

SHOOK, HARDY & BACON

MAY 8, 2024

Be Aware: Your AI Has Implicit Bias, Too

Presented By | Caroline Gieser and Yara Rashad

SHOOK
HARDY & BACON

2024

1

SHB.COM

SHOOK

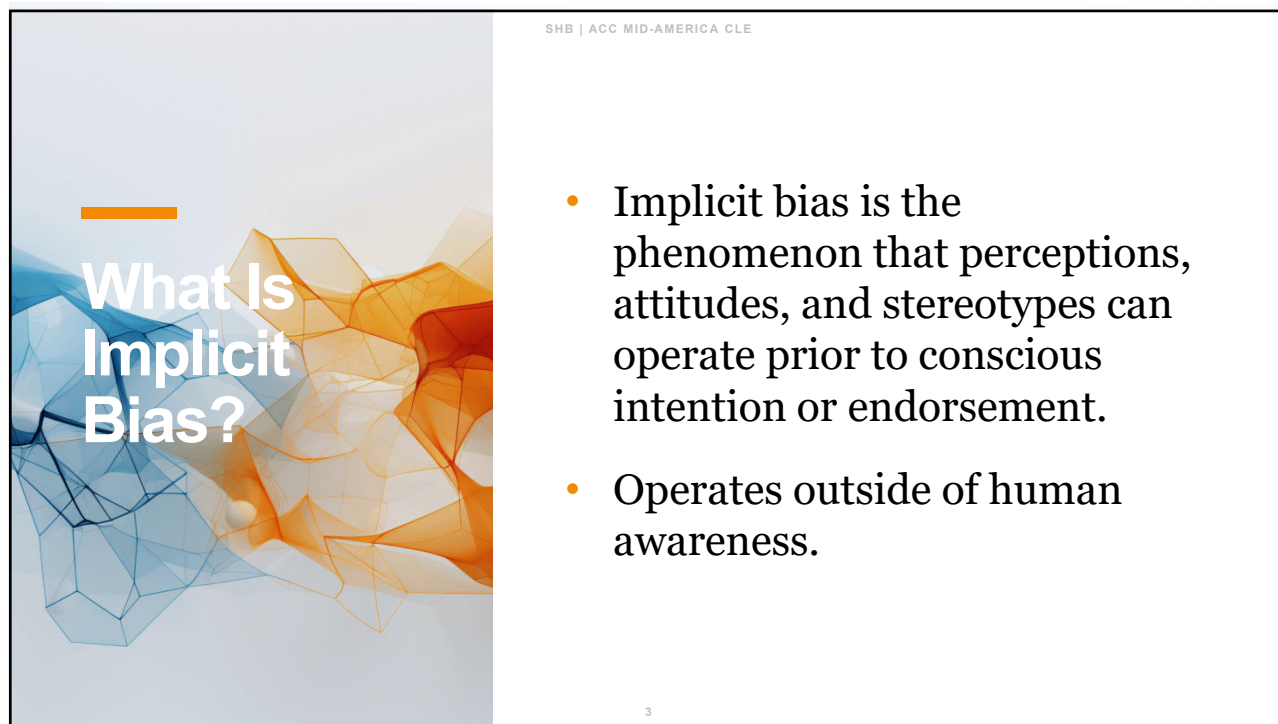
SHB | ACC MID-AMERICA CLE

Roadmap

- 1 Implicit Bias Defined
- 2 Implicit Bias In AI
- 3 Mitigating Risk of Implicit Bias in AI
- 4 Eliminating Implicit Bias in AI

2

2



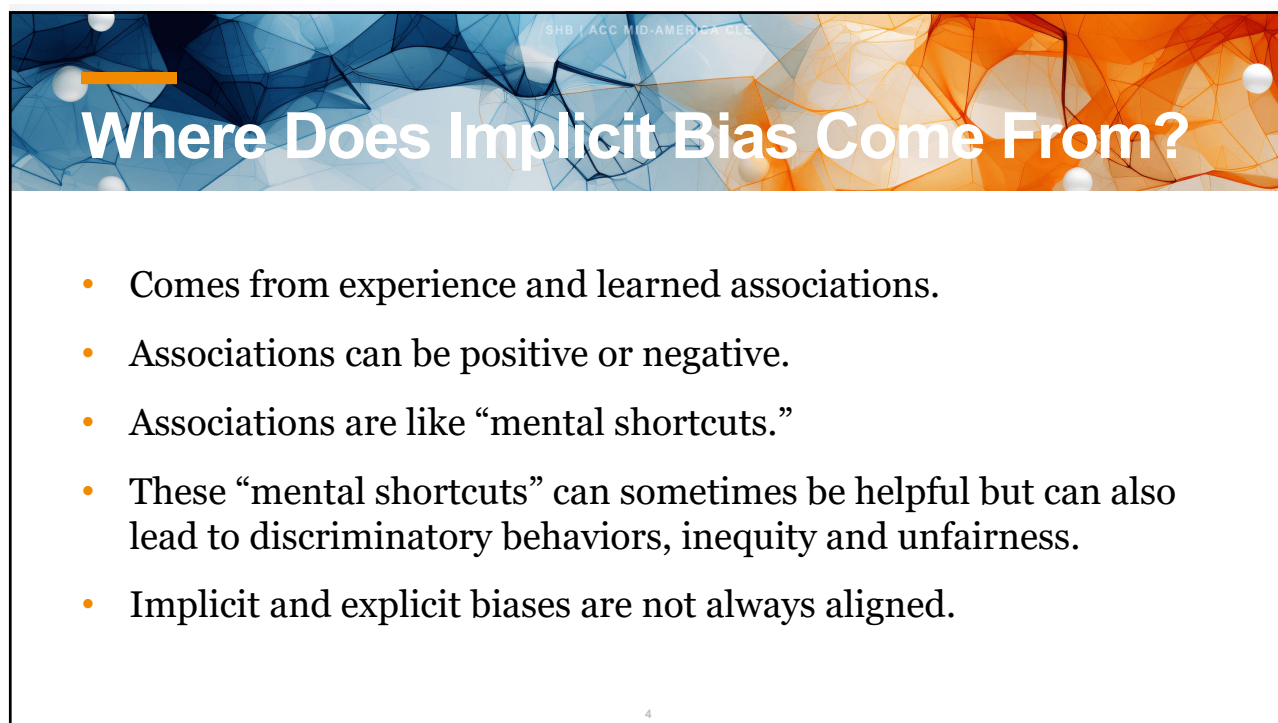
SHB | ACC MID-AMERICA CLE

What Is Implicit Bias?

- Implicit bias is the phenomenon that perceptions, attitudes, and stereotypes can operate prior to conscious intention or endorsement.
- Operates outside of human awareness.

3

3



SHB | ACC MID-AMERICA CLE

Where Does Implicit Bias Come From?

- Comes from experience and learned associations.
- Associations can be positive or negative.
- Associations are like “mental shortcuts.”
- These “mental shortcuts” can sometimes be helpful but can also lead to discriminatory behaviors, inequity and unfairness.
- Implicit and explicit biases are not always aligned.

4

4

Examples of Unconscious Bias Errors

ERROR	DEFINITION
Gender Bias	The tendency of individuals to make generalizations based upon personal values, perceptions and outdated, traditional views about men and women's roles in the workplace.
Generational or Age Bias	The tendency of individuals to make generalizations based upon personal values, perceptions, or views about different generations in the workplace.
Background or "Like Me" Bias	The tendency of individuals to positively view people who resemble themselves or have had similar experiences higher than others that have different backgrounds or experiences.
"Non-Native Speaker" Bias	The tendency of individuals to positively or negatively view people who do not speak the native language, or speak with an accent, lower than others that speak the native language.
Attractiveness Effect	The tendency of individuals to positively or negatively view people who are physically attractive.
First Impression Bias	The tendency of individuals to make an initial positive or negative judgment of a person and allow the first impression to positively or negatively view the individual at a later time.
Halos/Horn Effect	The tendency of individuals to make generalizations based upon one experience with an individual, and cause them to assume the same is true about all aspects of that individual.
Past Performance Error	Permitting an individual's poor (or excellent) performance in a previous rating period to color the manager's judgment about their performance in this rating period.
Stereotyping	The tendency of individuals to make generalizations based upon one aspect of diversity.

Ways Implicit Bias Creeps into AI

1. AI tools trained with biased data from biased human decision making.
2. Training data is not an adequately diverse sampling of data.
3. Data biased in the way it is gathered or chosen for use.

AI Tools Trained with Biased Data from Biased Human Decision Making

- Amazon found that one of its AI tools for hiring favored applicants who used the words executed and captured.
- Words more likely to be on the resumes of male candidates.

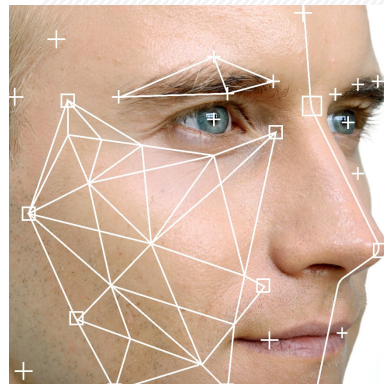


7

7

Training Data is Not an Adequately Diverse Sampling of Data

- AI tools with facial recognition technology are less accurate for ethnic minorities.
- Use by police departments has led to wrongful arrests.



8

8

Data Biased in the Way it is Gathered or Chosen for Use

- Correctional Offender Management Profiling for Alternative Sanctions Tool (COMPAS).
- AI tool used to assess recidivism risk for criminal defendants.
- Data inputs come from answers to a written questionnaire given to criminal defendants at booking.
- Black defendants who did not reoffend within two years after release were nearly twice as likely to have been classified as high risk by the COMPAS tool compared to white defendants.
- White defendants who did reoffend within two years after release were nearly twice as likely to have been labeled as low risk of recidivism than black defendants.

9

9

SHB | ACC MID-AMERICA CLE

Consequences of Implicit Bias

1. Inaccurate results
2. Discrimination
3. Disparate treatment of minorities

10

10

Inaccurate Results

- Numerous studies report that facial recognition technology is flawed and biased, with significantly higher error rates when used against people of color.
- Microsoft and IBM's facial recognition software had error rates of 35% and 21% when identifying darker female faces versus 0.3% and 0% when identifying white male faces.
- Google's speech recognition software (the highest performer as compared to Bing, AT&T, WIT, and IBM Watson Systems) was found to be 13% more accurate for men than women.

11

11

Discrimination

- In 2018, Google's facial recognition software identified two African Americans as "gorillas" in Google Photos.
- In 2010, Nikon's face detection cameras consistently displayed the message "did someone blink?" when an Asian face was photographed even when their eyes were wide open.

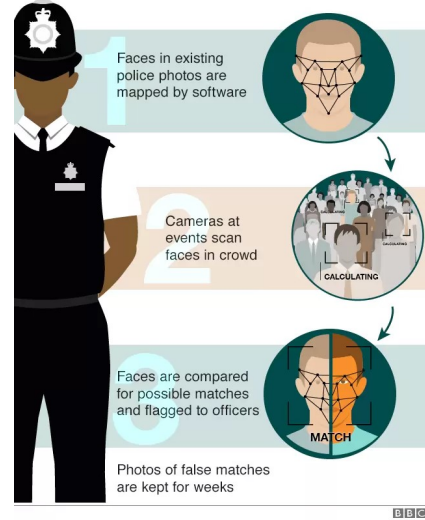
12

12

AI in Law Enforcement: Facial Recognition

- A study conducted by the National Institute of Standards and Technology (NIST) found “empirical evidence for the existence of demographic differentials in the majority of face recognition algorithms used by law enforcement.
- The majority of algorithms tested falsely identified African-American and Asian faces between 10 to 100 times more than Caucasian faces and misidentified African-American females when comparing their photo to others in a database.
- Several US cities have banned the use of such technology.

“How does live facial recognition work?”



AI in Law Enforcement: Risk Assessment Instruments

PATTERN	COMPAS	PREDPOL
<ul style="list-style-type: none"> • Uses factors such as current age, nature of the convicted offense, education status, incident reports, and financial responsibility to assess the likelihood that individuals will engage in crime upon their release (minimum, low, medium, or high risk). • NIJ researchers reported that PATTERN overpredicts the risk that Black, Hispanic, and Asian people will reoffend or violate their parole. 	<ul style="list-style-type: none"> • Used in many jurisdictions to make decisions about pretrial release and sentencing. • Issues a statistical score between 1 and 10 to quantify how likely a person is to be rearrested if released. 	<ul style="list-style-type: none"> • Created by UCLA scientists in 2008. • Identifies the areas in a neighborhood where serious crimes are more likely to occur during a particular period. • Used by more than 60 police departments across the country.

AI in Employment

- 55% of human resources leaders in the United States use predictive algorithms in hiring.
- Many companies employ algorithmic systems to:
 - Analyze resumes;
 - Improve how resumes appear to other algorithms;
 - Predict job performance through specialized assessments/questionnaires;
 - Transcribe recorded statements to text, then analyze those textual responses with natural language processing;
 - Determine retention, salaries, and promotion through “on-the-job surveillance.”
- Algorithms used to analyze natural language, such as in resumes or transcribed from interviews, have demonstrated biases against women and people with disabilities.
- Speech recognition models have demonstrated clear biases against African Americans and potential problems across dialectical and regional variations of speech.
- Algorithms that disseminate job postings can unintentionally result in biased outcomes against young women for STEM jobs and similarly ageism against older candidates.

15

15

AI in Consumer Lending

- Use of algorithmic lending systems can make significant progress in allowing disadvantaged populations—specifically black Americans—obtain housing loans.
- AI and algorithmic systems remove much of the human element from the underwriting process, lowering the potential for bias.
- A 2021 study found AI-based systems to reduce racial bias by 40% and show no discrimination in rejection rates.

16

16

Mitigating + Eliminating the Risk of Implicit Bias

LAW ENFORCEMENT

Train judges, parole boards and officers to understand the pertinent caveats when they're given the calculated probability a black suspect, defendant or convict will reoffend.

EMPLOYMENT

Reduce bias at the technical level within language models by masking race, gender and demographic information from applicants and rewording job descriptions to attract more diverse candidates or by changing the outcome variable that is being predicted.

CONSUMER LENDING

Widen the data criteria used for risk evaluation decisions.
 Include a diverse team of loan officers in the final step of the risk evaluation process.
 Implement alternative credit scoring models that include rental history payments.

Resource Guide

- [Black Students Face Racial Bias in School Discipline \(Forbes, April 5, 2019\)](#)
- [The 5 Biases Pushing Women Out of STEM \(Harvard Business Review, March 24, 2015\)](#)
- [ABA Diversity and Inclusion 360 Commission Toolkit](#)
- [Why Many CEOs are Tall People, The Height of the Matter \(Premium News Magazine, May 15, 2016\)](#)
- [Racial Bias, Even When We Have Good Intentions \(New York Times, January 3, 2015\)](#)
- [9 Resources for Teaching Unconscious Bias \(Inst. for Humane Educ., October 13, 2017\)](#)
- [The Problem of Algorithmic Bias in Autonomous Vehicles \(U. of Mich. Law and Mobility Program, March 12, 2019\)](#)
- [Facial Recognition Software Wrongly Identifies 28 Lawmakers As Crime Suspects \(npr.org, July 26, 2018\)](#)
- [Harvard Implicit Bias Test](#)
- [The Risks of Bias in Artificial Intelligence \(New York City Bar Center for Continuing Legal Education\)](#)
- [AI/ESQ: Impacts of Artificial Intelligence in Lawyer-Client Relationships. \(Oklahoma Law Review\)](#)
- [AI and Machine Learning – What Do We Know About the Biases in AI? \(Harvard Business Review, Oct. 25, 2019\)](#)
- [The Problem with Biased AIs \(and How to Make AI Better\) \(Forbes, Sep. 30, 2022\)](#)
- [Wrongfully Arrested Because Face Recognition Can't Tell Black People Apart, ACLU \(June 24, 2020\)](#)
- [Machine Bias \(Pro Publica, May 23, 2016\)](#)
- [How We Analyzed the COMPAS Recidivism Algorithm \(Pro Publica, May 23, 2016\)](#)
- [Microsoft, IBM Facial Analyses Struggle With Race and Gender \(WIRED, Feb. 2018\)](#)
- [Facial recognition fails on race, government study says \(BBC, Dec. 2019\)](#)
- [Voice Recognition Still Has Significant Race and Gender Biases \(Harvard Business Review, May 2019\)](#)
- [Report: AI Algorithms Can Reduce Bias, Create More Equitable Mortgage Lending – MortgageOrb](#)
- [Reducing bias in AI-based financial services | Brookings](#)
- [A Popular Algorithm Is No Better at Predicting Crimes Than Random People - The Atlantic](#)
- [Predictive policing algorithms are racist. They need to be dismantled. | MIT Technology Review](#)
- [What Happens When Police Use AI to Predict and Prevent Crime? - JSTOR Daily](#)
- [New study finds AI-enabled anti-Black bias in recruiting - Thomson Reuters Institute](#)
- [Research shows AI is often biased. Here's how to make algorithms work for all of us | World Economic Forum \(weforum.org\)](#)



SHB | ACC MID-AMERICA CLE

Contacts



Caroline Gieser

Associate
Atlanta
cgieser@shb.com



Yara Rashad

Associate
Chicago
yrashad@shb.com

