ASEE 2022 ANNUAL CONFERENCE
Excellence Through Diversity
MINNEAPOLIS, MINNESOTA, JUNE 26TH-29TH, 2022 ASEE
Paper ID #38054

(Work in Progress) Examining how students critically evaluate racial bias in a medical device in a first-year computing course

Tyrine Jamella Pangan (STEM Education PhD Student)

Tyrine Jamella Pangan is a STEM Education PhD student at Tufts University and a Graduate Research Assistant at the Tufts University Center for Engineering Education and Outreach (CEEO). She is interested in integrating social and emotional learning (SEL) in engineering, specifically within the elementary school context. Tyrine hopes to explore how Transformative SEL can be implemented to cultivate socially responsible engineers.

Chelsea Andrews (Research Assistant Professor)

Chelsea Andrews is a Research Assistant Professor at Tufts University at the Center for Engineering Education and Outreach.

(Work in Progress) Examining how students critically evaluate racial bias in a medical device in a first-year computing course

Introduction

Despite a growing consensus of the urgency to consider the social, economic, and political aspects of engineering in undergraduate engineering education, the technical/social dualism still prevails in engineering culture; technical competencies continue to be prioritized over so-called "soft skills" (Cech, 2014; Faulkner, 2007). As a result of this dualism, engineering students are misled into thinking that engineering exists within a vacuum (Johnson et al., 2019; Trevelyan, 2014). If their engineering courses focus solely on building their technical skills, students may enter the professional field with little experience in addressing bigger problems that affect members in society beyond the traditional stakeholders they were exposed to in their engineering programs, who are often those who hold the most power in society (Leydens & Lucena, 2017). Efforts to introduce the social dimensions of engineering are typically sprinkled into design and capstone courses (Leydens & Lucena, 2017) or designated to standalone ethics seminars (Hess & Fore, 2018). To better integrate these concepts, some researchers suggest starting at a smaller scale within existing engineering courses, such as engaging students in sociotechnical thinking through class assignments or rubrics (Claussen et al., 2019; Salzman et al., 2019).

Although academic institutions are now working to integrate the social, economic, and political dimensions of engineering in undergraduate engineering coursework, there remains a lack of fundamental research on how to engage students in these sociotechnical topics. This is particularly true within the scope of traditionally technical courses, such as the engineering sciences, where students are only expected to demonstrate technical competencies by the end of the course. This work-in-progress study explores the range of ways undergraduate students attended to sociotechnical dimensions in a first-year engineering computing course, by analyzing written reflection responses to readings focused on the racially biased outcomes of a ubiquitous medical technology, the pulse oximeter. These initial findings add to a growing body of literature on including sociotechnical topics within undergraduate courses, and will help inform pedagogical approaches to support students in developing sociotechnical ways of thinking within engineering.

Conceptual Framework for Developing Sociotechnical Literacy

This work-in-progress study is focused on a first-year computing course that has been redesigned to incorporate sociotechnical aspects of engineering alongside the original technical content. We aim to center engineering and data science within larger sociotechnical systems; to have students delve deeply into the social, economic, and political impacts of data, algorithms, and related technologies. A main research area of the overall project is to investigate different pedagogical structures and supports and explore how these relate to shifts in students' sociotechnical literacy over the course of the semester. In this work-in-progress paper, analyzing some of the pilot data of this project, we analyze students' sociotechnical understandings mid-way through the semester, as expressed in a reading reflection assignment.

As sociotechnical literacy is a fairly new research topic in engineering education, there are currently a multitude of dimensions researchers are investigating. Due to the data science focus of this re-designed course, this project is focused on sociotechnical literacy in three dimensions: *bias*, *differential impacts*, and *responsibility*. As described in the paragraphs below, these dimensions include both content knowledge and skills. For example, if students are sociotechnically literate along the *bias* dimension, they will be able to: identify potential sources of bias in a dataset or algorithm, computationally analyze the effect of that bias, and propose ways to remove or correct for that bias.

Bias

In the bias dimension, sociotechnically literate students recognize no data set, algorithm, or technology should be considered to be objective, neutral, ahistorical, or "true" (Duschl, 2008; Manz, 2016; JafariNaimi, 2018). All human artifacts contain bias, not only due to unconscious bias on the part of creators, but because they are created within and reflect unjust systems. Students should also learn to recognize that impact is more important than intent, and that, as Ruha Benjamin (2019) puts it, there need not be a racist "boogeyman" behind the scenes for technology to have racist outcomes; as she states: "People looking for the boogeyman [are] really trying to hinge the analysis on the intentionality to do harm" (47:24).

As part of the bias emphasis of the course, students learn the technical skills of identifying sources of bias in data (e.g., sampling and non-sampling errors), how bias can be amplified through algorithms (e.g., through feedback loops), and some approaches to tackle bias (Feng & Wu, 2019). They learn about the role of third-party algorithmic audits (Buolamwini & Raji, 2019), de-centering the creators as best situated to tackle their own code. We also want students to recognize that the solutions to these problems are not solely technical. They need to grapple with the notion that "removing bias…though it may generate less ethically troubling results, will not fix the underlying social injustices" (Feng & Wu, 2019, Conclusion section, para. 1).

Differential impacts

The differential impact dimension includes noticing that different groups of people (and possibly more-than-human actors) systematically benefit or are harmed to different degrees, and these impacts are often aligned with historic power differentials. Rather than simply classifying technology as having "good" or "bad" outcomes, it is important for students to understand that the effects of technology are not felt evenly. Recognizing differential impact means observing a pattern of outcomes that are systematically more negative for one group of people compared to another, and that this pattern often amplifies historical oppression or marginalization (Riley, 2008). To do this work, students must be aware of and able to reason about how issues of power, history, and culture impact the designed world (McGowan & Bell, 2020).

Responsibility

In the responsibility dimension, we are interested in whether and in what ways students identify someone or something as responsible for causing or for fixing the issue under consideration. For example, do students blame individual designers for differential impacts of a technology

currently in wide use, or do they point to policy? While there is a general trend in justice-oriented work, including in some circles of engineering education, to consider macro-ethics and systemic causes to injustice (Gupta, 2017; Herkert, 2005), this trend has not yet broadly permeated undergraduate engineering analyses. As Pawley (2019) states, "we are accustomed to thinking about ethics in engineering but mainly in terms of individual accountability, micro-ethics, and direct (not systemic) causality" (p. 6). Becoming more sociotechnically literate involves moving away from focusing on individual accountability and towards being able to recognize systemic causality, rooted in a historical context.

Beyond assigning responsibility for harm that has been caused by existing technologies, students are also asked to struggle with the idea of who they are responsible for when (if) they become designers and engineers. With much of engineering still client-driven, institutions and people in power choose the problems that are to be solved, often at the expense of the broader community; we need to teach students to ask: "for whom and by whom is engineering done?" (Riley, 2008, p. 97).

Across all three of these dimensions, it is important for students to grapple with what a just outcome would be in different cases. Authentic engineering and data science problems do not permit simple solutions, even when it is clear that the current situation is problematic (Costanza-Chock, 2020).

Grounded in this conceptual framework, we focus in this study on a first-year computing course that has been redesigned to incorporate sociotechnical aspects of engineering alongside the original technical content. In this work in progress paper we address the research question, "*In what ways do undergraduate students in a first-year engineering computing course attend to the sociotechnical dimensions of bias, differential impacts, and responsibility in written reflection responses about the racially biased outcomes of pulse oximeters?"*

Methods

Study Context

This study comes from the pilot year data of a three-year, NSF-funded study focused on improving sociotechnical literacy in a first-year engineering course. The "Introduction to Computing for Engineering" course has traditionally focused on teaching students a coding language as well as introducing them to the basics of data science. As part of this study, this course was re-designed to support sociotechnical thinking by adding three components: (1) weekly 20-minute in-class discussions reflecting on assigned readings, (2) a sociotechnical mid-semester project, and (3) including sociotechnical reflections in students' self-defined final project. Students received additional support in sociotechnical topics from upper-class undergraduate "Equity Learning Assistants" (ELAs). Equity Learning Assistants attend separate weekly seminars that provide them with tools to support the first-year students in having sociotechnical conversations, then lead the weekly small and whole-group discussions during class.

This work-in-progress study focuses on students' written responses to a lesson addressing findings of racially biased measurement errors of pulse oximeters. Students were provided two articles from major media outlets and were assigned a written reflection on the readings prior to an in-class discussion. Our study focuses on these written student responses to explore how students are grappling with sociotechnical concepts. In particular, by looking into these responses, we can get insight into students' ideas before participating in facilitated group discussions, which provided additional scaffolding for sociotechnical thinking.

Participants

This study was conducted at a medium-sized private college in the northeast. In the pilot year of the study, two of the five sections of the first-year computing course added the sociotechnical components, including the weekly sociotechnical readings, written reflections, and in-class discussions. The pilot year of this study occurred during the COVID-19 pandemic; as a result, courses were held virtually via Zoom or in a hybrid format. Student written reading responses were collected from an online educational platform, Canvas. Most of the students in the course were first-year engineering students who were required to take the class; a fraction were non-engineering students, who tend to be in later years of study. All 76 of the students in the two participating sections of the course were asked to participate in the study; 39 consented to analyzing their coursework, which was provided to researchers completely anonymized. The School of Engineering overall at this university is more diverse than the engineering schools at most Predominantly White Institutions; this diversity was reflected in this course. In a separate survey, 59 of the students in these sections provided gender information: 61% of these were women and 3% were non-binary, agender, trans* or gender fluid; 61 students provided racial and ethnicity information: 26% of these were underrepresented minorities (specifically, Black/African American, Hispanic/Latinx, Middle Eastern, or multi-racial including at least one of these categories).

Lesson context

This study uses student coursework from week 11 of the course, out of 14 weeks. At this point, students were comfortable with the weekly cycle of pre-readings, written reflections, and in-class small group and whole-class discussions. Early topics included disability justice in the designed world, equity in public transit access, and data analysis of environmental justice topics. This week was the first time students had read about racial bias in medical equipment.

Two recent articles on pulse oximeters were assigned to students for this week, both from popular news sources. The first article provided students with background information on the device, reviewed published research paper results showing racial bias, and situated these findings in historical context (Moran-Thomas, 2020); the second, and much shorter, article focused on the FDA response to research showing different readings for different skin tones, and criticized the FDA for not explicitly labeling this as racial bias (Brodwin & St. Fleur, 2021).

The first article, published in Boston Review, first briefly discusses the technology behind the pulse oximeter, including how the device uses both infrared and red light to estimate a patient's oxygen levels (Moran-Thomas, 2020). Moran-Thomas then focuses on a series study conducted by researchers at the University of California, San Francisco (UCSF) that found racially biased

results in pulse oximeter readings, specifically, "pulse ox bias was generally the greatest in dark-skinned subjects, intermediate for intermediate skin tones, and least for lightly pigmented individuals" (Feiner et al., 2007, as cited in Moran-Thomas, 2020, para. 7). The article describes how, despite this bias, in practice, pulse oximeter readings are often seen as objective data, used to inform decision making in hospitals, such as when/if a patient is admitted, and used by insurance to determine whether or not to cover at-home oxygen. While there is controversy amongst physicians on how significant these discrepancies are in the context of all the other data healthcare providers take into account, the author argues that this dismissal contributes to systemic structures that "were designed to quantify—and justify—racial hierarchies" (para. 30) that have historically favored white patients.

In the second assigned article, from STAT news, the authors criticize a recent alert published by the U.S. Food and Drug Administration (FDA) warning the public about the limitations of pulse oximeter readings without explicitly relating the technological issues to race (Brodwin & St. Fleur, 2021). The FDA warning states, "Be aware that multiple factors can affect the accuracy of a pulse oximeter reading, such as poor circulation, skin pigmentation, skin thickness, skin temperature, current tobacco use, and use of fingernail polish" (U.S. Food & Drug Administration, 2021, as cited in Brodwin & St. Fleur, 2021, para. 3). The authors argue that the errors in oximeter readings are best described as racial bias; they reference studies "that found oximeters were nearly three times as likely to miss hypoxemia in Black patients compared with white patients" (Brodwin & St. Fleur, 2021, para. 2). Several physicians weighed in with their concerns with pulse oximeters based on the results of the 2020 paper.

The two articles were chosen as resources for the students to inform their sociotechnical thinking about the complexities of finding racial bias in a common medical device.

Students submitted short written responses to the articles, typically one paragraph long, based on the prompt:

> "After reading, as a submission for this assignment, provide some of your personal thoughts about these articles. This could be one or more of the following:
>
> - Ideas in the articles that stood out to you
> - Issues you had with the substance of the articles
> - Questions you have about what the authors wrote
> - Other areas of further exploration these articles made you wonder about
>
> The format should be just a few written sentences or bullet points."

Data Collection and Analysis

We took a case study approach to analyzing this data (Creswell & Poth, 2018), looking at written responses from both participating course sections. We utilized a thematic analysis approach, searching for patterns across the student responses from both sections (Braun & Clarke, 2006). Specifically, because we are interested in how students demonstrate sociotechnical literacy in reflective responses, we took a "theoretical" thematic analysis approach, which is "driven by the researcher's theoretical or analytic interest in the area" (p. 84).

We began our analysis guided by our conceptual framework, focusing on students' sociotechnical understandings along the dimensions of bias, differential impact, and

responsibility. In an iterative process, each author considered the entire set of 39 student responses and considered whether the three dimensions were found throughout the responses and whether any aspects of the responses were not well captured by the dimensions. We roughly coded student responses phrase-by-phrase to ensure each dimension was found in multiple instances of data. We found many examples of each dimension, and did not feel a need to add additional dimensions. This is not surprising, both because the dimensions are fairly broad, and because they were already potential focus areas of the researcher team in deciding on course topics. We then focused on identifying and describing sub-themes within each dimension. As we are not interested in drawing conclusions based on frequencies of themes in student responses, we have not conducted a formal coding process with interrater reliability.

Bias coding
We identified responses as pertaining to bias when students pointed out that the algorithm, device, and/or device testing procedure was biased. Many students also pointed out that bias does not need to be intentional.

Differential impacts coding
We identified responses pertaining to differential impact when students addressed differences in the quality of medical care of patients, the difference in requirements for insurance to cover treatment, and the historic impact on the difference in treatment for nonwhite patients.

Responsibility coding
We identified responses as pertaining to responsibility if students explicitly identified individuals, organizations, or institutions as making decisions or having influence over the design/use of the device, and recognizing that those decisions potentially led to adverse consequences. A few students also identified actors they believed should be responsible for remedying the current situation.

Results

First, we describe the general reactions students had to the information in the articles. Then, we break down the results by each of the three dimensions of sociotechnical literacy, giving examples of each sub-theme from student written responses.

General reactions

Many students expressed surprise to find out that research has shown racial disparities in a widely-used piece of medical equipment, even if they were aware of racial disparities in other aspects of society:

> I had no idea that racial disparities exist even in widely-used, common-place devices, so this was somewhat shocking to hear. [Student 34[1]]
>
> Reading about the racial biases in the pulse oximeter furthered my understanding of how

[1] As student responses were reported anonymously, we use numbers to identify them. The numbers refer to the instructor's system of organizing all of the students, including those who did not consent and whose responses were not analyzed, thus the numbers go above 39.

> racial disparities appear in society in more ways than most people realize. I never thought about how skin color could affect readings from different medical devices, yet it is evident from decades of studies that this is a serious issue faced by people of color. It is difficult to comprehend that, with the abundance of revolutionary technology today, that the pulse oximeter has not been adjusted to ensure equality in readings. [Student 10]
>
> I found myself checking my privilege as I read these articles because I realized that these considerations about pulse oxes are things I would never even think about. [Student 14]

A majority of students also expressed surprise and frustration at the slow rate of change–many noted research on this topic was published over 15 years ago. Students also were struck by the fact that solutions have been proposed but not yet implemented.

> The most surprising thing about the information presented in the articles is that the racial disparity of oximeters has been known for over 15 years, but still there has been no major changes to address the differences in blood oxygen readings for both the technology and medical treatment as it relates to those readings [Student 37]
>
> As stated in the first article, engineers at MIT have stated that simple additions and enhancements to devices can easily be made to become for [more] equitable. Why, however, is this not occurring throughout the medical industry? Are there other common medical devices or tests that also result in inequitable results based on race? [Student 10]

In contrast, some students seemed to expect racial bias, and thus were not surprised to read about this specific case:

> Medicine as with many industries has a sordid history with race and racial bias. So it makes sense that the effects would still be around today. It's not good, and I wish people had done more to fix it before we got into a pandemic that interferes with oxygen levels, and we're stuck with devices that inaccurately measure the oxygen levels of people with darker skin. But, here we are. [Student 5]

Considering the diversity of the students in this course, it is not surprising that some expected to find racism in technology while others “had no idea” or “never thought” about the issue, perhaps because these students have internalized the prevailing ethos that data and technology are inherently objective and unproblematic.

Bias

Many students’ reflections included discussion of racial bias, which is not surprising, as “racial bias” was a key theme of both articles, and was in the title of the first article.

<u>Bias Theme 1: Devices can be biased</u>

Some students explicitly used the term “bias” or “racial bias” to describe to the device, contrasting this with the general expectation that devices are unbiased:

> I think it's crazy how a common medical device can be so inaccurate. It is not just race that can affect the results but also anemia, jaundice, poor circulation, and nail polish. You never think to make sure that a piece of equipment does not have a racial bias. [Student 3]
>
> We always think of machines as unbiased, but that's obviously not the case. [Student 5]
>
> If something that is supposed to be providing equal service to all like a medical device is racially biased how can we trust society to do good in the other most crucial aspects? [Student 6]

Other students pointed out issues with the device "not working" or having "flaws," but did not use the term "racial bias." As the articles made it clear that the device generally works well for light-skinned individuals, these responses also seem to be pointing towards racial bias.

> A lot of the time, I think people expect technology and equipment to just work, and they often don't question the circumstances under which they might not work. I would say that this is true for the hospitals and consumers of the oximeter [Student 48]
>
> Technologies that are created with flaws, no matter how small, end up having larger impacts as future technology incorporates the old technologies or is based on them. [Student 4]

<u>Bias Theme 2: Device testing can be biased</u>

Many students brought up device testing, specifically whether enough dark-skinned individuals were part of the test group. Some students linked this potential bias in testing to bias of the device itself:

> I find it so strange that companies won't test more diverse groups of people with their product before they send it out into the world. I imaging [sic] that if the entire population is able to access the device, it would make more sense for the testing groups to display that same diversity [Student 35]
>
> Why doesn't training data match the general population and only reflect a certain group of people? [Student 57]
>
> Eye opening to consider again how many things are created with light-skinned people as the primary users, and how those devices were likely originally or mostly only tested on light-skinned people, therefore putting people of color using the device in danger if the device (especially when for medical purposes) doesn't work properly [Student 51]

Some students made inferences as to why the device was not tested on a diverse enough population, using an assumption that the designers are not diverse (which was not stated in the article):

> These articles are yet another potent example of what happens when the people designing the devices are not diverse. If you don't have a diverse design team there won't be a push

> to test devices on a diverse group of people. [Student 22]
>
> Technology amplifies the bias of those who create it. I'm sure that the creators were mostly, if not all white. [Student 54]

Related to device testing, but not explicitly to the testing sample, one student explicitly took up the notion of acceptable margins of error, and pointed out their concern with "science and statistics" being considered to be "objective sources":

> This sentence really stuck with me: "how to build worlds that don't normalize *any* margin of error that would disproportionately obfuscate patients' vital signs based on the color of their skin." I think with science and statistics, they're often presented as objective sources. When the statistical test results say "this margin of error is acceptable," it may be "acceptable" in the mathematical sense, but that doesn't translate to being acceptable in the real world. For each situation what is an acceptable margin of error changes and is not fixed in stone by some p-value table. [Student 44]

One of the course goals was for students to understand different ways that collected data can be biased, including sampling bias. We were therefore pleased that students attended to those passages in the article. However, the students recommended that the testing sample reflect the population, which is actually consistent with the current guidelines. As the article states, the FDA recommends "including at least two people with "darkly pigmented" skin in a group otherwise 85 percent white" (Moran-Thomas, 2020, para. 27); this is a somewhat misleading phrasing of the FDA recommendation: "Your study should have subjects with a range of skin pigmentations, including at least 2 darkly pigmented subjects or 15% of your subject pool, whichever is larger" (U.S. Food & Drug Administration, 2021, Interpretation and Limitations of Pulse Oximetry section, para. 5). This recommendation would result in a range of skin pigmentation that is likely more diverse than the general population of the United States, which is about 12% Black or African American (2020 U.S. Census Bureau, as cited in Jones et al., 2021). While students are right to focus on the makeup of testing samples, they may be overestimating the percentage of the population that has darkly pigmented skin (which is likely no higher than the population of Black people).

Along the *bias* dimension, students were thinking about the different ways in which bias can come up during the creation and testing of the pulse oximeter, and how this is aligned with historical patterns of medical device development and testing practices.

Differential Impacts

Along the differential impacts dimension, we looked for student responses that pointed to how the racially biased device measurements impacted groups of people differently, in this case, specifically related to skin tone. Note that while the Boston Review article and the research article it draws from were careful to describe skin tones (e.g., "dark-skinned subjects", "intermediate skin tones", "lightly pigmented individuals") and wrote about bias and history in terms of Black and White people, student responses also included terms like "POC" (people of color), which is less precise in this context.

Differential Impacts Theme 1: Impact of small errors on insurance coverage

The first article discusses insurance coverage cutoffs based on oxygen saturation, which can lead to small errors in the pulse oximeter reading having large implications for treatment, which resonated with many students:

> Nonwhite people have to be sicker to get the same treatment that White people get because things like reimbursement and treatment are based on readings [Student 4]
>
> The Medicare threshold is an example of how this racial bias becomes systemic – a Black patient would be statistically less likely to be reimbursed for home oxygen due to this equipment error, turning a technical issue into a systemic advantage [Student 31]
>
> I'm sure that the creators were mostly, if not all white. It's incredibly scary to imagine not knowing how low your saturation really is because the numbers have such a big range of uncertainty for POC. This could also lead to refusal of treatment for someone who needs it because the numbers don't qualify. [Student 54]

Differential Impacts Theme 2: Resulting differences in medical care

Some students took up the article's argument that differences in pulse oximeter readings may result in differences in treatment, including misdiagnoses, for darker skinned patients.

> - oximeters have trouble with dark skin- why can't we fix this now? especially since Black people are at a higher risk for covid complications
> - these errors could have MASSIVE consequences in triage situations, a barrier to equitable care
> - combined with other prejudices, this could quite literally mean the difference between life and death [Student 2]
>
> Poor readings of the pulse ox can lead to mistreatments, delayed recoveries, and worse outcomes. [Student 3]
>
> Not only are patients with darker skin being dismissed more simply because the ox readings are generally inflated/higher than they normally would be with a white patient, but it seems that these individuals are also being misdiagnosed for harmful respiratory conditions, which is even more problematic during the current pandemic [Student 23]

While the article was careful to state that these are *potential* impacts, and that there are not currently conclusive studies on these downstream effects of the pulse oximeter errors, some students did not include qualifiers such as "could" and "can" in their responses, such as Student 23 above. It is unclear whether this student was simply less careful with their wording or if they interpreted the article to mean that these impacts are clearly and directly related to the documented pulse oximeter bias.

Differential Impacts Theme 3: Medical community best serves White people

Some students tied the points about this particular medical device to larger themes of racism in the medical field. Students noted how the medical system has historically positioned White individuals to be the default, leading to inequitable treatment of nonwhite patients:

> [A] lot of medical research historically has been done on fairly health [sic] white males, and this means when diseases manifest differently across genders or races they aren't detected as much. … The point of these articles is to highlight how these issues exist across races, and it is a really big issues [sic] that is resulting in people of color getting sub par medical treatment. Medical racism and the general ways the medical community best serves wealthy, white, men as a whole is a big issue that the medical community need [sic] to take seriously because people are getting sick and dying because of it. [Student 40]

> Having a light-skinned patient as the standard in medicine is inevitably a poor choice even if errors associated are small [Student 28]

> If any other device worked most of the time but was inaccurate for certain cases, I'm pretty sure it would come with all sorts or warning labels, and doctors would definitely know about it. However because this is racism, it seems easier for people to turn a blind eye and say "as long as it works for white people, were [sic] fine." [Student 5]

Within these three themes of the *differential impacts* dimension students showed their understanding of how this one device is embedded in larger systems of racism in medicine. They noted how reimbursement and treatment can be based on specific device readings, how this could lead to differences in care for Black patients, and how this ties into historical patterns in which medicine as an institution was designed for white people.

Responsibility

While students were not explicitly asked about responsibility (the generic prompt asked for their ideas, issues and questions they had, and what they wanted to know more about), many students considered responsibility for the situation in their responses. Students explicitly named three groups in their responses: healthcare providers, the device designers, and the FDA.

<u>Responsibility Theme 1: Healthcare providers</u>

Many students were struck by the author's description of how some of the healthcare providers she talked to were less concerned than she expected about the findings of racial bias in the device. While these doctors and nurses claimed that the pulse oximeter readings were only one data point, and that healthcare providers use many pieces of information when deciding on care, the author was concerned that this stance was based on anecdotes, and found that in her own care, nurses seemed to rely heavily on that one metric. The students generally took the same stance as the author, and were not convinced that these small errors in pulse oximeter readings were negligible in practice.

> My main concern is how complacent the healthcare providers seems about this issue. They brushed off the journalists concerns about racial bias in medical devices saying that there are other factors to look at if you are a trained professional. That might be true to some extent, but using a machine that is known to provide inaccurate readings for certain patients seems like a practice that would have been fixed, or at least had doctors and nurses informed about it. [Student 5]

> The doctors in hospitals claimed to use multiple methods to determine whether or not patients need oxygen, but in reality, they tend to just use the oximeter to make their decisions. [Student 4]

> A recurring theme I observed was how all the medical professionals seem to dismiss this issue, however, based on what I have read from the article, it is a major issue of racial bias that is negatively impacting people of color. It should be addressed because the health and wellbeing of people are directly being affected by this problem! [Student 23]

One student took a different stance, that doctors should *not* rely solely on devices and should be expected to give proper treatment regardless of technological advances:

> The first article talking about the oximeters, was very unsettling. Overall, it seems as if no matter how accurate the machines become, it is up to medical personnel to give proper treatment to darker-skinned patients. [Student 60]

These students noted their concern with how some healthcare providers downplayed the risks from the racial bias in the pulse oximeter readings. However, these students do not blame the doctors or nurses for the situation, or for using devices which are now known to give biased results in some situations.

<u>Responsibility Theme 2: Engineers/designers</u>

A number of students specifically called out the engineers/designers/creators (different students used different terms) of the pulse oximeter as responsible for creating a device without racial bias:

> I think scientists and engineers should make it their priority to make sure that medical devices work for everyone before releasing a device that can put [sic] potentially put someone in danger. [Student 49]

> Everyone has put their trust in the developers who made the technology (rightfully so), and doesn't question whether or not it works all the time. While some might say that it's just an unfortunate circumstance that the type of technology used is not as accurate with darker skin, the problem is the creators of the oximeter didn't think about the implications of their technology and probably didn't test it on a diverse enough group of people. A better mindset might be: why didn't they use a different type of technology to measure oxygen levels? [Student 48]

One student argued that responsibility to do better in the future rests with the next generation of designers and biotech companies:

> I think that covid has brought up a lot of racial biases in the biotech industry and even in the over health system and it is up to the new generations and current major biotechs to start defying this [sic] standards [Student 6]

In contrast to the reflections discussing healthcare providers, where students generally seemed to be pointing out concerns but not placing blame, many students placed blame for the device's racial bias on engineers/designers/creators of the technology. It is not clear if students believe the same people are in charge of designing, testing, and releasing a device for public use.

Responsibility Theme 3: FDA

Aligned with the stance of the authors of the second article, the vast majority of students expressed criticism of the FDA for not explicitly linking the "limitations" of the pulse oximeter to racial bias in the Safety Communication.

> The second article says that the FDA's release 'does not specifically mention racial disparities in the accuracy of the devices', which is both dangerous and upsetting. It's dangerous because it is a *major* issue with pulse oximeters. Without explicitly stating which users are more likely to have inaccurate readings, the alert doesn't really help anyone. It's upsetting because the first step toward change is admitting and [sic] problem, and if the FDA can't admit that racial bias can change the accuracy of the readings, then there's no indication that change might occur… In order for change to occur, the problem must be recognized. The FDA tried to side-step the issue, and that's unacceptable. [Student 31]

> I think it's also wrong that the FDA didn't implicitly [sic] say that the limitations of the pulse oximeter are the racial bias behind the inner workings of the device because by not mentioning it, they're not bringing attention to the issue. [Student 49]

> Although the FDA issued alert on 'limitations' of pulse oximeters it still does not solve the issue, They didn't even mention the word race in their statement. [Student 3]

One student brought up questions about possible consequences of FDA wording and whether a more explicit warning tied to racial bias would have the desired outcome:

> Why did the FDA avoid the words race or racial in its communication, even though there is evidence that oximeters discriminate against patients with darker skin color? What would be the consequences/impacts if they did acknowledge racism as an issue here? Would positive changes be made, or would there potentially be backlash from manufacturers? [Student 44]

While no students blamed the FDA for creating this situation of widely-used pulse oximeters giving racially biased readings, nearly every student who mentioned the FDA believed the FDA does have a responsibility for labeling the errors in the devices as linked to race. Most students

believe that a necessary first step in fixing this inequity is to be clear that the studies found clear racial bias in the data.

Within these three themes of the *responsibility* dimension, students are thinking about the different actors involved around the design and use of the pulse oximeter and how their different roles play an impact on the consumers of the product.

Discussion

As this study data was gathered during the pilot study and initial year of adding sociotechnical content to a previously solely-technical computing course, the research team was unsure how students would take up these sociotechnical content. Would students already be aware of these issues, from their high school educations or from the news? Would they readily accept the many examples of inequities in technology and engineering products? Would they push back on these ideas, sticking to technocratic ideologies that technology is inevitable and the right way to solve problems? Unsurprisingly, we found a range of responses in the diverse collection of mostly first-year, mostly engineering students.

In this specific case of the topic of racism in medical equipment, we found that many students were surprised to hear of racially biased readings from a widely-used medical technology, while others expected to find racism in most (all?) aspects of society. Overall, students were able to grasp the main sociotechnical points of the articles, and many comments showed students understood the systemic nature of racism in medicine.

We found that students generally did not argue with the points of these particular articles, in contrast to readings on other sociotechnical topics in this course. They did not problematize the author's arguments and instead tended to repeat the article's claims and voice their alignment. Some students put the article's findings in personal context.

Most students seemed to understand the limitations of the study cited in the Boston Review article (Moran-Thomas, 2020): the study was conducted in a laboratory setting with healthy volunteers, and found racial bias in the device's outputs–there was a clear pattern of great bias for dark-skinned subjects and least for the lightest skin tones, and this bias was worse at lower oxygen levels. There is no reference to studies on the implications of these findings, although the author spells out how this documented bias could lead to (and have already have been a factor in) inappropriate medical care for Black patients. In some students' responses, it was unclear whether they saw this distinction, or if they believed that Black people have had worse outcomes with Covid-19 as a direct result of errors in pulse oximeters, which is plausible, but certainly not a claim of the article.

Overall, students felt that the engineers or designers creating these devices have a huge role to play in making sure medical technology is equitable. They seem to believe those conducting testing have a responsibility to test new devices on a wide range of individuals, going beyond what is required by the FDA. Students also showed that they are beginning to understand how regulatory agencies, such as the FDA, can have a strong impact on technology.

While students generally seemed to understand the main arguments of the articles, they did not show that they understood just how hard it is in the real world to solve a problem of this magnitude. Students were dismayed that the problem has been known for over a decade and that these devices are still being used. Students do not seem to have an understanding of how long it takes to develop, test, and get approval for new medical devices, or how much it would cost and how long it would take for a single hospital to replace every pulse oximeter with a new, less biased model (such as one proposed by MIT researchers and mentioned in the article, that is more complex and would cost more than existing models). This lack of knowledge is entirely reasonable, as many of the students in this course are in their late teens and have little experience with these processes.

In the year following the collection of this pilot study data, the course designers chose different articles for students to read: the Boston Review article was replaced by an article that more clearly spelled out the findings from the research study that showed racially biased results, but, in a trade-off, did not place the findings as clearly in the sociohistorical context. For the second reading, students read an abbreviated version (unnecessary sections were removed by the course designers) of the actual FDA guidelines for testing pulse oximeters. Specific reading response questions highlighted for students the invasive and risky nature of the test required to verify new devices, so that they could see that the solution is not as simple as testing on many more people, and particularly more Black people. The goal of these changes was to make it more clear to students just how complex it is to address a problem of this magnitude, even in situations where racial bias is clearly leading to worse outcomes, which is not necessarily true in this situation.

Conclusion

In this work-in-progress paper, from a pilot study interested in the impact of including sociotechnical topics in a first-year introduction to computing course for engineering students, we analyze student written reflection responses to a pair of popular news articles focused on findings of racially biased readings of pulse oximeters. We were interested in how students interpreted and made sense of these articles, including how they related the information to what they knew or expected of the real-world impacts of technology. We unpacked three dimensions of students' developing sociotechnical literacy: bias, differential impacts, and responsibility.

We found that first year students were thinking about multiple, complementary aspects of the impacts of the medical device. They considered how neutral-seeming technology can be biased, likely as a result of testing the device on mostly light-skinned individuals during development. Students recognized the differential impacts that resulted from this biased device within the ecosystem of healthcare and insurance reimbursements, and some students noticed how these disparate impacts aligned with historical patterns of technology privileging white individuals. Students started to attribute responsibility to different actors, most strongly the designers, and also criticized the FDA for issuing an alert that did not explicitly call out racial bias.

In future iterations of this course, we hope that students will also begin to appreciate just how difficult these problems are to solve. Students who are in their late teens will understandably find it hard to believe that a problem that was identified over a decade ago has not yet been solved.

As part of developing students' sociotechnical literacy, they will need to appreciate how difficult change is, and to understand the many trade-offs that are always a part of sociotechnical systems.

In future research, to get a deeper understanding of student reasoning, we will investigate other data sources, including video recordings of in-class small group and whole-class discussions on sociotechnical topics. We will also look at student written responses over the entire semester as students consider a range of current real-world application examples, to examine changes in students' sociotechnical literacy skills (McGowan & Bell, 2020).

Acknowledgements
This material is based upon work supported by the National Science Foundation, IUSE-2110727. Any opinions, findings, and conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We would like to thank Dr. Ethan Danahy and Dr. Jenn Cross for their course design work and overall support of this project.

References

Benjamin, R. (2019, October 17). *Ruha Benjamin on "The New Jim Code? Race, carceral technoscience, and liberatory imagination"* [Speech transcript]. Other & Belonging Institute at the University of California, Berkeley. https://belonging.berkeley.edu/video-ruha-benjamin-new-jim-code-race-carceral-technoscience-and-liberatory-imagination

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Brodwin, E., & St. Fleur, N. (2021, February 19). FDA issues alert on 'limitations' of pulse oximeters, without explicit mention of racial bias. *STAT News*. https://www.statnews.com/2021/02/19/fda-issues-alert-on-limitations-of-pulse-oximeters-without-explicit-mention-of-racial-bi/

Buolamwini, J., & Raji, I. D. (2019). Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. In Proceedings of the 2019 Conference on Artificial Intelligence, Ethics, and Society.

Cech, E. A. (2014). Culture of disengagement in engineering education? *Science, Technology, & Human Values*, *39*(1), 42–72. https://doi.org/10.1177/0162243913504305

Claussen, S. A., Tsai, J., Boll, A., Blacklock, J., & Johnson, K. (2019). Pain and gain: Barriers and opportunities for integrating sociotechnical thinking into diverse engineering courses. In Proceedings of the 2019 ASEE Annual Conference and Exposition.

Costanza-Chock, S. (2020). *Design justice: Community-led practices to build the worlds we need*. The MIT Press. Cambridge, MA.

Creswell, J. W., & Poth, C. N. (2018). Qualitative inquiry and research design: Choosing among five approaches. SAGE Publications.

Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education*, *32*(1), 268-291.

Faulkner, W. (2007). 'Nuts and bolts and people': Gender-troubled engineering identities. *Social Studies of Science*, *37*(3), 331-356. https://doi.org/10.1177/0306312706072175

Feiner, J. R., Severinghaus, J. W., & Bickler, P. E. (2007). Dark skin decreases the accuracy of pulse oximeters at low oxygen saturation: The effects of oximeter probe type and gender. *Anesthesia & Analgesia*, *105*(6), S18-S23. https://doi.org/10.1213/01.ane.0000285988.35174.d9

Feng, A., & Wu, S. (2019, May 1). The myth of the impartial machine. Parametric Press, Issue 1. Retrieved from https://parametric.press/issue-01/the-myth-of-the-impartial-machine/

Gupta, A. (2017). A practitioner account of integrating macro-ethics discussion in an engineering design class. *Proceedings of the American Society of Engineering Education Annual Conference*, Columbus, Ohio.

Herkert, J. R. (2005). Ways of thinking about and teaching ethical problem solving: Microethics and macroethics in engineering. *Science and Engineering Ethics*, *11*(3), 373-385. https://doi.org/10.1007/s11948-005-0006-3

Hess, J. L., & Fore, G. A systematic literature review of US engineering ethics interventions. *Science and Engineering Ethics*, *24*(2), 551-583. https://doi.org/10.1007/s11948-017-9910-6

JafariNaimi, N. (2018). Our bodies in the trolley's path, or why self-driving cars must *not* be programmed to kill. *Science, Technology, & Human Values*, *43*(2), 302-323. https://doi.org/10.1177/0162243917718942

Johnson, K., Leydens, J., Walter, J., Boll, A., Claussen, S., & Moskal, B. (2019). Sociotechnical habits of mind: Initial survey results and their formative impact on sociotechnical teaching and learning. In *Proceedings for the American Society for Engineering Education Annual Conference*, Tampa, Florida.

Jones, N., Marks, R., Ramirez, R., & Ríos-Vargas, M. (2021). *2020 Census illuminates racial and ethnic composition of the country*. United States Census Bureau. https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html#:~:text=In%202020%2C%20the%20Black%20or,million%20and%2012.6%25%20in%202010

Leydens, J. A., & Lucena, J. C. (2017). *Engineering justice: Transforming engineering education and practice*. John Wiley & Sons. https://doi.org/10.1002/9781118757369

Manz, E. (2016). Examining evidence construction as the transformation of the material world into community knowledge. *Journal of Research in Science Teaching*, *53*(7), 1113-1140. https://doi.org/10.1002/tea.21264

McGowan, V. C., & Bell, P. (2020). Engineering education as the development of critical sociotechnical literacy. *Science & Education*, *29*(4), 981–1005. https://doi.org/10.1007/s11191-020-00151-5

Moran-Thomas, A. (2020, August 2). How a popular medical device encodes racial bias. *Boston Review*. https://bostonreview.net/articles/amy-moran-thomas-pulse-oximeter/

Pawley, A. L. (2019). "Asking questions, we walk": How should engineering education address equity, the climate crisis, and its own moral infrastructure? *Advances in Engineering Education*, *7*(3), 1–11. https://advances.asee.org/wpcontent/uploads/vol07/issue03/Papers/AEE-25-Opinion-Pawley.pdf

Riley, D., & Claris, L. (2003, June). Pedagogies of liberation in an engineering thermodynamics class. In Proceedings of the 2003 American Society for Engineering Education Annual Conference & Exposition, Nashville, Tennessee. https://peer.asee.org/11854

Riley, D. (2008). Engineering and social justice. *Synthesis Lectures on Engineers, Technology, and Society*, *3*(1), 1-152. https://doi.org/10.2200/S00117ED1V01Y200805ETS007

Salzman, N., Winiecki, D., & Jain, A. (2019, October). Assessing community in an undergraduate computer science program using social network analysis. In *2019 IEEE Frontiers in Education Conference (FIE)* (pp. 1-4). IEEE.

Trevelyan, J., P. (2014). *The making of an expert engineer: how to have a wonderful career creating a better world and spending lots of money belonging to other people*. CRC Press.

U.S. Food & Drug Administration. (2021, February 19). *Pulse oximeter accuracy and limitations: FDA safety communication*. https://www.fda.gov/medical-devices/safety-communications/pulse-oximeter-accuracy-and-limitations-fda-safety-communication