*

306. The authors of a paper started with the premise that a decision is fair towards an individual if it is the same in (a) the actual world and (b) a counterfactual world where the individual belonged to a different demographic group. Matt J. Kusner et al.,

protected attributes, social biases are taken into consideration.[307] Essentially, two models are created: one including potentially biased factors (such as gender and race) and one without. An outcome is considered fair if it is the same in both models. In other words, the outcome would be the same whether the individual was male or female or one race or another.[308] While this study examined fairness in law school admissions, it could equally be applied in employment decisions. Another way of discovering and mitigating the effects of biased data is by including additional data. This is known as the text classifier model, where wrongful negative associations are supplemented with nontoxic labels.[309] In this study, a text classifier was built to identify toxic comments in Wikipedia Talk Pages but the program would also flag nontoxic references, such as to the word "gay." In order to eliminate false positives, a number of non-toxic data points were entered the training data set such as "I am gay" and "I am a gay person" to counteract the many instances of toxic uses of the word "gay" in the data set and preventing the flagging of these neutral statements as toxic language. The study demonstrates the ability of tools to mitigate bias without reducing the accuracy of the model's results. This type of program could be used to eliminate inappropriate wording in job ads and employee reviews.[310] Another method, adversarial debiasing, removes undesired biases in training data while meeting a fairness test.[311] In other words, an algorithm

---

Counterfactual Fairness, PROC. 31ST INT'L CONF. ON NEURAL INFO. PROCESSING SYS. (Dec. 4-9 2017), https://perma.cc/HR4Z-NC3C.

307. *Id.*

308. *Id.*

309. Lucas Dixon et al., *Measuring and Mitigating Unintended Bias in Text Classification*, PROC. AAAI/ACM CONF. ON AI, ETHICS, AND SOC'Y 67 (Feb. 2-3, 2018), https://perma.cc/8FWL-7RRU.

310. Recently, Cygnet System came under fire for putting out a job ad for a tech position indicting that the desired candidate would be "preferably Caucasian." Gabrielle Sorto, *A Company Posted a Job Ad Seeking 'Preferably Caucasian' Candidates*, CNN (Apr. 30, 2019), https://perma.cc/NC3P-DFYE. By using algorithms to both examine wording and correct inappropriate wording, candidates and employees can be treated more fairly.

311. Brain Hu Zhang et al., *Mitigating Unwanted Biases with Adversarial Learning*, PROC. AAAI/ACM CONF. ON AI, ETHICS, AND SOC'Y 335 (Feb. 2-3, 2018), https://perma.cc/4ETA-4CPN. In a greatly oversimplified explanation, using census data (X), to predict income bracket (Y), the authors wanted to maximize the algorithm's ability to predict (Y) while minimizing the model's ability to also predict a protected attribute such as gender (Z). Because X contained unwanted biases reflecting Y, the researchers sought to remove the generalizations about the protected attribute from the data. Because they were able to accurately predict Y without predicting Z, they met the accepted measurements of fairness. They were successful in training a model to be

could be designed to predict an output variable based on an input variable without predicting a protected variable (such as gender or race). The ability to predict the output variable without predicting the protected variable would signal that the algorithm produces a fair result. This type of algorithm could be used to test that the AI being used to predict required skills for a particular position is fair for all genders and races.

Due to the attention that scholars have brought to this issue, many companies are now working on these types of solutions.[312] IBM's AI Fairness 360, for example, is an open source library of tools available online to people and entities who want to identify and mitigate bias in their own machine learning programs.[313] The tools include seventy fairness metrics and ten "state-of-the-art" bias mitigation algorithms.[314] A number of other organizations are actively working on methods to detect and mitigate potential discriminatory results in machine decision-making, including

---

"demonstrably less biased" while still performing "extremely well" on predicting Y. *Id.*

312. Kalev Leetaru, *Can Algorithms Save Us From Bias?,* FORBES (Jul. 31, 2016), https://perma.cc/5M7C-YFMG.

313. Jesus Rodriguez, *What's New in Deep Learning Research: Reducing Bias and Discrimination in Machine Learning Models with AI Fairness 360*, TOWARDS DATA SCI. (Sept. 24, 2018), https://perma.cc/9SF7-3B9V. *See also* Cynthia Dwork & Deirdre K. Mulligan, *It's Not Privacy, and It's Not Fair*, 66 STAN. L. REV. ONLINE 35 (2013) (describing tools to test for classification bias).

314. *AI Fairness 360 Open Source Toolkit*, IBM, https://perma.cc/59YJ-3ZEV (archived May 20, 2019). Represented research includes work by Rachel K. E. Bellamy. Rachel K. E. Bellamy et al., *AI Fairness 360: An Extensible Toolkit for Detecting, Understanding, and Mitigating Unwanted Algorithmic Bias* (2018); Flavio P. Calmon et al., *Optimizing Data Pre-Processing for Discrimination Prevention*, PROC. 31ST CONF. ON NEURAL INFO. PROCESSING SYS. (Dec. 4-9, 2017), https://perma.cc/87ET-8XXQ; Michael Feldman et al., *Certifying and Removing Disparate Impact*, PROC. 21ST ACM SIGKDD INT'L CONF. ON KNOWLEDGE DISCOVERY & DATA MINING (Aug. 10-13, 2015), https://perma.cc/3HC4-JJ3W; Moritz Hardt et al., *Equality of Opportunity in Supervised Learning*, at 3315-28, PROC. 30TH INT'L CONF. ON NEURAL INFO. PROCESSING SYS. (Dec. 5-10, 2016), https://perma.cc/W2AQ-XWCN; Faisal Kamiran & Toon Calders, *Data Preprocessing Techniques for Classification Without Discrimination*, 33 KNOWLEDGE INFO. SYS. 1 (2011); Faisal Kamiran et al., *Decision Theory for Discrimination-Aware Classification*, PROC. 2012 IEEE 12TH INT'L CONF. ON DATA MINING (Dec. 10, 2012), https://perma.cc/38Q4-YKNM; Toshihiro Kamishima et al., *Fairness-Aware Classifier with Prejudice Remover Regularizer*, at 35, JOINT EUR. CONF. ON MACHINE LEARNING & KNOWLEDGE DISCOVERY IN DATABASES (Sept. 23-27, 2012), https://perma.cc/6Q45-PFM3; Geoff Pleiss et al., *On Fairness and Calibration*, PROC. 31TH INT'L CONF. ON NEURAL INFO. PROCESSING SYS. (Dec. 4-9, 2017), https://perma.cc/2ES9-ZNRJ; Rich Zemel et al., *Learning Fair Representations*, at 28, PROC. 30TH INT'L CONF. ON MACHINE LEARNING (June 16-21, 2013), https://perma.cc/KE4D-XHMH; Hu Zhang, *supra note* 311. *See also* Emily James, *Practical Steps to Addressing Algorithmic Bias*, FLEXMR BLOG, https://perma.cc/K8U2-R3CA (archived May 20, 2019) (explaining algorithmic bias and potential tools for its mitigation). See also research on debiasing word embedding in Bolukbasi, *supra* note 305.

Facebook's Fairness Flow,[315] Pymetrics' open-source Audit AI Tool,[316] Accenture's Toolkit, and Google's What-if Tool.[317] Overall, a responsible AI program to reduce bias in employment decisions will start with the careful consideration of the design of the algorithms, the ongoing monitoring and correcting of data, and the auditing of outcomes for potential discriminatory results.

The focus on whether AI is biased is somewhat misplaced. The more important question is: Does the use of AI result in less biased, more consistent, and more accurate decisions than flawed human decision-making?[318] No amount of mentoring in the tech industry is going to fix its diversity problem. No amount of training is going to fix flawed human decision-making. By focusing on the potential harms of AI, we miss opportunities to make workplace safer and more equitable for all employees.[319] According to Tim O'Reilly, who is an uncanny predictor of trends in technology, "AI is not the machine from the future that is hostile to

315. Dave Gershgorn, *Facebook Says It Has a Tool to Detect Bias in Its Artificial Intelligence*, QUARTZ (May 3, 2018), https://perma.cc/4Q7P-MYUW.

316. Khari Johnson, *Pymetrics Open-Sources Audit AI, an Algorithm Bias Detection Tool*, VENTUREBEAT (May 31, 2018), https://perma.cc/SL9U-WUP6.

317. *The What-If Tool: Code-Free Probing of Machine Learning Models*, GOOGLE AI BLOG (Sept. 11, 2018), https://perma.cc/9H7J-W3CJ. For more recent developments that may have been created since the publication of this article, see arXiv.com, the largest open source database of scientific papers.

318. Algorithms beat individuals about half the time. And they match individuals about half time," Kahneman said. "There are very few examples of people outperforming algorithms in making predictive judgments. So when there's the possibility of using an algorithm, people should use it. We have the idea that it is very complicated to design an algorithm. An algorithm is a rule. You can just construct rules." Paul McCaffrey, *Daniel Kahneman: Four Keys to Better Decision Making*, ENTERPRISING INVESTOR (June 8, 2018), https://perma.cc/YYD7-ZUB3.

319. While this paper focuses on gender, further research must be done on gender in combination with racial bias. Sheryl Sandberg notes that "More companies prioritize gender diversity than racial diversity, perhaps hoping that focusing on gender alone will be sufficient to support all women . . . But women of color face bias both for being women and for being people of color, and this double discrimination leads to a complex set of constraints and barriers." Courtney Connley, *Why the Gender Pay Gap Still Exists 55 Years After the Equal Pay Act Was Signed*, CNBC (June 10, 2018), https://perma.cc/WBH7-8EU9. For the types of issues that URMs face in the tech industry, see Erica Joy, *The Other Side of Diversity*, MEDIUM (Nov. 4, 2014), https://perma.cc/9HBB-4LP5 (describing the need for more research into "the psychological effects of being a minority in a mostly homogeneous workplace for an extended period of time"); Aston Motes, *Why Aston Motes, Dropbox's First Employee, Chose MIT Over Caltech*, FAST COMPANY (Nov. 14, 2014), https://perma.cc/5WRY-MUAY (explaining why Silicon Valley lacks diversity); Salvador Rodriguez, *Why Silicon Valley Is Failing Miserably At Diversity, And What Should Be Done About It*, I.B. TIMES (July 7, 2015), https://perma.cc/BRR8-MT3W (describing one woman's experience about the lack of diversity in Silicon Valley).

human values and will put us all out of work. AI is the next step in the spread and usefulness of knowledge, which is the true source of the wealth of nations. We should not fear it. We should put it to work, intentionally and thoughtfully, in ways that create more value for society than they disrupt. It is already being used to enhance, not replace, human intelligence."[320]

IX. LEGAL CONCERNS IN INCORPORATING AI INTO YOUR D&I PLAN

It should be noted that many scholars have already addressed Title VII in connection with using AI in employment decisions and concluded that liability under current law is not likely.[321] Disparate treatment requires "intent," and a plaintiff likely cannot show that a machine intended to discriminate. The consensus is that any claim of algorithmic discrimination would fall under disparate impact theory due to a facially neutral practice (using AI) but would most likely be excused as job related and consistent with business necessity.

With a claim of disparate impact, courts would almost certainly use the test set forth in *Griggs,* which requires that there be a disproportionately negative effect on a statutorily protected group.[322] If the screening or testing results in a more diverse workplace, there is no discrimination. If more women and URMs are disproportionately screened out, the algorithm could

320. Tim O'Reilly, *What Will Our Lives Be Like as Cyborgs?,* THE ATLANTIC (Oct. 27, 2017), https://perma.cc/8Q46-YTGL. In addition to technical solutions to mitigating AI risks, more attention is now being paid to how human-centric AI can be developed. The European Commission is in the process of creating ethical rules for the development and use of AI. The goal is to create AI which is trustworthy and that supports human rights rather than harms them. This human-centric approach recognizes the potential of AI to improve society. The two main requirements are that AI serve both an "ethical purpose" and be "technically robust" meaning that it should do what it purports to do. It is very important for companies developing AI solutions to reduce bias and noise while increasing diversity to keep these considerations in mind. EUROPEAN COMM'N, ETHICS GUIDELINES FOR TRUSTWORTHY AI (2019).

321. *See* Barocas & Selbst, *supra* note 256 (discussing how disparate treatment as a cause of action would not work with an allegation of discriminatory AI due to the requirement of "intent" and that a claim of disparate impact would likely be defeated under the business necessity test); Kim, *Data-Driven Discrimination*, *supra* note 262, at 910-11 (arguing for a new theory of liability in response to classification bias); Lehr & Ohm, *supra* note 265, at 666 (explaining Barocas & Selbt's argument that Title VII is not sufficient to address algorithmic discrimination); Sullivan, *supra* note 13 (concluding current disparate treatment law would not fit with AI and disparate treatment claims could be overcome with a showing of business necessity). *Cf*. Stephanie Bornstein, *Antidiscriminatory Algorithms*, 70 ALA. L. REV. 519 (2018) (arguing that stereotyping theory, a form of disparate treatment could result in a finding of discrimination by AI).

322. *Griggs*, 401 U.S. at 431.

be reviewed, and any bias detected and mitigated. The company must be able to show that the algorithms used in the AI are valid for use and accurately measure what they purport to measure.[323] Another concern raised is that setting a goal of increasing diversity in your algorithm could result in a reverse discrimination suit. This argument is not likely to succeed, due to the valuable and quantifiable business reasons to employ a diverse workforce.[324] AI helps organizations hire the best candidates. Provided the organization uses AI in a responsible and intentional fashion, they would meet this standard.[325]

An additional concern that has been raised is that the auditing for bias and/or correction of biased outcomes from a machine decision would violate the holding in *Ricci v. DeStefano*.[326] In *Ricci* the court held that the City's decision to throw out the results of a test used to determine promotions because of a fear of a disparate impact suit (no URMs scored high enough to be promoted and only white men would have been promoted) was a violation of Title VII and constituted disparate treatment.[327] The analogy is that correcting a biased outcome by fixing the algorithm is similar to throwing out a test because of biased results.[328] However, Pauline Kim suggests that any such auditing would be to prospectively revise the algorithm to reduce bias thus there would be no adverse employment action.[329] In addition, she argues that because Title VII encourages "voluntary efforts to comply with nondiscrimination goals," auditing algorithms for bias would not run afoul of discrimination law.[330] Because the use of AI would be implemented to increase diversity and

323. Matthew T. Bodie et al., *The Law and Policy of People Analytics*, 88 U. COLO. L. REV. 961, 1023 (2017).

324. Courts have found affirmative action plans as a legitimate nondiscriminatory reason for an employment action permissible under Title VII countering a reverse discrimination claim. Johnson v. Transp. Agency, 480 U.S. 616 (1987); United Steelworkers of Am. v. Weber, 443 U.S. 193 (1979), as cited in Bodie et al., *supra* note 323 at 60.

325. See Part VI for discussion on how AI can identify the best candidates.

326. 557 U.S. 557 (2009).

327. *Id*.

328. Kroll, *supra* note 300, at 694-95 (arguing that any such auditing would violate the holding in Ricci).

329. Kim, *Auditing Algorithms, supra* note 300 at 197-202. *See also* Mark MacCarthy, *Standards of Fairness for Disparate Impact Assessment of Big Data Algorithms,* 48 CUMBERLAND L. REV. 102, 125-29 (2017) (explaining how recent case law would not prevent the "development, modification, or use of algorithms that have a lesser disparate impact through aiming toward statistical parity or equal group error rates").

330. Kim, *Auditing Algorithms*, *supra* note 300, at 191.

create more uniform employment decisions, rather than correct a specific problem, it would not be analogous to throwing out the results of a test (as in *Ricci*) to avoid a disparate impact suit.[331]

Another benefit to using AI over subjective human decision-making is that noise can be reduced or eliminated. This will help organizations avoid liability for inconsistent employment decisions.[332] Claims of discrimination based on noise occurs when multiple reasons are given for an employee's termination or employees in identical situations are treated differently.[333] This variability in decision-making can lead to a lawsuit for discrimination. For example, when multiple reasons are given for an employee's termination, courts are likely to find that this was pretext for discrimination.[334] When an employee in a protected class receives an adverse employment action, but a similarly situated employee who is not in the class does not, courts would likely find discrimination.[335] Using AI to create consistency in employment decisions will avoid these types of lawsuits.

---

331. *Id*.

332. *See* William T. Bielby, *Minimizing Workplace Gender and Racial Bias*, 29 CONTEMP. SOC. 120, 123-27 (2000) (citation omitted) ("[P]ersonnel systems whose criteria for making decisions are arbitrary and subjective are highly vulnerable to bias due to the influence of stereotypes—as, for example, when individual managers have a great deal of discretion with little in the way of written guidelines or effective oversight."); Barbara F. Reskin, *The Proximate Causes of Employment Discrimination*, 29 CONTEMP. SOC. 319, 323-27 (2000) (arguing organizations can minimize stereotyping and bias by using "objective, reliable, and timely information that is directly relevant to job performance in personnel decisions"); Bornstein, *Reckless Discrimination*, *supra* note 90 at 1098 (suggesting that reducing the opportunity for subjective decision-making can be effective in reducing the effects of stereotyping and implicit bias).

333. I use the term "noise" to maintain consistent terminology in this article. As explained, noise is unjustified inconsistency and variability in human decision-making. Although this term has not yet been used in the context of employment discrimination law, I anticipate it will become a rich vein of research in future law review articles after Kahneman's book NOISE comes out in 2020 or 2021.

334. Velez v. Thermo King de Puerto Rico, 585 F.3d 441 (1st Cir. 2009). *See also* Pierson v. Quad/Graphics Printing Corp., 749 F.3d 530 (6th Cir. 2014) (reversing summary judgment because multiple conflicting reasons for the termination would present the jurors with a triable issue of fact); Hitchcock v. Angel Corps, Inc., 718 F.3d 733 (7th Cir. 2013) (reversing summary judgement on grounds that the jury should be allowed to determine whether the reason for termination was a pretext due to the four conflicting reasons given for her termination).

335. Int'l Bhd. of Teamsters v. United States, 431 U.S. 324 (1977) (setting forth the standard for "disparate treatment" cases which turns on whether or not the Plaintiff is treated differently than someone "similarly situated.")

One particularly promising use of AI to avoid discrimination suits is in the area of promotions. Under Title VII, an action may be brought when an employee in a protected class is passed over in favor of a similarly or lesser qualified employee who is not in a protected class.[336] Courts have held that subjective promotion criteria can raise an inference of discrimination.[337] Because promotion criteria is seldom explicit in tech start-ups, for example, there is a great deal of inconsistency in the criteria used for promotions.[338] Although unconscious biases may give rise to men being preferred for promotion over women, as Part II *supra* indicates, the courts have been very inconsistent in their treatment of gender discrimination based on unconscious bias evidence. However, plaintiffs have been more successful in cases where inconsistent treatment of employees due to the use of subjective criteria for promotions can be demonstrated.[339] As discussed earlier the rates of promotion for men and women in the tech field are significantly different.[340] An algorithm can be used to evaluate employees using objective criteria to avoid these types of lawsuits as well. As mentioned earlier, AI can also be used to achieve more objectivity and consistency in conducting interviews through the use of online structured formats or chatbots. Another benefit is that a chatbot used to conduct interviews will not ask illegal questions as some one in five humans do

336. Under Title VII, to prove promotion discrimination, an employee who is qualified, but in a protected class, must demonstrate that they were passed over for promotion in favor of a person with similar or lesser qualifications. Reeves v. Sanderson Plumbing Prods., Inc., 530 U.S. 133, 148 (2000); Jacobs v. N.C. Admin. Office of the Courts, 780 F.3d 562, 575 (4th Cir. 2015).

337. Watson v. Fort Worth Bank & Trust, 487 U.S. 977 (1988) (ruling that Title VII provides relief for denial of promotion due to subjective criteria); Garrett v. Hewlett-Packard Co., 305 F.3d 1210, 1217 (10th Cir. 2002) (ruling that evidence of pretext may be demonstrated with the use of subjective criteria); McCullough v. Real Foods, Inc., 140 F.3d 1123, 1129 (8th Cir. 1998) ("[S]ubjective criteria for promotions 'are particularly easy for an employer to invent in an effort to sabotage a plaintiff's prima facie case and mask discrimination.'" (citing Lyoch v. Anheuser-Busch Cos., 139 F.3d 612, 615 (8th Cir. 1998))).

338. Pawel Rzmkiewicz, *Recruiting Methods for Startups: Balancing Objectivity & Subjectivity for Tech Roles,* MODERN RECRUITER (June 1, 2017), https://perma.cc/U3BJ-K9JZ.

339. Butler v. Home Depot, Inc., Nos. C-94- 4335 SI, C-95-2182 SI, 1997 WL 605754, at *7 (N.D. Cal. Aug. 29, 1997) (sustaining certification of a sex discrimination class action challenging hiring and promotion practices and quoting expert testimony explaining that "[i]n the context of a male-dominated culture, relying on highly arbitrary assessments of subjective hiring criteria allows stereotypes to influence hiring decisions").

340. McKinsey, *supra* note 50.

during the interview process.[341] By incorporating AI to make consistent decisions, the risks of variability in employment actions will be lessened.[342]

Despite the numerous advantages from a legal perspective to using AI to achieve more objectivity in employment decisions, there are a number of additional legal issues that warrant further attention. The following briefly recognizes several of these issues. Organizations must know where their data is coming from and actively seek to examine and remedy bias and variability in it. Those seeking to incorporate AI into their employment decision-making should guard against data mined from the internet, especially social media, and from data brokers, as it is likely to be biased and error-prone.[343] If using internal data, examine it for balance, meaning that the set is not predominantly one race or sex. The most important aspect to reducing algorithmic bias is making a significant investment in clean data. In addition, take care in designing the program. Avoid irrelevant classifications, vague outcomes (such as a "good employee") or permitting a single or homogenous group of programmers to create algorithms.[344] Determine the desired outcome in advance and test any assumptions. In other words, organizations should not assume that those with a college degree will be better employees than those without. Auditing of outcomes is also recommended to uncover any discriminatory results from the use of

341. Questions about race, age, disability, marital status, number of children, place of worship, are not permitted during the interview process. Chatbots are given specific questions to ask resulting in the solicitation of identical categories of information from each candidate (known as structured interviews) which has been shown to not only eliminate bias, but also avoid illegal questions. Anthony Tattersall, *Using Analytics to Stamp Out Workplace Bias*, LAUNCHPAD, https://perma.cc/5B22-MHG5 (archived Apr. 16, 2019).

342. As Kahneman explains, noise can be an invisible problem. "When you have a judgment and it is the noisy judgment, your judgment is determined largely by chance factors. You don't know what these factors are. You think you're judging based on substance, but in fact, that's just not the case." Matias, *supra* note 172. In order to prevent errors, Kahneman suggests using an algorithm rather than a person. He notes, however, that people "hate to be replaced by an algorithm." *Id.*

343. *See* Siegel, *supra* note 255; Chandler, *supra* note 212; Pasquale, *supra* note 297; Mayer-Schönberger & Cukier, *supra* note 255; and O'Neil, *supra* note 255.

344. Firms, such as Pymetrics, match skills discovered with certain positions, not skills with the definition of a "good employee." When the game determines someone's risk comfort, this information would be used to suggest appropriate positions. For example, you would want someone who is more risk averse in your legal or accounting department and less risk averse in your sales department. One of the benefits of games is that, unlike personality exams, you cannot guess what answers the employee is seeking. Chris Ip, *To Find a Job, Play These Games*, ENGADGET (May 4, 2018), https://perma.cc/3YJB-T95N.

an AI program.[345] If someone wants to override a machine decision, the reasons should be documented and only in exceptional circumstances. Furthermore, creating an AI Council ensures the quality of your data, examine classifications for legitimacy, and continually monitor outcomes for bias.[346]

When conducting employee surveys, employers should inform what the surveys will be used for and make sure to note that questions regarding race, gender, health, and veteran status are voluntary.[347] If the employer is gathering information from emails and DMs used on company equipment, counsel will need to examine state law and the Stored Communications Act and the Electronic Communications Privacy Act to ensure that such monitoring is permitted.[348] Anytime a business uses and maintains data, there are privacy and security issues.[349] If any type of employee testing is done, statutes such as the Americans with Disabilities Act must be considered to make sure that the use of any analytics program, especially games, does not factor these individuals out in a way that violates the law.[350]

---

345. See Kroll, *supra* note 300 and Kim, *Auditing Algorithms*, *supra* note 300 for discussion on legality of auditing algorithms.

346. For guidelines on the creation of trustworthy AI, see EUROPEAN COMM'N, *supra* note 320.

347. For examples of illegal interview questions see *Illegal Interview Questions*, BETTERTEAM (Jan. 6, 2019), https://perma.cc/HM8W-89GB

348. Society for Human Resource Management, *State Surveillance and Monitoring Laws*, https://perma.cc/U6QH-FFHR (archived Apr. 16, 2019). *See also* Ifeoma Ajunwa, *Algorithms at Work: Productivity Monitoring Applications and Wearable Technology as The New Data-Centric Research Agenda for Employment and Labor Law,* 63 ST. LOUIS U. L.J. (2019, forthcoming) (examining potential legal issues regarding productivity monitoring of employees).

349. *See, e.g*., Ajunwa, *supra* note 348; Ifeoma Ajunwa, Kate Crawford & Jason Schultz, *Limitless Worker Surveillance*, 105 CALIF. L. REV. 735 (2017) (discussing worker privacy issues due to advances in big data analytics, communications capture, mobile device design, DNA testing, and biometrics); Ronald C. Brown*, Measuring Worker Performance Within the Limits of Employment Law in the Changing Workplace Environment of Industry 4.0* (May 25, 2018) (unpublished manuscript), https://perma.cc/6EGT-SRKY (discussing privacy and other legal consideration in using technology to evaluate employee performance); Citron & Pasquale, *supra* note 297 (arguing that, because of the lack of legal oversight for automated decisions, due process safeguards should be implemented); Pauline T. Kim & Erika Hanson, *People Analytics and the Regulation of Information Under the Fair Credit Reporting Act*, 61 ST. LOUIS U. L.J. 17 (2016) (discussing limitations on employee data collection activities under the Fair Credit Reporting Act); Karl M. Manheim & Lyric Kaplan, *Artificial Intelligence: Risks to Privacy and Democracy*, 21 YALE J. L. & TECH. 106 (2019) (discussing risks of AI to decisional and informational privacy and security).

350. *See, e.g*., Allan G. King & Marko J. Mrkonich, *"Big Data" and the Risk of Employment Discrimination*, 68 OKLA. L. REV. 555 (2016) (describing the potential for

If a discriminatory outcome is detected, rather than scrap the entire program, like Amazon did, conduct further investigations as to the source of the biased outcome and seek to remedy it. Research confirms that machines are capable of not only making more accurate decisions than humans,[351] it also confirms that in the area of employment decision-making, their ability to override bias and noise will results in a greater diversity of hires, fairer promotion decisions, and better retention of employees through early detection of unhappiness.

X. CONCLUSION

Although the tech industry holds itself out as being committed to diversity, it has failed to make any meaningful progress since the first diversity report came out in 2014. It is estimated that U.S. companies will be unable to fill the 1 million open tech positions in 2020. Because the tech industry accounts for 20% of the country's output, if the United States is to remain an economic competitor in the world, it must be able to fill tech jobs by vastly expanding its applicant pool beyond the usual suspects.

The tech industry and legal system have failed women miserably. The excuses given for the lack of women in tech do not hold up. There are a significant number of women from around the world who graduate with computer science degrees and U.S. colleges have made advances in recent years in increasing their numbers. However, many women are alienated during the recruiting process due to gendered job ads and sexist behavior during the interviews. The conditions women face in the tech field, widespread disrespect, sexism, harassment, stereotyping, exclusion from

---

violating the Americans with Disabilities Act using data analytics).

351. *See, e.g.*, Wu Youyou, Michal Kosinski, & David Stillwell, *Computer-Based Personality Judgments Are More Accurate than Those Made by Humans*, 112 PROC. NAT'L ACAD. SCI. 1036, 1036 (Jan. 27, 2015) ("This study compares the accuracy of human and computer-based personality judgments, using a sample of 86,220 volunteers who completed a 100-item personality questionnaire. We show that (i) computer predictions based on a generic digital footprint (Facebook Likes) are more accurate (r = 0.56) than those made by the participants' Facebook friends using a personality questionnaire (r = 0.49); (ii) computer models show higher interjudge agreement; and (iii) computer personality judgments have higher external validity when predicting life outcomes such as substance use, political attitudes, and physical health; for some outcomes, they even outperform the self-rated personality scores."); Nathan R. Kuncel et al., *Mechanical Versus Clinical Data Combination in Selection and Admissions Decisions: A Meta-Analysis,* 98 J. APPLIED PSYCHOL. 1060 (2013) (stating that using mechanical means, such as algorithms, to predict job or academic performance is 50% more accurate than holistic methods, such as using experts or subjective human judgment).

networking events, and an inability to move up in the organization, cause half of the women who enter the field to leave. In addition, courts have shown a reluctance to find companies liable for non-overt discrimination leaving a gaping hole in remedies. Class action suits are rarely certified, and most tech companies have arbitration or confidentiality requirements that prevent women from getting their day in court.

Although some believe that the tech industry's homogenization is intentional, it is more likely that the lack of diversity is due to the unconscious biases and noise present in human decision-making. Because most leaders in the tech field are men who rely on gut feeling and personal opinion to make employment decisions, their unconscious biases have created a vicious circle of hiring and promoting young white men. As Daniel Kahneman explains, human judgments are untrustworthy due to cognitive biases of which they are unaware. Mental shortcuts result in decisions influenced by prejudice and bias. In addition, human decision-making is also inconsistent. This noise presents itself as chance variability. Because these flaws are not easy to counter in humans, the situation never improves. No lawsuits, legislation or new theories of liability are going to solve this crisis. This urgently needed solution must come from within the tech industry itself.

It is only by introducing consistent objective decision-making into talent-management decisions that bias and noise can be mitigated. This is where AI is superior to humans. The use of AI in talent-management decisions has shown success in not only creating more successful hires, but in also creating a more diverse slate of candidates and employees. While some companies have embraced these new technologies, others fear that AI may actually cause discriminatory outcomes. As discussed, the phenomena of "garbage in, garbage out" is real, but can be addressed paying attention to the data sets by using known sources and making sure the sets are balanced and representative of all groups. Additionally, the algorithmic process and outcomes must also be monitored. The "black box" problem can be addressed in multiple ways including testing for bias, providing indications of influence, and auditing for fairness. By increasing the diversity of programmers in the tech industry bias can be considered and prevented in the creation of AI programs. As Fei-Fei Li warns, diverse programmers must be hired now: "Trying to reverse [biased systems] a decade or two from now will be so much more difficult, if not close to impossible."

The growth of AI applications is not going to slow down, however, we need to ensure that it is developed and used responsibly. With new collaboration between disciplines such as law, psychology, economics, business, engineering, technology and the social sciences, data sets are being developed which more accurately reflect the demographics of the society in which they exist and open source fixes are being created to remedy potentially biased outcomes. By increasing diversity in the tech industry, we will have more eyes and heterogenic perspectives overseeing the development of AI. To be clear the answer is not to replace all human decision-making with machines, but rather take advantage of the ability of a machine to make decisions without noise (because an algorithm will provide the same outcome for any given input, unlike the variability in outcomes of human decision-making) and with less bias than humans (because algorithms can be designed to review only the relevant employment criteria unlike with human decisions).

Do the risks of incorporating AI into employment decisions outweigh the benefits? From a purely legal standard point, it seems that despite claims of increased discrimination using AI, scholars believe the risk of liability is very small; however, the truth is more nuanced. Any potential discrimination detected in outcomes will most likely stem from human biases contained in data and not by virtue of the use of AI itself. AI alone will not fix the diversity problem in tech but addressing the unconscious biases prevalent in this industry is a first and vital step. It does not work to shame management or require diversity training. Nor does it serve to delay incorporating AI into your decision-making because of fear of discriminatory results. What works is removing subjective criteria from employment decisions, improving working conditions and the culture for everyone in these tech companies, and providing oversight and accountability for creating a diverse working environment. Most importantly, while it is clear that bias cannot be removed from human decision-making, it can be mitigated with machine decision-making. In fact, with the rapid development of responsible AI, there may come a time in the not so distant future when courts will find companies still using human decision-making in employment to be a prima facie showing of discrimination.

While this paper mentions specific solutions as examples of ways to incorporate AI into employment decision-making, it is not meant as a limitation, but rather a starting place. It is intended to serve as an alternative more optimistic view in contrast to the dire warnings about the use of AI. AI has enormous potential to address societal ills by quickly and efficiently discovering where bias exists and how to root it out. This can have enormous implications for correcting the societal injustices befalling historically disadvantaged groups.

As Malcolm Gladwell wrote in TIPPING POINT, "Social change has always followed this pattern; slowly emerging theories or paradigms accelerate exponentially, suddenly appearing everywhere." There is no stopping the development of AI at this point, and there is no reason to. I encourage the tech industry to take the lead in developing it responsibility—not only for the benefit of their own organizations—but for the benefit of society and our economy. We have jobs that need to be filled, and we know that human decision-making is flawed. Addressing the underlying problem of noise and bias in human decision-making will shift the needle towards a more diverse and inclusive workplace. It is my hope that companies will make these fixes open source, share best practices, and advance public understanding of how AI can be used for the greater good. The question is not whether AI should be incorporated into decisions regarding employment, but rather why in 2019 are we still relying on faulty human-decision making.