



# Min Kyung Lee

## Who Is Included in Human Perceptions of AI?: Trust and Perceived Fairness around Healthcare AI and Cultural Mistrust

Min Kyung Lee  
School of Information  
University of Texas at Austin  
Austin, TX, United States  
minkyung.lee@austin.utexas.edu

Kate Rich\*  
School of Information  
University of Texas at Austin  
Austin, TX, United States  
k.rich@utexas.edu

### ABSTRACT

Emerging research suggests that people trust algorithmic decisions less than human decisions. However, different populations, particularly in marginalized communities, may have different levels of trust in human decision-makers. Do people who mistrust human decision-makers perceive human decisions to be more trustworthy and fairer than algorithmic decisions? Or do they trust algorithmic decisions as much as or more than human decisions? We examine the role of mistrust in human systems in people's perceptions of algorithmic decisions. We focus on healthcare Artificial Intelligence (AI), group-based medical mistrust, and Black people in the United States. We conducted a between-subjects online experiment to examine people's perceptions of skin cancer screening decisions made by an AI versus a human physician depending on their medical mistrust, and we conducted interviews to understand how to cultivate trust in healthcare AI. Our findings highlight that research around human experiences of AI should consider critical differences in social groups.

### CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**.

### KEYWORDS

Perceptions of Algorithmic Decisions, Trust, Fairness, Healthcare AI, Group-Based Medical Mistrust Scale (GBMMS), Black Perspectives

### ACM Reference Format:

Min Kyung Lee and Kate Rich. 2021. Who Is Included in Human Perceptions of AI?: Trust and Perceived Fairness around Healthcare AI and Cultural Mistrust. In *CHI Conference on Human Factors in Computing Systems (CHI '21)*, May 8–13, 2021, Yokohama, Japan. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3411764.3445570>

\*The second author is a Ph.D. student in the Department of Communication, University of Washington. The research was conducted at The University of Texas at Austin.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

CHI '21, May 8–13, 2021, Yokohama, Japan

© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 978-1-4503-8096-6/21/05...\$15.00 <https://doi.org/10.1145/3411764.3445570>

### 1 INTRODUCTION

Artificial Intelligence (AI) is increasingly automating decision-making in diverse industry sectors. In many high-stakes domains such as healthcare, education, criminal justice systems, organizational management and public assistance [8, 9, 25, 40], AI systems are automating or augmenting decisions that human experts used to make. In order to better understand people's reactions to this change, many scholars have investigated how people perceive algorithmic decisions as compared to human decisions [6, 10, 20, 23, 24, 26]. Their findings suggest that people tend to perceive algorithmic decisions as inferior to human decisions and are resistant to following them. In several studies, people trusted algorithmic decisions less than human decisions and were less likely to adopt them, particularly when tasks were deemed to require a human's unique capabilities [23], be subjective [6] or require attention to individual uniqueness [27].

However, not everyone has an equal level of trust in human decision-makers. Those who experience marginalization from other humans may have less faith in human decisions. Specifically, anti-Black racism across multiple human-led institutions has created a climate of mistrust among many Black and African American people in the United States [31]. Do people who mistrust human decision-makers also perceive them to be more trustworthy and fairer than algorithmic decisions? Or do they trust algorithmic decisions as much as or more than human decisions?

In this paper, we examine the role of cultural mistrust related to human systems in people's perceptions of algorithmic decisions. We focus on a healthcare context, group-based medical mistrust, and Black people living in the U.S., who have been shown to have higher medical mistrust than other populations. We conducted a between-subjects online experiment to examine people's perceptions of skin cancer screening decisions made by an AI versus a human physician depending on their medical mistrust. We recruited a balanced pool of Black and white participants. The result suggests that the previous literature's finding is replicated with participants with low mistrust in human systems, but not with those with high mistrust. Participants with low mistrust trusted human decisions more than algorithmic decisions and regarded them as fairer. However, participants with high mistrust in human systems perceived algorithmic and human decisions to be equally trustworthy and fair. We conducted interviews with 21 participants to understand what contributes to mistrust in healthcare AI and what information might cultivate their trust in healthcare AI.

In this work, we make a contribution to research around human experiences of AI in the fields of Human-Computer Interaction (HCI), psychology and communication. Our work offers new insight