International Journal of Engineering Sciences Paradigms and Researches (Volume 47, Issue: Special Issue of January 2018)
ISSN (Online): 2319-6564 and Website: www.ijesonline.com

An investigation of the big data ethics of aspiring mathematics teachers with a focus on data access

1.Mrs Soubhagini Mohapatra 2.Dr. Meenakhsi mohanty
Nalanda Institute of Technology, Bhubaneswar
Dept. of Basic Science & Humanities
E-mail ID: soubhagini@thenalanda.com
MeenakshiMohanty@thenalanda.onmicrosoft.com

A R T I C L E I N F O

Keywords:
Teacher education
Ethics
Big data analysis
Critical thinking
Social sustainability
Access to data

A B S T R A C T

Big Data analysis applications are changing society. New efforts should be made in mathematics education to teach the new technology and to provide a critical and ethical understanding of its consequences in light of the creative ways that mathematics is used in society. This interview study explores the ethical considerations of pre-service teachers in data science environments, concentrating on issues related to access to the data that underpin the technology. Results demonstrate that pre-service teachers present a wide range of ethical justifications for data access that guide their efforts to think critically about oppressive circumstances. Yet, there is also evidence that their ability to reason may be constrained by a lack of knowledge of the relevant data science approach, suggesting that more of this should be covered in mathematics teacher preparation..

The rise of Big Data Analytics (BDA) and related methodologies, have collectively been labeled the fourth industrial revolution. BDA is a new phenomenon that sits at the intersection of statistics, computer science, and domain knowledge. Using advanced statistical and computer science methods, data scientists are able to extract stories about human behavior and world phenomena to predict the future (Matthews, 2019). Notably, the ability to predict human behavior has reshaped the world economy into a data driven one in which leading entities are those that have the financial, technological, and human resources to be proactive instead of reactive (Agrawal, 2017). However, as analogous to previous industrial revolutions, rapid societal changes have led to negative ramifications for marginalized groups, including immigrants, the impoverished (O'Neil, 2016), women and non-binary gendered peoples (D'Ignazio & Klein, 2020), and people of color (Sabbagh, 2020). Relating to the theme of the special issue, *Mathematics in Society: Exploring the mathematics that underpins social issues*, the perspective taken in this study is grounded in the necessity to understand the potentially harmful and liberatory aspects of the mathematical components of this 4th industrial revolution in order to fully reap the benefits of the new technologies and BDA methodology. Since students will become both consumers and producers of these new technologies, we identify mathematics education as an arena for making connections between ethics, critical thinking and the mathematics of BDA. Thus, pre- and in-service teachers' understanding of such connections constitute important factors for consideration and connects to the question raised in the call for papers for the special issue, *how can research in mathematics education support a more robust and critical understanding of the world?* As such, this study seeks to investigate pre-service teachers' (PSTs) ethical and mathematical reasoning in the context of BDA, with particular emphasis on their considerations of *accessibility* to data and algorithms.

This study serves both as a phase of a collaborative, Design Based Research project initiated by the Charlotte Ethical Mathematics

Design Team (C-EMDT) that seeks to understand how to design curricula for Critical Mathematics Consciousness (CMC) and ethical reasoning in the context of BDA, and as a point of departure in a PhD project situated in Malmö, Sweden centered around fostering democracy through critical discourses in mathematics education. Our aim for this article is to investigate PSTs perspectives on access to data in relation to their discussions of ethical and mathematical considerations in BDA. The research questions that guide this study, and reflect different aspects of the aim, are:

1. How do preservice teachers reason ethically and mathematically about access to data in data science contexts?
2. How does preservice teachers' demonstrated Critical Mathematics Consciousness in data science contexts relate to their reasoning about access to data?
3. How might the *accessibility* construct be refined to reflect common dispositions related to access to data?

1. Literature review

In the sections that follow, we will briefly describe the literature related to discrimination in machine learning, access to data in machine learning, and research-based arguments regarding the need for ethical mathematics education in the Age of Information. The literature review intends to illuminate the role of data in the data driven practices of BDA, how existing injustices can be reproduced by the new technology, and how ethical aspects can become pivotal when teaching and learning about mathematical aspects of the new technology. The literature review ends with identifying a gap in the related research that corresponds to the focus of this study on perspectives of access to data.

Discrimination and access to data in machine learning

The mathematical algorithms used for predictive models in BDA are developed on training data and tested for accuracy on a testdata set. Data scientists typically choose the algorithm that predicts with the highest accuracy. Importantly, if the training data is biased towards particular groups, then the algorithm itself will produce biased outcomes.

Well known data sources used in BDA include a digital compilation of users' behaviors and characteristics. Such "digital traces"generate a user profile constituted by personal attributes and predicted traits in the absence of desired data. As a result, the demand for more data to improve convenience and efficiency has led to companies' increasing surveillance of peoples' behaviors (Zuboff, 2019). Today, organizations are able to access trends in our movement through mobile phone locations, internet traffic on company websites and third party cookies that track internet browsing. In addition, the data collected often goes beyond its initial purpose, as exemplified by the court conviction of the Swedish Police authority for gathering ethnicity data (which is sensitive and protected data in Sweden)in ways that went beyond the purpose of police work (Renfors & Justitieombudsmannen, 2015). Because of this, some scholars have stressed that *access* to such personal information can make individuals and groups more vulnerable (O'Neil, 2016; Zuboff, 2019), while others have countered that the precision in the statistical predictions are too low to pinpoint and exploit specific individual's char-acteristics and that only characteristics on a group level are predictable to any meaningful degree (Sumpter, 2018). Despite arguments about precision, the collection of sensitive and/or behavioral information on individuals (e.g. race, ethnicity, gender, etc.), whetheraccurate or not, has been shown to yield predatory and discriminatory ramifications for some groups (O'Neil, 2016).

An additional concern relates to the ability of machine learning methods to predict characteristics of individuals based on their available data. That is, if sensitive data, such as race or gender, is omitted from the training data, the algorithms can still adopt ste- reotypes from proxies in the data (O'Neil, 2016). To clarify, information is contained in that data that can be connected to the de- mographic characteristics of the person, regardless of whether or not those characteristics are explicit. For example, a person's zip code can be connected to their neighborhood of residence, and potentially their race or ethnicity. This constitutes a problem for democracy when such algorithms are used to make important automated decisions for citizens (Villani et al., 2018), creating tensions between different perspectives of fairness (e.g., the right not to be evaluated on the basis of one's race on the one hand, and not being discriminated against because of racially biased data on the other, (Corbett-Davies et al., 2017). Naturally, this begs consideration of the ethical implications of using personal data when making decisions in society. As a result, we argue an approach to mathematics education in which students are positioned to consider the ethical implications of their mathematical work on individuals and social groups. The literature related to this argument is discussed next.

Impetus for an ethical mathematics education

The rapid growth in our capacity to analyze data has catalyzed a shift towards data science in the job market and in the amount of data literacy required by the general public. As it stands, the BDA industry is populated by data scientists who come from predominantly privileged backgrounds (white and/or male; upper income) resulting in what D'Ignazio and Klein (2020) term a *privilege hazard*. That is, phenomena in which teams of data scientists are primarily composed of people from dominant groups in privileged positions. Designs created in these contexts often reflect the dominant perspectives, experiences and values of the privileged creators at the expense of nondominant identities and viewpoints (D'Ignazio & Klein, 2020; Noble, 2018). Having limited or no experience with social and/or financial struggle or the lived-experiences of the masses, privileged data scientists are often ill-equipped to identify oppressive situations in the world resulting in a lack of consideration for the impact of their models on societal groups (D'Ignazio & Klein, 2020). Consequently, a salient threat to democratic society today is the *hard coding of discrimination* in the processes used by world governing

entities (D'Ignazio & Klein, 2020).

To combat the negative effects of BDA and the privilege hazard, there is an increasing amount of research that adopts a socio-political perspective on BDA congruent with the aim of this study. Such research stresses the relevance of power in relation to data ineducation (Rubel et al., 2021), proposing that power and responsibility may be hidden when people interact with mathematics,statistics, and technology (Straehler-Pohl, 2007). To clarify, the algorithms used in BDA to make decisions about users are eitherhidden behind a user interface and/or the code is uninterpretable by the common citizen, making it impossible for the user to critiquethe variables, methods, and/or claims made by the organization (O'Neil, 2016). This phenomenon is further reflected in mathematicscurricula where, for instance, the societal and democratic aims for mathematics education are expressed in general terms and do notcall for any of the specific mathematical understandings required to perceive how power can operate in BDA (Andersson, et al., 2022).

This article draws on an emerging body of research that calls for the explicit grounding of mathematics education in ethics (Atweh & Brady, 2009; Boylan, 2016; Ernest, 2018). Such scholars view mathematics as undeniably integrated into the technical, political, industrial, military, social, etc., facets of the world having real effects on individuals and groups in society (D'Ambrosio, 1998; Lengnink, 2005). Further, given the mathematical nature of data science and its prevalence in the social, economic, and politicalaspects of globalized society, it has become imperative that students become proficient in its techniques and algorithmic ways of thinking (Boaler & Levitt, 2019; Cobb, 1999; O'Neil, 2016; PISA, 2020). Initiatives to develop a standardized K-12 data science curriculum are taking shape across the globe. These include efforts to transform the current mathematics curriculum, or provide for the eventual shift to more relevant, modern schooling (Boaler & Levitt, 2019; Koh, 2020; Tong & Yong, 2015). K-12 data science programs, however, have yet to fully integrate ethics and social justice into their coursework. They typically focus on developing students' foundational understanding of the data life cycle and the basic statistical and/or computer science concepts necessary for navigating it (Gould et al., 2016; Heinemann et al., 2018; Tong & Yong, 2015). As such, we adopt methods from Critical Mathematics Education (CME) to guide our research in developing ethically grounded data scientists and citizens. However, there is a shortfall of research in CME related to peoples' personal data, their perspectives on access to data, and their consideration of, and responsibility toward ecologies and the other (Boylan, 2016; Andersson et al., 2022; Register et al., accepted). Given the expansive reach of BDA, its implications for marginalized groups, and its overall effect on society, we find it imperative to expand the scope of CME and to address the aforementioned gaps.

2. Theoretical orientation

The essence of what we consider an ethical mathematics education is tied to the *transformative ideal* for mathematical education. That is, we see mathematics education as a means to "create the world in a new way" (Atweh & Brady, 2009, p. 270), going beyond *responding* to injustice (typical of CME and Teaching Math for Social Justice) and focusing on responsible *creation*. Drawing from relational ethics, Boylan (2016) proposes four ethical dimensions that mediate the learning of mathematics and include relationships with *others*, the *societal and cultural*, the *ecological*, and the ethical *self* (Boylan, 2016). Notably, in Western cultures, considerations of the other are typically subordinated to neoliberal and individualistic attitudes concerned with individualism, power, success, and self-preservation (Wiggan, 2012). In contrast, Boylan (2016) suggests that ethical reasoning concerns issues of fairness and choice, but also must consider relationships outside of those that exist within our own communities and beyond human interactions. Thus, mathematics education should also develop a sense of *social response-ability* that includes "the ability to respond to the demands of our own well-being and the ability to respond to the demands of the other" in our current and future lives (Atweh & Brady, 2009, p. 269). Considering the ethical implications of the mathematically grounded sciences in the Digital Age, it is essential that mathematics is learned in the context of *ethical design* where students are encouraged to consider the micro and macro-level implications of their mathematical products.

Critical mathematics education (CME)

Given its explicit focus on identifying and dismantling oppressive relations and power structures, ethical mathematics education adopts the core tenets of Critical Mathematics Education (CME) as a basis for promoting socially response-able mathematicians and data scientists (Atweh & Brady, 2009). Skovsmose (1994) identified several core tenets of CME that we see as essential to the development of ethical mathematics education. Namely, schooling should serve as preparation for active participation in political life. Thus, mathematics education should center applied mathematics as it is used in the increasingly globalized world. Further, mathematics should be promoted as a tool for exploring and analyzing critical societal features, but that can also be used in unethical ways. Considering the potentially problematic function of mathematics in modern technology and society, mathematics itself should be rejected as objective and neutral since humans employ mathematic and should be critiqued with regard to the reproduction of societal inequalities and discrimination. Regarding these concerns, CME scholars (Gutiérrez, 2013; Gustein, 2006; Frankenstein, 1983; Rubel et al., 2016; Skovsmose, 1994) argue for a contextualized mathematics curriculum which is based on students' macro and micro-level realities for the purposes of developing a critical consciousness of, and empowerment to dismantle, the oppressive forces in one's life.

Critique of CME

Boylan (2016) suggests that within CME there is a lack of discourse related to consideration of the *other* as an ethical imperative for mathematics. Critical Mathematics pedagogies typically focus on liberation of the self or that person's social group. They adopt Freire's (1970)/(2018) viewpoint that people in oppressed positions cannot be set free by their oppressors, but must develop the literacy, reflection, and agency to free themselves. Such studies typically focus on groups of racially and/or ethnically homogenous and

historically marginalized populations of students who explore social justice issues relevant to their community through mathematics (Berry, 2003, 2004; Gutstein, 2006; Rubel et al., 2016; Rubel, 2017). However, while Freire certainly opposed the concept of teachers and/or oppressors as saviors, he suggested that to dismantle oppressive systems, the oppressed and the oppressors must *work together to* transform reality. That is, the oppressors must also free themselves from their oppressive role through their own development of critical consciousness. With that being said, Freire's work typically conceptualized oppression on the basis of economic class, receiving criticism for his lack of explicit attention to race and gender (Gutstein, 2006; Hooks, 2014). The present lens considers oppression in terms of unearned societal advantages based on a diversity of privileges including but not limited to race, gender, socioeconomic status, dis/ability status, education, and their complex intersections (Crenshaw, 1991).

The outlined ethical dimensions of mathematics illuminate gaps in mathematics education as it currently stands. While attempts to integrate the societal and cultural dimension into mathematics education have been gaining traction (Berry, 2003, 2004; Esmonde, 2014; Gutstein, 2006; Kokka, 2020; Rubel et al., 2016; Rubel, 2017), significant work needs to be done to promote students' consideration of the other, the ethical self, and the ecological impact of their mathematical products, and to understand how such reasoning can be supported in classrooms. For this work to be effective, it is necessary to understand the diverse ethical orientations of those who will implement such curricula (e.g., pre-service teachers) and the characteristics of ethical reasoning that promote the consideration of the ethical dimensions of mathematics (Boylan, 2016). Thus, the results of this study will serve to inform future frameworks used for analysis of individuals' ethical reasoning in mathematics and data science, for making curricular design decisions, and for preparing teachers to implement such curricula. As a starting point, the following CMC analytic framework is used to understand how individuals think critically in mathematics and data science contexts which have micro and macro-level ethical implications.

Critical mathematics consciousness (CMC)

The development of ethical mathematicians and data scientists may be aided by their development of *critical mathematics consciousness* (CMC) (Stephan et al., 2021). Derived from Freire's conception of *critical consciousness*, the definition for CMC used in this research attempts to critically analyze the mathematical action and/or product from an *ethical* and *sociopolitical* perspective, and refers to.

the awareness that human beings do mathematics; thus, there are potential ethical dilemmas and implications of mathematical work which may affect entities at the individual, group, societal, and/or environmental level. CMC includes *sociopolitical*, *ecological*, and *communicative* mathematical awarenesses (Table 1) *and a willingness and commitment to act* (i.e., critical mathematics agency) (Register et al., 2021; Stephan et al., 2021).

Analytic frameworks

Assessing CMC

Importantly, people may exhibit different levels of consciousness at different times and in different contexts (Stephan et al., 2021). Their consciousness is not a part of their personality but is fluid, dynamic, and tied to the context of the issue and their relevant experiences. Within a given context, an individual's reasoning can be classified as one of six levels of consciousness (see Fig. 1. CMC Growth Framework).

Shor (1993) defines *critical transitivity (CT)* as the actualization of full critical consciousness where the individual engages in critical reflection, recognizes the systemic influence on oppression, and is empowered to act by attacking the causes of oppression at the systems level. Freirean theory maintains this as the goal of a liberatory education. *Semi-transitivity* includes critical thought on the ethical and critical aspects, may see the systemic cause of the oppression, but does not attack it at its root. Rather, they may perceive and attack the cause as isolated incidents, *isolated semi-transitive (IsT)* or put their faith in other individuals or groups to change oppressive situations, *systemic semi-transitive (SsT)*. Individuals who demonstrate *disempowered (Di)* consciousness exhibit some critical thought, but feel restricted in their ability to act, likely due to a lack of power. Similarly, those who exhibit intransitive consciousness,feel disempowered to act, but do not exhibit critical thought, perceiving their oppressive situation as a consequence of God's will or bad luck, rendering their agency as irrelevant. Finally, Stephan et al. (2021) incorporate King (1991) construct of *dysconsciousness (Dy)* referring to a lack of critical thought caused by a "distorted vision of oppression" where inequity is the "natural order of the world"(p.3). They believe that the oppressed are at fault for their current situation and are thus solely responsible for overcoming their oppression. As King explains, dysconsciousness may be taught through familial and educational influences, or by media effects.

Ethical reasoning in mathematics [ERiM] Principles

Stephan et al. (2021) and Register et al. (2021) argue that high levels of CMC imply an awareness of ethical implications to be

Table 1
CMC Awarenesses.

Sociopolitical	Mathematics is used to model and interpret the real world and can be used to make decisions (both at the individual and systemic levels) that may further disenfranchise (or liberate) marginalized groups.
Ecological	Mathematics has been socially constructed by human beings and thus has implications for humans, animals, the environment, and its interconnected ecologies.
Communicative	Mathematical communication has the power to educate and mis-educate society and encourage the masses to act in certain ways.

Source: (Register et al., 2021; Stephan et al., 2021).

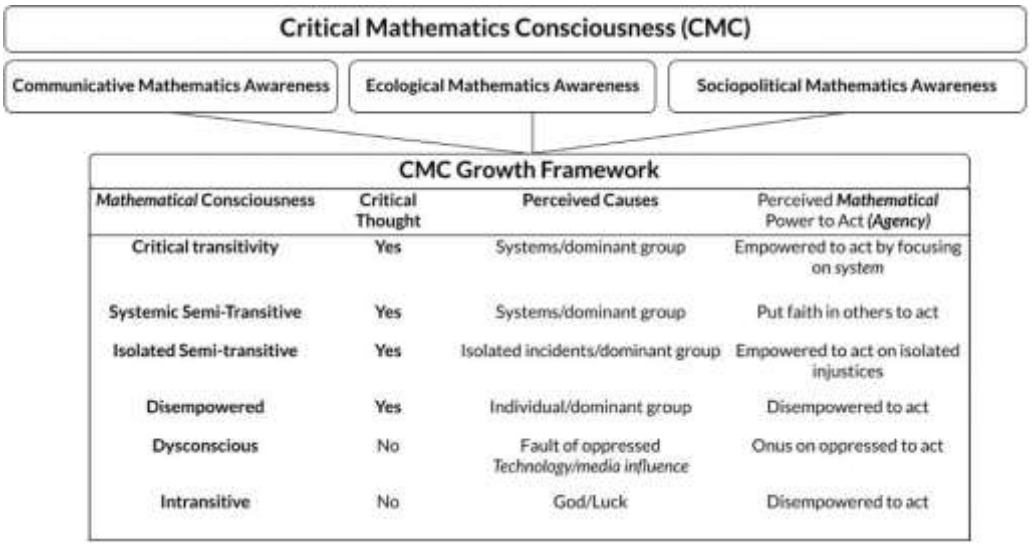

Fig. 1. CMC Growth Framework.

considered in the process of making mathematical decisions and a sense of personal responsibility to do right by the affected parties. As such, the Ethical Reasoning in Mathematics [ERiM] Principles Framework is used to identify the specific ethical considerations thatindividuals' make in order to help classify their demonstrated levels of CMC in specific contexts (Stephan et al., 2021). Unlike CMC, ERiM does not identify individuals as having an exhibited level of consciousness, but rather identifies specific categories of ethical principles involved in data science contexts as identified by discipline experts. For instance, the ERiM framework contains ethical principles that speak to considerations made in the process of making data-based decisions. These include violations of privacy (privacy), fair access to benefits (fairness), the accuracy of data and/or predictive algorithms (accuracy), who is accountable for theeffects of these algorithms (accountability), whose property is the data to sell (property), whether decisions made are loyal to specific individuals, groups, or entities (loyalty), whether algorithms are biased or objective (algorithm bias), if the algorithms are readily available for inspection (transparency), what ecological impact they may have (ecological), how they effect employment (employ- ment), whether they are discriminatory (discrimination), and who has (or should have) access to the data, algorithm, or findings (accessibility). By identifying contexts in which individuals' express concern for the ethical impact of their mathematics products and decisions, we anticipate that the ERiM Principles may support the design of curricular materials that foster their development of critical consciousness. For this study, however, we explicitly focused on the accessibility element of the ERiM framework in hopes ofspeaking to the forms of reasoning that PSTs demonstrate with regard to who should have access to personal information, and in what forms.

Access to data

Accessibility is an element of the ERiM framework that we explicitly focused on in the development of the task-based interview questions in the study and in our analysis of PSTs reasoning. In the analysis, a new construct emerged which allowed us to betterclassify PSTs considerations of access to data. To encompass two perspectives on the role of data in PSTs answers, we consulted literature on BDA in oppressive situations and formulated two related concerns, *dearth* and *excess*, collectively called access to Data (AtD):

- *Dearth*: The problem can be understood as being caused or amplified by limited possession of reliable and valid data, e.g. too few data points, biased data, or data that is impossible to analyze (Williams et al., 2018).
- *Excess:* The problem can be understood as being caused or amplified by an *excess* of reliable data (or what is portrayed as reliable data) possessed by a potential adversary (Darragh, 2021; Zuboff, 2019).

Importantly, *dearth* and *excess* are not to be understood as mutually exclusive. It is possible to construe the role of data as satisfying both perspectives, for example O'Neil (2016) identifies the dangers of an adversary having too much data and using algorithms that are perceived to have the capability to predict risk and performance (*excess*), but simultaneously lacking sufficient data for these predictions to be fair for all groups (*dearth*). Like CMC, AtD does not necessarily represent an individual's general feelings towards access to data, but rather their reasoning in that specific context.

Framework interactions

The discussed analytic frameworks are used reflexively to identify an individual's level of CMC within a given context. Recall that our goal for mathematics education is to foster high levels of CMC in our students and teachers. Thus, it is important to understand

what constitutes and motivates high levels of CMC. When used together, the ERiM framework (including AtD) allows us to identify the ethical considerations that PSTs make in their reasoning, which speaks to the mathematical awarenesses that they possess in that context (potentially motivating the decisions they make). However, being aware that there is an issue does not necessarily imply recognition of the systemic factors that cause or uphold injustice, or the personal agency to dismantle such systems. Therefore, the CMC growth framework helps us to characterize these essential components, speaking to their CMC.

3. Methodology

This study employs a Design Based Research (DBR) approach which implies the development of learning theories that either do not exist or are in need of reform (Bakker & Van Eerde, 2014). Such theories speak to both the process and means of supporting learning in dynamic and diverse classrooms (Cobb et al., 2003). The findings of the wider DBR project at UNC Charlotte have illuminated that secondary students demonstrate a range of ethical concerns in their reasoning related to BDA which may serve as a starting point for the development of curricula intended to foster students' CMC in data-based contexts (Register et al., 2021; Stephan et al., 2021). Given that PSTs are positioned at the forefront of education in the digital age, we chose to focus our study on PSTs' ethical reasoning and CMC in hopes to provide more specific information about the preparedness of the future generation of teachers to support the emergence of ethical mathematicians, citizens and data scientists. As aligned with the tenets of CME, such teacher preparation includes the deconstruction of former identities to redevelop teacher identities that support diverse and innovative minds. This involves teachers' examination of privilege, personal biases, the history of opposing cultures and belief systems, and the development of political knowledge of both themselves and their prospective students (Gutiérrez, 2013). It also includes the development of a consistently reflective, adaptive, and dialogical practice that accommodates the changing world and needs of their students. Thus, our goal is to provide a means to support teachers in becoming revolutionary leaders, intent on "unveiling reality" and "coming to know it critically" while discovering themselves as "its permanent recreators" who are committed to participating alongside their students (Freire, 1970/2018, p. 69). Finally, we chose to focus on *accessibility* due its permeating influence on society and the lack of literature related to issues of access to data in mathematics education.

Author positionality

The authors of this paper serve as doctoral students at the PSTs' respective Universities in the Mathematics Education program. The first author grew up in a lower class, social-democratic area in Western Sweden. He moved to a segregated area of southern Sweden and formed a racially mixed family. The second author currently resides in the "deep South" after growing up in a lower-middle class family in a small, and highly-liberal state, in the Northeastern U.S. Both authors are White, middle class, first generation University students who moved from predominantly White areas to more urbanized and segregated communities to receive their education credentials and teach high school mathematics. This experience catalyzed their teaching and research towards social justice and ethics, where they see education as an arena for bridging the divide between privileged and unprivileged groups. Their overlapping research in the marginalizing effects of BDA and digital traces served as a catalyst for this study.

Participant selection methods

Participants were sought by asking for volunteers for interviews on ethics and BDA within the mathematics education teacher preparation programs from the researcher's respective universities. Both researchers recruited PSTs in the last two years of their program from courses in which they served as instructors. Available demographics of the participants are listed in Table 2 below. Note that participant demographics related to the race/ethnicity of Swedish (SWE) PSTs are unavailable due to Swedish policy restrictions and the sensitive aspects of collecting race-based data (as opposed to being a common practice in the United states (U.S.).

Interview task development

The semi-structured, task-based interviews consisted of a series of six tasks developed to reflect common and/or troubling discourses in the media and data science industry in order to trigger participants' critical/ethical reasoning. Specifically, they were intended to elicit PSTs notions of systemic injustices related to the mathematical classification of individuals according to social,

Table 2
Participant Demographics.

U.S. Participant Demographics					SWE Participant Demographics				
Pseudonym	Gender	Race/ Ethn.	Age Range	Other	Pseudonym	Gender	Race/ Ethn.	Age Range	Other
Cruz	Male	Hispanic	20–25	Multi- lingual	Alice	Female	N/A	20–25	
Jenny	Female	White	20–25		Birgitta	Female	N/A	35–45	
Rachel	Female	White	20–25		Carl	Male	N/A	25–35	
Erin	Female	White	20–25		Daniel	Male	N/A	25–35	Data Scientist
Josephine	Female	White	20–25		Erik	Male	N/A	> 40	
Katherine	Female	White	20–25						

political, behavioral, and economic stereotypes which are often communicated through data-based representations (O'Neil, 2016) and to consider such notions when making data-based decisions. To develop the tasks, interview questions from the pilot studies conducted by the C-EMDT with high school students were adapted to align with the mathematical knowledge of undergraduate mathematics PSTs (Stephan et al., 2021). Their purpose was not to "teach" but to gauge how PSTs reason through and make decisions in real world, BDA contexts. To achieve this, the tasks positioned PSTs as *decision-makers* in order to promote agency while drawing on each of the CMC

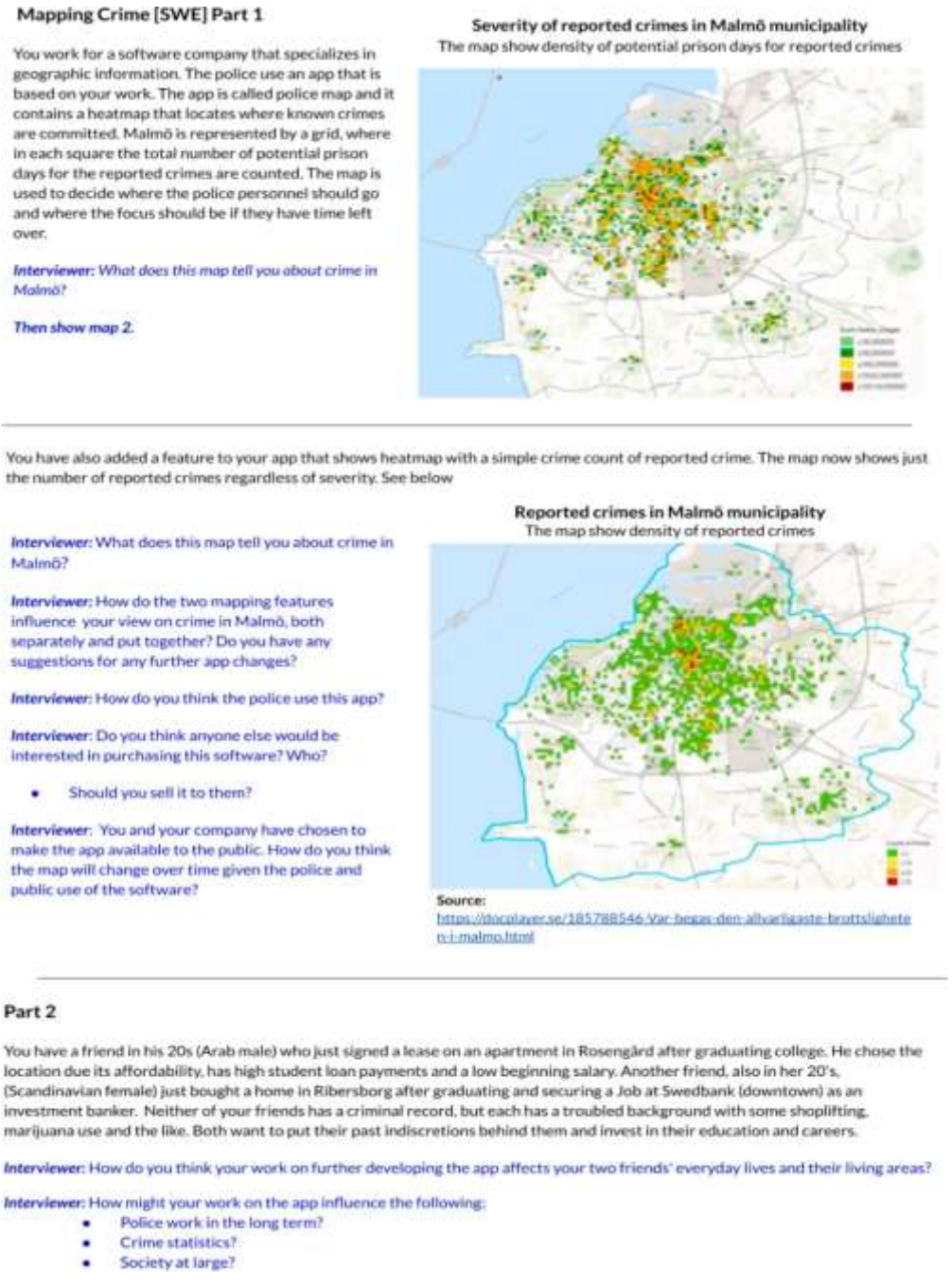

Fig. 2. Mapping Crime, SWE Version.

awarenesses and ERiM principles. The decision making aspect of the tasks is critical in that it places responsibility on the participant to make, what they consider, an ethical decision, rather than leave the interview having internalized biased or discriminatory claims about social groups. Given that what individual's deem ethical is related to their individual experiences, values, beliefs, cultures, etc. (BBC, n.d.; Chonko, n.d.; IEP, n.d; Noddings, 1988; Norlock, 2019), the participants could not tailor their responses to our notions of what is ethical or not. Thus, we expected that their responses would likely reflect the reasoning that they would typically exhibit in

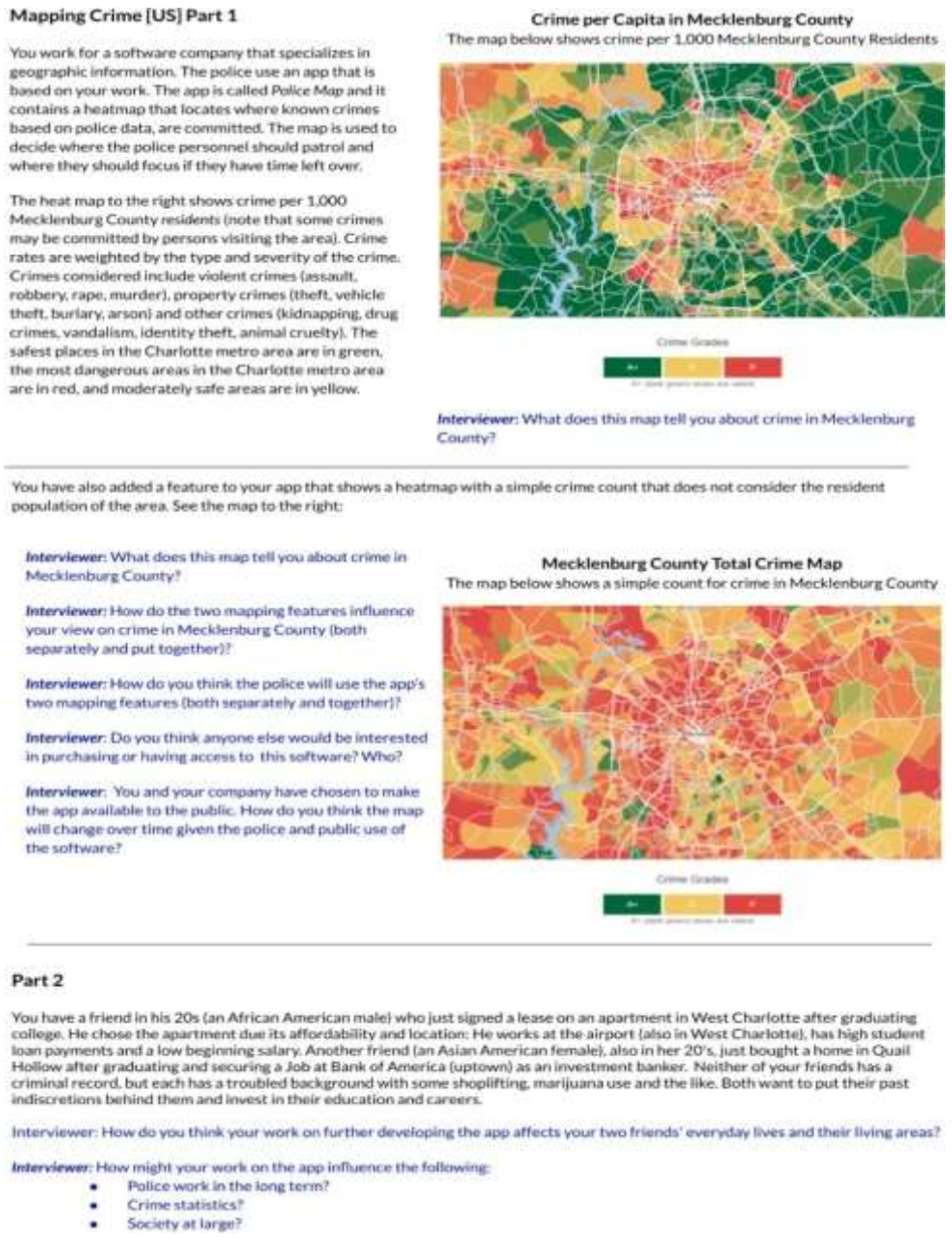

Fig. 3. Mapping Crime, US Version.

similar contexts outside of the University environment. Prior to implementation, the tasks were tested on community members and doctoral advisors from both countries to select those that promoted ethical discourse and elicited a diversity of ethical considerations from the participants. This study will report the findings of one of six tasks, called Mapping Crime.

Mapping crime

The Mapping Crime task was developed to draw on PSTs notions of historically biased data, police targeting, reinforced stereotypes, and discrimination. Heat maps were sourced from public data (see: https://docplayer.se/185788546-Var-begas-den-allvarligaste- brottsligheten-i-malmo.html, © Persson & Hallström and https://crimegrade.org/safest-places-in-charlotte-nc-metro/, ©Open- StreetMap contributors). Mapping Crime was created as a local task, in that the maps represented crime in their respective regions, and reflected dominant and/or historical discourses related to the perceived criminality of certain groups in each city (based on race in the U.S. and immigration in SWE). For instance, Black populations have been historically criminalized by society in the U.S., while Arab and immigrant populations have been so in Sweden (Anderson, 1988; Lainpelto, 2019). The goal for this task is not to reinforce common stereotypes, but draw attention to them so that PSTs may apply reasoning to critique and dismantle related discourses using data based representations that promote such misleading discourses (e.g. by drawing on their knowledge of the systemic contributions to crime rates including aggressive policing, police targeting and surveillance). Though we cannot speak to the participants' dispo- sitions after the interviews, the testing and implementation phases indicated that participants reacted to the questions in the intended way, by questioning rather than upholding the illustrated stereotypes.

A salient goal of this task was for PSTs to recognize how data representations may influence both the perceptions and behaviors of individuals within a system. To probe their consideration of access to data, PSTs were asked who might be interested in having accessto the application, should they have access, and how they feel the maps themselves might change over time given police and public useof the crime mapping application. Part 2 was intended to gauge PSTs sociopolitical awareness of the potential for negative conse- quences on marginalized populations and society. We drew on our knowledge of local stereotypes related to the perceived criminality and economic ability of social groups, positioning two hypothetical friends in locations on the map that would trigger the PSTs considerations of reinforced stereotypes, social immobility, discrimination, and ecological impact due to targeted policing. Here, we hoped that the participants may speak to the cyclical nature of BDA in which what the user observes in the data will influence their behavior, which then influences the data in an ongoing feedback loop (O'Neil, 2016).

Mapping Crime centered around a geo-localization application for data driven police work. It is thus related to the trend of pre-dictive policing based on machine learning, by potentially having the same feedback mechanism between data use and data collection that also intersects with existing prejudices and inequalities. Such applications have been described as promising for reducing crime since they would enable the police to better plan *where* to use their resources, but also critiqued since they give little input on *how* towork and could therefore increase the risk of discriminatory practices, profiling, skewed depictions and stigmatization of minorities due to focusing on correlations rather than causality (Meijer & Wessels, 2019). Although Meijer and Wessels (2019) found little empirical evidence for the benefits and none for the drawbacks of this kind of data driven police work, we still see a utility of such applications for discussing ethical aspects of BDA since the use of such applications are contested in public debate (e.g. Heaven, 2020) and statistical and mathematical literacy is required to engage in profound discussion about them. Specifically in Mapping Crime, there are several concepts that participants can build their arguments on. This includes but are not limited to; area density, a general un-derstanding of distributions and reading graphical representations of them, proportional reasoning for comparing across neighbor- hoods when both crime density and population density change, per capita versus a simple count, the role of bias in data collection (e.g. when arrests are being overrepresented for some minorities due to racism), and finally probability, since arrests and therefore, theconsequences of arrests, are probabilistic in nature.

In addition to the concerns with data driven police work mentioned above, Mapping Crime addresses the exacerbated complexity of potential ramifications that is added by the dimension of *who* has access. An example of this, is an expressed concern within the predictive policing community that organized crime could potentially adopt the same kind of data driven methodology if suitable data is available to them (Pearsall, 2010). In addition, the follow up questions within the task (blue text in Figs. 2 and 3) were meant to draw attention to the different ethical and mathematical aspects mentioned above. We conjectured that such questions would encourage participants to develop rich and diverse answers, but also recognize that they may draw attention away from other possible answers, e.g. expressing the opinion that not even the police should have access to the data discussed in the task. From the viewpoint of the research questions, and that we wanted the task to portray the reality of the data industry (that by definition must chose to work with data and sell the outcome), we see these limitations as acceptable, as we believe that there is still ample of opportunities to express advanced CMC and perspectives on access to data.

Data collection methods

The approximately one hour interviews were recorded in person or through Zoom. The researchers began by building rapport, describing the expectations for the interview, and asking a series of pre-interview questions to gauge PST their dispositions towards the role of mathematics and mathematics education in society, controversial topics, and key understandings/skills that they feel their students need to succeed in the Digital Age. Interview expectations were to employ rough draft thinking and reason outloud. PSTs were shown each task, asked to interpret the information, and then respond to the designated questions. The role of the researcher was to create a safe space, provide directions and prompt PSTs to elaborate on their reasoning in cases where their perspective was unclear. PSTs were allowed to skip questions that they deemed uncomfortable.

Importantly, all of the PSTs knew their respective interviewer from coursework and the student-teacher relationships may have

impacted their responses to the tasks. However, the contexts of the tasks had not been seen or discussed with participants prior to the interviews and their prior instruction did not include topics related to ethical reasoning or BDA. Our decision to interview our own PSTs was grounded in our socio-emotional understanding of our own PSTs and our desire to make them feel comfortable discussing controversial topics. Additionally, proponents of DBR argue that designs must respond to the needs of participants in a particular context. Thus, we were concerned only with the results of the participants we studied, recognizing that the specific results may not transfer to populations that are considerably different from our studied population.

Process of analysis

The coding process was guided by the CMC and ERiM Principles analytic frameworks. The two authors independently coded the interview transcripts from their respective PSTs. The Swedish interviews were first transcribed and coded in Swedish, then translated

Table 3
Coding Process.

Process	Purpose and Considerations	Example
1. Determine what decisions the PST would make in response to the prompts.	The participant's decision is integral to determining their CMC level. Though they may make ethical considerations at a high level of critical thought, the decision that they make speaks to their willingness to act on those considerations (i.e. agency).	that task.
2. Identify any mathematical reasoning demonstrated in their response.	To understand how participants leverage mathematics to understand the issue at hand. This is integral to CMC as the pilot studies (Stephan et al., 2021; Register, 2021) showed that mathematical knowledge can either support or hinder their critical thought in specific contexts (e.g. proportional reasoning is essential for analyzing representation and identifying issues related to disproportionality, inequity, and discrimination).	
3. Identify ethical considerations made using the ERiM framework	To determine what ethical issues are at the forefront of PSTs reasoning as it relates to data.	
4. If accessibility (AtD) was considered, determine whether they demonstrated an excess perspective, dearth perspective, or a combination of the two.	To determine how PSTs reason about access to data. That is, *who* should have access to it, in what *format*, and *how much*.	
5. Identify demonstrated CMC awarenesses (sociopolitical, ecological, communicative).	To determine what types of awareness PSTs demonstrate with regard to communication, sociopolitical and ecological impact of data science.	
6. Classify PTS's demonstrated reasoning according to its probable CMC level.	Holistic analysis of interview, including ethical considerations, CMC awareness, and decisions made-to determine their overall level of critical thought, on whom they place of onus of responsibility for change, their role in catalyzing change, and at what level that change needs to occur (individual, local, system, etc.). Note: a PST who demonstrated all three awarenesses, but either did not recognize the systemic influence on the issue, or did not see themselves as agents of change may still be classified as having lower a level of CMC for	

Which maps to include; Who should have access to theapplication; Adaptations to be made to the application.

Proportional reasoning when comparing the two maps; *StatisticalReasoning* related to the effects of the application on society (distribution changes, probability of future crimes); *Data scienceknowledge* about filtering methods and the effects of real time data.

General concern for an action being unfair for individuals →*fairness*. Concern about disparate outcomes between specificsocial groups, or marginalized populations → *discrimination*. Concern for accuracy of the maps → *accuracy*.
Dearth: All people should have access to the data to make informed decisions about where to live or frequent; Police should have open access so that they can keep neighborhoodssafe.
Excess: Open access to the Mapping Crime app might be dangerous because it could be weaponized by criminals/police,to target specific neighborhoods and/or reinforce harmful stereotypes about the people who live in certain neighborhoods.
Both: People should have access to the information to make informed decisions, but there are potential repercussions of doing so on marginalized communities which should be resolved.
Communicative: concern for how the information is communicated or miscommunicated (map accuracy).
Sociopolitical: concern for marginalized groups or individuals(police targeting, reinforced criminal stereotypes).
Ecological: concern for the effect on entities or ecologies(migration patterns).
CT: Recognizes application's influence on marginalized groups (reinforced stereotypes, feedback loop), and migration patterns.EXplains how they could reduce negative effects by employing safeguards related to who has access to the application, the amount and format of information represented, and how that information is used.
SsT: ^ but places responsibility of safeguards in others hands. *IsT*: ^ but identifies a solution that either does not attack thesystem or translate outside of the local context (show only the simple count).
Di: ^ no solution presented or solution does not protectmarginalized communities.
Dy: Identifies criminals as the issue. Does not address discrimination, targeting or stereotyping.

Int: Sees the entire problem as the way of the world.

into English in order to preserve connotations of the language. This enabled a sensibility to cultural and local contexts. After coding and classifying our respective PSTs, we coded the interviews from the other group, then met to determine a consensus on codes that differed. In a final step, coding patterns within answers were discerned and related to PSTs choices when making decisions, allowing us to determine their overall CMC level for the task. See Table 3 below.

An excerpt of the coding of Cruz is found in Table 4 below (see the appendix for the full coding for Cruz). Notice that the codes and the coding comments in this excerpt partly overlap and support each other. This is a feature of the step by step procedure explained in Table 3; the next step typically builds on previous steps. For example, once the decision is coded and described, ERiM follows from the rationale of the decision, which in turn informs which kind of CMC awareness that is displayed. In this case, Cruz' exploration of the topic is related to different perspectives conveyed by the two maps (communicative awareness) and possible migrationary impacts

Table 4
An excerpt of the codes for Cruz.

Mapping Crime Transcript	Decision	Math Reasoning	ERiM Codes	AtD Codes	CMC Awareness	CMC Level
Cruz: For the more affluent one [Asian female], the [mapping application] would likely cause prices in that area to go down. Cause, um, I remember we were talking about red lining and stuff before. Maybe some areas could be redlined as safer locations. Like only letting certain people live here, because if you have a certain type of job, you're probably not trying to commit crimes, like blue collar crime, I suppose. If your reasoning is that areas with higher housing prices will have less crime… but you find out, oh, crime here is the same as it was before the housing prices went up. you start to lose the validity of that argument for other areas. Um, the prices might go up because comparatively, this location seemed to have so much crime based on the proportion of residents compared to others… but when you look at a simple count of how many crimes occur… Well, in that case, it's the same as everywhere else.	argument changes depending on what you're looking at.	Economic changes (to the housing market) based on movement as a resultof public access to the map.			Fairness Accuracy Ecological (economic)	CommunicativeEcological
Interviewer: Okay, so what about with, uh, your friend who lives in west Charlotte?						CommunicativeEcological
Cruz: So for them, the cost of their property might go up because someone might be basing their price on this idea of, oh, it's not a safe area –. It's less safe than other areas because of the amount of crimes committed here because of the proportion of crimes per thousand residents. So, they think the area is not as good. So therefore we have to have lower prices to have people move in because people don't want to be here unless there's a good reason to. But when you look at the simple crime count for the map, the simple maps, then that doesn't really, that's not really a valid argument. The		Economic changes (to the housing market) based on movement as a resultof public access to the map.	Proportional reasoning		Accuracy Ecological (economic)	

Disempowered - Perspective of fairness in the housing market (i.e. all areas have crime so shouldn't have more or less attention from the police/ buyers). Does not mention discrimination or reinforced stereotypes

Disempowered - Looking at this from the perspective of the housing market. Does not mention potential discrimination or reinforced stereotypes

(ecological awareness).

4. Findings

Mapping crime

The findings from Mapping Crime indicate that PSTs in both the US and SWE demonstrated a range of ethical considerations, CMC levels, and perspectives on *accessibility* that were influenced by their personal experience, mathematical conceptions, and understanding of the BDA industry. CMC levels ranged from *dysconscious* to *critically transitive*. There was also a range of responses related to the PSTs' belief in how much data should be accessible, who should be able to access it, and in what format. An overview of PSTs CMC levels and exhibited awarenesses as well as their reasoning related to *accessibility* can be seen in Table 5. Details of PSTs' ethical and mathematical reasoning specific to their CMC and considerations of access to data will follow. An acronym after a PST name refers to the CMC level of the task.

Access to data

PSTs pondered over who, outside of the police, should have access to the mapping application, what specific information the maps should include, and the format of the maps themselves. Their responses and proposed solutions were classified according to how they described the problem, that is, can it be understood as being caused by a limited possession of reliable and valid data (*dearth*) or as being caused by an *excess* of data in the hands of a potential predator or adversary (*excess*)? Significantly, those who were classified as demonstrating higher levels of CMC (6 PSTs) considered both the *dearth* and *excess* perspectives on AtD as well as discrimination and the ecological impact of police and public use of the mapping application.

Dearth perspective

PSTs who demonstrated the *dearth* perspective on *accessibility* were primarily concerned with (1) providing open access to the application, (2) providing more disaggregated data, and (3) using both maps to provide a more accurate representation of crime in the region. Regarding the first, several PSTs argued that who is granted access to the data is not a cause for concern, and in some cases, is beneficial. For instance, after noting that insurance companies would be interested in accessing the data, Erik (*Di*) suggested that not only should they be allowed to access it, but that the information should be accessible by all. Similarly, Erin and Rachel *(IsT and Dy)* commented on how people who are moving or those who work in real estate would also be interested in having access. Notably, both argued that if their access was granted, it may be useful to include more disaggregated information about the historic crime rates in the area. Whereas Rachel considers the effects of *accessibility* on the consumer, Daniel and Carl *(IsT and CT)* discuss the effects of *accessibility* to disaggregated information on police distribution for the purposes of more effective crime prevention (Table 6).

A final characteristic of AtD in Mapping Crime involves PSTs considerations of how the maps should be used. Those who demonstrated the *dearth* perspective argued that greater access to information provides a more accurate representation of crime in the area and thus argued that both maps should be used together.

Table 5
Mapping Crime Ethical Considerations and Demonstrated CMC.

	SWE Participants			US Participants		
CMC Level	PST	Access to Data	CMC Awareness	PST	Access to Data	CMC Awareness
Critical Transitivity (CT) 1 SWE 3 US	Carl	Excess	Ecological Sociopolitical Communicative	Jenny Josephine Katherine	Excess, Both	Ecological Sociopolitical Communicative
Systemic Semi-Transitive (SsT) 1 SWE 0 US	Birgitta	Both	Ecological Sociopolitical Communicative	-	-	-
Isolated Semi-Transitive (IsT) 1 SWE 1 US	Daniel	Excess	Ecological Sociopolitical Communicative	Erin	Both	Ecological Sociopolitical Communicative
Disempowered (Di) 2 SWE 1 US	Alice Erik	Dearth Excess	Ecological Sociopolitical Communicative	Cruz	Dearth	Ecological Communicative
Dysconscious (Dy) 0 SWE 1 US	-	-	-	Rachel	Dearth	Ecological Communicative

Note: Rows indicate which PSTs were classified at each CMC level. Columns indicate the AtD reasoning and ethical awareness demonstrated by those PSTs within each CMC level. For instance, One SWE PST and three US PSTs were classified as Critically Transitive for this task. The SWE participants

demonstrated an *excess* AtD perspective while the US PSTs demonstrated either an excess perspective or drew on both *excess* and *dearth* AtD perspectives. All PSTs (SWE and US) demonstrated ecological, communicative, and sociopolitical awareness in their reasoning.

Excess perspective

PSTs who demonstrated the *excess* perspective on AtD were primarily concerned with the negative effects of having open access to the mapping application, and limiting which maps should be accessible. For instance, Erin (*IsT*) argued against the police having access to both maps to guide patrols, stating that since the maps tell a different story, only the second map should be used. Erin believed that the second map should be considered more accurate for the purpose of police targeting because, in her opinion, population should not be the deciding factor over where the police patrol given similar crime counts.

PSTs who wanted to limit access to the application to specific groups argued that open access could invite potentially malevolent and/or ignorant users. For instance, Cruz *(Di)* argued that open access to the application could invite criminals to track where the police will monitor, potentially influencing them to commit their crimes elsewhere. Similarly, Alice (*Di*) argued that by providing access to the public, criminal organizations (terrorists) could use the information to recruit other terrorists within the high crime areas. Notably, those who were primarily concerned with criminality did not reach higher levels of CMC because they did not explicitly consider the potentially discriminatory effects of the application. (Table 7).

AtD as it applies to discrimination, and ecological impact

For those who displayed higher levels of CMC, ecological impact and discrimination were major concerns related to AtD. For instance, several PSTs reasoned that open access to the data could potentially contribute to biased targeting by police and/or the public. For Katherine, this bias came in the form of higher police presence in certain areas contributing to increased reports of crime, but not necessarily reflecting the criminality of the area. Similarly, Daniel argued that open access to the application could reinforce police and public discrimination due to preexisting social prejudices, especially for criminalized populations living in areas with historically high crime rates. (Table 8).

Both dearth and excess

Those like Daniel and Katherine, who discussed the discriminatory effects of open access to the application, seemed to consider both the *dearth* and *excess* perspectives in their reasoning. Of the six PSTs who considered both types of *accessibility*, all displayed levels of CMC that include critical thought related to the influence on oppressed groups, and some agency to dismantle oppressive systems. For instance, Erin (*IsT*) argued for more disaggregated data within the maps (*dearth*) but stated that she would only use the simple crime count map to show that crime is everywhere and to avoid profiling (*excess*). Similarly, Birgitta (*SsT*) argues that open access to the application could be useful (*dearth*), but too much accessible information could also have negative impacts on society (*excess*), like criminal use of the application, the potential for discrimination, and effects on the housing market. Both Erin and Birgitta demonstrated ethical considerations in their reasoning. However, Birgitta's proposed solution placed the onus of responsibility on another person or entity, stating that she did not "want to be the one to decide, whether it should be released or not." Similarly, Erin's solution, though aimed at reducing discrimination, did not necessarily address the oppressive system itself. Rather, she pinpointed an isolated characteristic of the system that she felt might reduce the potentially marginalizing effects of the application. That is, she would "rather give the public and the police the map that doesn't include the population or the residencies" because "it shows crime can happen anywhere".

PSTs who displayed critical transitivity in this task considered discrimination, ecological impact, and both the *dearth* and *excess* perspectives on AtD. In addition, they demonstrated critical thought, attention to oppressive systems, and proposed solutions at the systemic level (components of CMC). For instance, Katherine (*CT*) ponders both the benefits (*dearth*) and constraints (*excess*) of providing open access to the mapping application, but ultimately decides that its use by the police or public is not worth the risks of discrimination and de facto segregation through the housing market (*excess*). Specifically, she states that if the public has access to the application, people may flock to low-crime areas that are typically wealthy, resulting in more resources for those areas. In addition, she disagrees with the assumption that targeting "*crime-ridden areas*'' will lower their crime rates. Rather, the targeting of those areas specifically, may actually make the crime rates go up, and deter people from moving there.

Josephine (*CT*) further discusses the benefits and constraints of open access, including the negative effects on the housing market and racial profiling by police (*excess*). (Table 9).

Josephine ultimately determines that both maps should be used because she believes that more information is always better (*dearth*). Her considerations of racial profiling by police due to the nature of the application was also evident in Carl and Jenny's (both *CT*) reasoning. They argued that providing statistical information to a population that may not have the appropriate statistical

Table 6
Example Responses (Dearth).

Rachel *(Dy)*	It'd be cool to look at the historic and the current rate of crime [.] Could we [also] have it separated based on the type of crime? […] because if one car gets broken into [its not ideal] but if you have a bunch of murders, or like attempted murders, or assaults or whatever, like you might want to steer away from that area.
Erik (*Di*)	I think this can be an open information as I see it. So that, no I do not see that there is any major problem actually.
Daniel *(IsT)*	Does this include things like tax crimes and tax evasion which can also have a very high penalty? [.] If you are going to use it to distribute where the police should be, then maybe you should probably look at the type of crime that is solved by a patrol. So you would need some filter to sort out crimes.

***Carl** (CT)* If this is to be used to distribute police presence in the city, one would probably have needed to take more account of the frequency of crimes, and *to what extent these crimes could have been prevented by police presence*.

Example Responses (EXcess).

***Cruz** (Di)*	So in the end, if you were to give people the information, people would flock to safer places. And maybe someone who makes a living with cat burgling is going to be like, oh, people think they're safe. So they're not going to buy home security. I'm going to rob them now.
***Alice** (Di)*	Yes, terrorist organizations would be one that I would keep it away from [accessing the data][.] Because of reduced recruitment.
***Erin** (IsT)*	I just would rather give the public and the police the map that doesn't include the population or the residency's. It shows crime can happen anywhere. It doesn't happen in just one singular location.

Table 8
Example Responses (AtD, Discrimination and Ecological Impact).

***Daniel** (IsT)*	People always have a lot of prejudices. And if you come from there [high crime area] and work at a good job, there is nothing to suggest that that person would do something bad. [.] If this is public [.] and unfortunately, if there are any sort of racist ideas, then there are people who can draw such preconceived notions.
***Katherine** (CT)*	I also was just thinking about the fact that if they're using these mapping features to identify [.] what areas they are going to patrol more, there's going to be more reported crime in areas where there's more police because there's [.] more police to catch things that are happening.

knowledge to interpret them could have potentially negative effects.

As intended, *accessibility* was a consideration made at every CMC level, albeit in slightly different ways. However, the differentiating factors between those who displayed higher levels of CMC were the ways in which the PSTs considered the ecological impact of using the maps and their potential for discrimination in the community. Importantly, such considerations were noticeably dependent on the PSTs' demonstrated proportional and statistical reasoning.

Mathematical reasoning

Up until this point, we have sparsely discussed the role that PSTs mathematical reasoning played in their demonstrated ethical reasoning. Recall the first map represented the density of crimes by population and/or severity of the crime, whereas the second map represented a simple frequency of crimes in the same area. In their reasoning, PSTs drew on their understanding of proportions, probabilities and distributions, and one, drew on his knowledge of data analytics to understand the maps and their effects on society.

As a general trend, all PSTs drew on their proportional reasoning to understand what was being represented in the two graphs. (Table 10).

Additionally, several PSTs used their statistical knowledge to reason that the mapping application is a probabilistic representation of where crime is likely to occur based on past data from the police. They drew on their understanding of distributions to discuss how the map may change according to real time data depending on how the public and police act on their interpretation of the results. For instance, people may move to low crime areas to feel safe and police may target high crime areas to either reduce crime or fill a quota, both of which will influence migration patterns within the area. As mentioned above, Katheriene (*CT*) described how the maps could make the police *find and report* more crimes in certain areas. In contrast, Daniel's expressed reasoning reflected his experience as a data analyst and his objective trust in statistics, a common theme that we observed with high school students in the pilot studies and a common notion in Western cultures. (Table 11).

As mentioned in the dearth perspective section, Daniel expressed his data science knowledge by arguing that filtering methods should be applied to the algorithm to break the information down according to which crimes could actually be addressed through police patrols. In addition, he discussed the function of real time data and its potential effects both on the mapping application itself, and on the community as a result. Specifically, Daniel argues that if people choose where to live based on the adaptive maps, you will not only see these changes in real time, but the changing populations may illustrate a faulty depiction of the criminality of some regions given that crime is typically more concentrated in densely populated areas. Such reasoning was shared by several other PSTs who discussed their concerns about the behavioral (ecological) effects of open access to real time data, a characteristic of BDA that is not a common consideration in the theoretical mathematics of the secondary school classroom.

Table 9
Example Response (Both Dearth and Excess).

***Josephine** (CT)*	I would definitely say for the [Black male living in a high crime, low income area], he's probably going to see a higher police presence [.]. He may have a harder time putting his background behind him, because if he lives in an area where they already have the mindset that the people who live here are around more crime, then any small offense might be taken more seriously. Whereas for the girl living in the better part of town, she might be around police officials less often. And if she were to ever encounter any problems, like a speeding ticket or something, they might not look any further into it if she is in an overall better part of town.
Interviewer	And is that because of the area they live in or their identity?
Josephine	I think both [.] But, I think that if everybody saw a map more like the one on the right [count], it might challenge their perception on certain neighborhoods and the safety of them.

Example Responses (Proportional Reasoning).

Map 1	
Jenny (*CT*)	So when that says, like per 1000 residents, that means like they're basing off of like how many people out of a thousand for that area. [.] Crime rates are a weight of type and severity of the crimes considered… The safest places in the Charlotte Metro area are in the green. The most dangerous areas in Charlotte are the red and moderate, the yellow.
Map 2	
Jenny	So this isn't the ratio. This is just a count. Well, this kind of shows that….like overall there's a general pattern of not good.
Alice (*Di*)	The first picture says what type of crime it is and where in the city it is most common [.] and where to put specific efforts against specific crimes. The second picture rather says how common it is with crime in general. Whether there are certain regions where it is much more common, where you may have to have more police presence.

The role of mathematics in their ethical decision-making

Notably, proportional reasoning allowed PSTs to enter the task, whereas their knowledge of data science and statistical reasoning related to probabilities and distributions allowed them to elaborate on the marginalizing effects of the mapping application in society. Interestingly, their mathematical and statistical reasoning only partially served as a basis for their ethical decisions but was essential for reaching higher levels of CMC. That is, PSTs drew on their contextual knowledge of local crime, police targeting, and common stereotypes to guide their analysis, often contributing to a state of cognitive dissonance over which reasoning (contextual or mathematical) to trust. For instance, when asked which map(s) should be used by the police and in what capacity, several of the PSTs toggled between their typically positive orientation towards the use of proportional representations and their understanding of the context at hand. Many indicated that the use of both maps would provide a more well-rounded picture of crime in their respective cities. However, when asked which maps the police should use to target their patrols, they maintained that police should attend to the simple crime map because if crime occurs everywhere, it is unfair to target areas with high crime grades since those areas are often densely populated urban areas with a higher proportion of historically marginalized and/or criminalized people by society. (Table 12).

Partial mathematical and statistical conceptions may have also played a part in their decision making. While it is clear that many of the PSTs attempted to limit the negative effects of the application on marginalized communities, their lack of understanding related to the techniques of BDA restricted several PSTs from making efficient and effective decisions about how to improve the application. For instance, choosing to restrict the application to the simple crime map avoids discriminatory practices based on the map, but does not provide a solution for reducing crime in the city. Though not confirmed, we believe that such generalized responses may be due to a lack of knowledge related to the design and analytic possibilities within the data and computer science fields. Additionally, many mentioned that crime is typically concentrated in densely populated areas, but ignored this in their decision-making. In an interesting case, Birgitta chose not to make a decision about who should be allowed to use the mapping application because she did not want to be responsible for its outcomes. Leading up to this statement, she indicated a flawed understanding of probability, a key concept which lies at the heart of data science. Birgitta implies that since the events on the map have already happened, their probability of occurring again would be low, potentially contributing to her lack of agency. Thus, it seems that in order to make ethical decisions in BDA contexts that are effective, an understanding of the facets of BDA, including its design and analytic possibilities, its probabilistic nature, the effects of filtering and aggregating data, and among others, the effects of real time data on human behavior, are essential.

5. Discussion

The results indicated that most PSTs were able to draw on their mathematical and statistical knowledge to understand data-based representations and related claims that have ethical implications in society. Some of the PSTs further demonstrated their ability to draw on such knowledge to make ethical arguments. Thus, mathematics was used to both understand the issue and to suggest a solution. Additionally, PSTs made coherent mathematical connections between the role of the data in the mapping application and access to the data for different actors, (e.g., arguments for which map the Police should have access to based on mathematical interpretations of the maps). While we expected that PSTs at the end of their teacher preparation program could use mathematics in this way, less expected was the richness in perspectives when discussing access to data.

We explicitly asked the PSTs about access to data and expected that they would give reasons to either grant access or to deny it. As it

Table 11
Example Responses (Statistical Reasoning).

Daniel (*IsT*)	There may not be as many who would like to move into the center city, or where the criminals are [.] But [the app] may make it a little clearer. *At least for me, I trust the statistics more than someone saying that, or writing that you should not live there.* So I could have been persuaded not to live here… If this data had been available.
Interviewer	So you're saying it's not that there's different amounts of crime. Is it that if they target those areas, they're going to be reported more often or be convicted more often?
Katherine *(CT)*	Potentially yes, or maybe there is less crime, but it's going to make it seem like those areas where it's already yellow might become red just because if there's more police around, there's more chances to get caught…. When you're thinking of the African-American male, if he is living in that area that's red … and the police are going to be using this information to target more of the red areas, I guess, that it might be more difficult. Like even if he is like, putting his past discretions behind him, he's still at risk of being targeted by police just because he is black and living in an area that is, (I might put this in quotes) "known for its crime".

Table 12
Example Responses (Cognitive Dissonance between Mathematical and Contextual Knowledge).

Cruz (*Di*)	If used separately, the first map that has a proportion will probably be somewhat misguiding because they'd probably be more focused on areas like this and might disregard areas like this, where crimes are still happening, but due to the population, it might not be as visible but it's still important because this is a group of people that you are simply leaving to suffer.
Birgitta (*SsT*)	It has already happened. So then the probability is not so high that it happens again. Yes, tricky. I do not want to be the one to decide, whether it should be released or not.

were, several PSTs gave elaborate descriptions that held both dearth and access perspectives concerning the same decision, giving detailed accounts of both the pros and cons. In this way, their answers clearly challenged the notion of general truths about access to data such as 'more data is always better' (dearth) or 'data should be protected' (excess). Instead, the PSTs drew from the specific context of the questions to guide their ethical reasoning thereby acknowledging that the complexity of real world situations require a more nuanced approach. Seeing such complexities and considering both dearth and excess perspectives were more prevalent among PSTs that demonstrated higher CMC levels while other PSTs demonstrated less advanced perspectives on AtD leading to scant ethical reasoning. This suggests that although advanced reasoning on AtD occurred among some of the PSTs, it cannot be expected to occur automatically and consistently among all PSTs. Rather we see that there is potential to develop PSTs reasoning as it relates to who should have access to data and in what format and quantity.

The long-term effects and consequences of data driven practices (referred to as ecological impact in the Findings) was another consideration we explicitly asked PSTs to make in their reasoning. Some responses here were also surprisingly advanced. Several PSTs described long-term interaction mechanisms between human behavior and the data. This is best illustrated by Katherine's reasoning, quoted in the findings section, on how the data can affect police behavior, which then affects future data collection. Understanding such feedback loops demonstrates insights into how access to data and mathematics can interact. In Mapping Crime however, PSTs mathematical concerns were primarily focused on the accuracy of the graphical representation of the crime data and how those representations would be used. A more robust analysis, however, is revealed in Katherine's reasoning. She realizes that the long-term effects of data driven practices are not only dependent on (1) *what the data supposedly portray*, or (2) *who has access* to it, she points at a phenomenon where (1) and (2) interact. This is because the color on the map changes when the police get access (i.e., the color no longer portrays only reported crime) it now also reflects the fact that the Police *have access* to the data and the concurrent change in patrolling as a result of this. This consequence analysis operates on another level than, for instance, saying that one of the maps is preferable over the other since it foregrounds more important perspectives, or that some persons should (not) have access to the data because of what they can do with it. Instead of drawing from only either mathematical understanding of the maps, or concerns for data access in this societal context, she combines them to produce a synthesis that goes beyond what either approach can achieve alone. While not all of the PSTs demonstrated this reflexive mathematical and AtD reasoning, we observed that PSTs have this capability, indicating a potential to foster such pluralistic reasoning.

One could perhaps argue that advanced reasoning on the interplay between mathematics and data access is too much for the needs of teachers, and that it would be sufficient with a lay citizen's caution about data in the hands of an adversary. This is reminiscent of how Skovsmose (1990) discusses reflective knowledge as important for citizens' ability to evaluate and critique, and as such, may occasionally be separated from technological knowledge. Instead, we argue that a general awareness *that* it might be a bad idea to give data/information about oneself to a potential adversary, may be enough to form a 'first line defense' of critical thinking without also understanding *why* or *how*, but it is not enough to develop rigorous ethical reasoning as it relates to BDA. The mathematics in Mapping Crime is comprehendible for PSTs. But in other applications of BDA, mathematics can go far beyond their current training. By tweaking Mapping Crime to involve predictions of future maps, we could for instance have introduced predictive statistics with machine learning. For such examples, knowing the involved mathematics is a requirement for doing the kind of insightful synthesis of AtD and mathematics that Katherine did. Such synthesis can uncover important perspectives. We therefore argue that a lay citizen's caution regarding data in the hands of an adversary, does not cover all needs for education. Rather, we suggest that this level of understanding, while potentially sufficient for consumers of BDA, is insufficient for future industry professionals and policymakers. Thus, a contribution of this study to the CME research field is that teachers need to understand the *why* and *how* behind common BDA practices, since they will teach students who will either become producers or consumers of such practices.

Implications

Our findings have implications for this special issue's aim to support a more robust and critical mathematical understanding of the world, in that they show the possibility and value of PSTs in discussing access to data in relation to ethical reasoning on BDA. As was described in the literature review, accessibility to data is an important topic for understanding how BDA can operate in society. Our results show that it is possible for PSTs to discuss this important concept in depth, and hence it ought to be possible to include AtD in teacher education. Our results also indicate that ethical reasoning can be enhanced if it is simultaneously informed by both understanding the mathematics involved, and AtD in the particular context. We therefore argue that an Ethical Mathematics Education in the digital age ought to include instruction on the key concepts and skills of data science and BDA.

This raises two challenges for mathematics education. First, teacher education must encompass more of the mathematical tech-

nologies used in BDA. This includes probability and statistics, that in the case of the USA have been situated in the periphery of mathematics education (Boaler & Levitt, 2019). Second, teacher education may have to reconsider how data is discussed in general. In

mathematics education, we typically hold the perspective that more data and/or more information is better. We rarely discuss the benefits of regulating and restricting data from a mathematical or ethical standpoint. It is important to integrate the ethical underpinnings of this viewpoint as it relates to privacy, property, discrimination, ecological impact, etc.

In addition, our findings indicate that an ability to consider multiple perspectives on a societal issue may be connected to higher levels of CMC. This is indicated by the combination of excess and dearth perspectives held by PSTs at the critical levels of CMC. Therefore, a context of a highly complex authentic societal issue, can become a resource to draw from when developing PSTs' CMC, equivalent with using wicked problems in mathematics education (e.g., Steffensen et al., 2018).

To summarize, an additional awareness of what outcomes are possible with BDA could draw attention to the value of personal data. This includes assessing the ramifications of providing open access to data, or providing access to specific groups. Teachers would then be able to focus their instruction on the nuances of BDA that promote the use of individuals' personal data, having significant consequences for individuals and groups in society. Such knowledge may promote new possibilities for ethical mathematics education in the Digital Age.

Declaration of Competing Interest

None

Acknowledgments

We are grateful for advice from Lisa Björklund Boistrup MAU, Helena Roos MAU, Michelle Stephan UNCC, MEWS at MAU, and Gerald Goldin Rutgers University. This article reports on research conducted within the Charlotte ethical mathematics design team and UNCC Center for STEM Education (second author), and the graduate school *Relevancing mathematics and science education* (first author), the latter funded by the Swedish National Research Council.

Appendix

Table A.1
Example Coding Process for Cruz.

Mapping Crime Transcript	Decision	Math Reasoning	ErIM Codes	AtD Codes	CMC Awareness	CMC Level
Interviewer: I'm just going to have you read the paragraphs and then take a look at this map and let me know if you have any questions. Then respond to the questions in blue on the bottom. Cruz: All right. So the maps based on the information here indicated is, uh, when it comes to crime, it seems like a good portion of it is in this near central Western area…Yeah, it's located there. Whereas the Southern part is considered safe. Cruz: That's what the map tells me. It's that for every thousand residents, a crime is there. **And if it's darker red, the darker it is, then more crimes are committed there as a result, but it's proportionately more crimes proportionally committed there in comparison to someplace that's green[1].** Interviewer: Okay. And, and just to clarify, this is, you know, crime rates there's, what's in	here, it's, you know, per 1000 Mecklenburg county residents. They're considering the	Crime more concentrated indarker areas				

Proportional Reasoning Graphical Analysis

Mapping Crime Transcript	Decision	Math Reasoning	ErIM Codes	AtD Codes	CMC Awareness	CMC Level
population in it. Um, so this next map, um, we've added a feature that shows a heat map with a simple crime count that does not consider the resident population of the area. So what does this map tell you about crime in Mecklenburg county?						
Cruz: Well, this **map [2] is a lot clearer. It's telling me that crime is equivalent just about everywhere.** Given how much red is.	Crime is equivalent everywhere Crime count clearer				Communicative	
Interviewer: So, how do you think these two mapping features influenced your view on crime and Mecklenburg county either separately or put together?						
Cruz: Well, **the total crime map gives you a far more accurate view because it's just counting crime regardless of the population there, which is important because crimes, when people commit a crime, they're not exactly concerned with how many people are living there. They're concerned about who lives there.** If someone's going to rob a house, they want to rob the person with the most expensive stuff, not the person with the most amount of humans in there, or the least crimes, I guess, are more agnostic to those type of concepts. It's rather if it's wealthier here, take care, but the map here shows that crimes are committed everywhere. **The other map makes it less apparent. Because it hasn't scaled towards proportional how many thousand people that live there. If you have an equal amount of crimes, if not more than another area, but if you have double or triple or even quadruple the population, then it's not going to be that apparent because here you have areas in orange, red that were once green on the other map.**	Crime count clearer than crime rate	Proportional Reasoning	Accuracy		Communicative	
Interviewer: Right. Okay. So how do you think the police will use the application's mapping features either separately or together?						
Cruz: Well, if you use them separately, the first map that has **a proportion will probably be somewhat**					Fairness[2,3] Accuracy[3]	

Ecological[2]
Sociopolitical[2]
Communicative[1]

Mapping Crime Transcript	Decision	Math Reasoning	ErIM Codes	AtD Codes	CMC Awareness	CMC Level
misguiding[1] because they'd probably be more focused on areas like this [high crime rate] and might disregard areas like this [low crime rate], where crimes are still happening, but due to the population, it might not be as visible or as noticeable, but it's **still important because this is a group of people that you are simply leaving to suffer**[2]. The **other map is more accurate**[3] because it shows like, "Hey, you have to just look everywhere. There is crime everywhere." **You can't just focus on certain spots because of what a proportion tells you. A crime count is fair**[3]. It lets you count what's happening where, and that's it. There's nothing peculiar about a proportion. What it's basing off of what the population is made of. That's important.	down to this area since,					
Interviewer: Yeah. Okay, I can see that.						
Cruz: **If you use them both together, then I suppose you'd have a somewhat more accurate method of gauging them**[1] because **you can use this total crime map in conjunction with the other one**, just to, okay. So this area has high crime, but also has a higher population. So it looks less. Then possibly it might still require more police to be there simply because there's a higher population to keep safe, should a crime be something that could affect a populous, like, I guess, domestic terrorism, something like that.			Use both graphs together	Accuracy	Accessibility	
Interviewer: Right. Okay. That makes sense. Okay. Do you think that anyone else would be interested in purchasing or having access to the software?				Accessibility		
Cruz: Um, yeah. I believe there's a lot of people that would be interested simply to gauge. I mean, even simply to know where to live, if they want to live in a place with less crime, they might want something that just tallies up how many crimes are committed in a certain location and decide from there within the most recent year, how many crimes were committed. Okay. And then just measure it like historically, has crime gone						

Dearth

Mapping Crime Transcript	Decision	Math Reasoning	ErIM Codes	AtD Codes	CMC Awareness	CMC Level
whenever I suppose, how has that changed? Does this mean it's a safer era area, comparatively to other? That would be important to some people.	**they think they're safer. Andnow more crimes will be**					
Interviewer: Okay. Do you think that they should have access to that?						
Cruz: Um, yeah, I think, **I think with a lot of statistics, it's always important to have access to them because you can, if you know how to look at them and how to read them, you can make a much more informed decision on what you want to do.**			Public should haveaccess			
Interviewer: Okay. All right. So you and your company have chosen to make the app available to the public. So now given the public and the police-use of the software, how do you think the map will change over time?						
Cruz: If you give it to the public, then you're going to have people moving to areas that are green, that have more, they have less crime rates that are considered safer, but as the population density of those areas grow, you're having as many Law abiding citizens moving in with as many possible criminals, I guess, really not the best way to say it, but oh yeah. I mean…						
Interviewer: Right, we don't really know what crime they committed or if they were rightfully convicted, so I guess we could say "criminals" in quotations just in case.						
Cruz: Yeah. Quote-unquote criminals. Since there's a higher density, that area might have a higher crime count in the future. I don't know. Okay. **So in the end, if you were to give people this information, people would flock to safer places. And maybe someone who makes a living with cat burgling is going to be like, oh, people think they're safe. So they're not going to buy home security. I'm going to rob them now. And I think if you give people this information, it'll just start to, as time goes on, level out everywhere to have about the same crime, because people will just move where**						

				e	aggressive policing, police targeting, etc.).
				t h a t c r i m i n a l s a r e t h e p r o b l e m . D o e s n o t m e n t i o n d i s c r i m i n a t i o n (e . g .	
Accessibility	Dearth				
Ecological gical	Ecolo				
Ecological		Disempowered: D e m o n s t r a t e s p e r s p e c t i v			

Mapping Crime Transcript	Decision	Math Reasoning	ErIM Codes	AtD Codes	CMC Awareness	CMC Level
committed there because more people live there. So there's more potential for crimes to happen. Okay. It's not safe anymore. They'll move somewhere else. There's not really a good method.	move in because people don't want to be here unless there's		access to the map.			
Interviewer: Okay. I understand what you're saying. All right, next part of this question. Please read the excerpt…. So how would your work on developing the app affect your two friends' everyday lives and living areas?						
Cruz: For the more affluent one [Asian female], the [mapping application] would likely cause prices in that area to go down. Cause, um, I remember we were talking about red lining and stuff before. Maybe some areas could be redlined as safer locations. Like only letting certain people live here, **because if you have a certain type of job, you're probably not trying to commit crimes, like blue collar crime, I suppose. If your reasoning is that areas with higher housing prices will have less crime… but you find out, oh, crime here is the same as it was before the housing prices went up. you start to lose the validity of that argument for other areas.** Um**, the prices might go up because comparatively, this location seemed to have so much crime based on the proportion of residents compared to others… but when you look at a simple count of how many crimes occur… Well, in that case, it's the same as everywhere else.**		Economic changes (to the housing market) based on movement as a result of public access to the map.				Fairness Accuracy Ecological (economic)
Interviewer: Okay, so what about with, uh, your friend who lives in west Charlotte?						Accuracy Ecological (economic)
Cruz: So for them, the cost of their property might go up because someone might be basing their price on this idea of, oh, it's not a safe area –. It's less safe than other areas because of the amount of crimes committed here because of the proportion of crimes per thousand residents. So, they think the area is not as good. So therefore we have to have lower prices to have people		Economic changes (to the housing market) based on movement as a result of public			Proportional reasoning	

		or reinforced stereotypes
CommunicativeEcological	Disempowered - Perspective of fairness in the housing market (i.e. all areas have crime so shouldn't have moreor less attention from the police/buyers). Does not mention discrimination or reinforced stereotypes	
CommunicativeEcological	Disempowered - Looking at this from the perspective of the housing market. Does notmention potential discrimination	

Mapping Crime Transcript	Decision	Math Reasoning	ErIM Codes	AtD Codes	CMC Awareness	CMC Level
a good reason to. But when you look at the simple crime count for the map, the simple maps, then that doesn't really, that's not really a valid argument. The argument changes depending on what you're looking at.						
Interviewer: All right. That makes sense. One more question about this one. How do you think that your work on app will influence either police work in the long term, crime statistics, or just society at large?						
Cruz: Well, this would give a tool to analyze the area in which you live ….for people in which that information matters. Then it might be really influential in their lives, crime per capita. If someone cared about specifically that then, okay, you have a perfect map to help you analyze where to live based on crime per capita. But for most people, just knowing how much crime occurs in an area without anything that affects it, just simply the total number, then they're just going to see everywhere that crime is about the same as anywhere else so their decisions might not be as effective…. Each map produces a different view and one makes it seem like other areas are safer. The other one's simply more pragmatic.						

References

Agrawal, V. (2017). *Why is Big Data so important in today's world?* Data Floq. https://datafloq.com/read/why-is-big-data-so-important-in-todays-world/2674 .

Anderson, J. (1988). *The education of blacks in the south* (pp. 1860–1935). University of North Carolina Press,.

Atweh, B., & Brady, K. (2009). Socially response-able mathematics education: Implications of an ethical approach. *Eurasia Journal of Mathematics, Science & Technology Education*, *5*(3), 267–276. https://doi.org/10.12973/ejmste/75278

Bakker, A., & Van Eerde, H. A. A. (2014). An introduction to design-based research with an example from statistics education. In A. Bikner Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Doing qualitative research: methodology and methods in mathematics education* (pp. 429–466). Springer. https://doi.org/10.1007/978-94-017-9181-6_16 .

Andersson, C.H., Björklund-Boistrup, L., Roos, H., (2022) Mathematics curriculum discourses on democracy: critical thinking in the age of digital traces. Twelfth Congress of the European Society for Research in Mathematics Education (CERME12), Feb 2022, Bozen-Bolzano, Italy. https://hal.archives-ouvertes.fr/hal-03747769/document.

BBC. (n.d.). *Ethics: Guide to moral and ethical issues*. BBC. Retrieved March 24, 2022, from https://www.bbc.co.uk/ethics/guide/ .

Berry, R. Q., III (2003). *Voices of African American male middle school students: School mathematics students. (Unpublished Doctoral Dissertation)*. Chapel Hill, NC: University of North Carolina at Chapel Hill,.

Berry, R. Q., III (2004). The equity principle through the voices of African American males. *Mathematics Teaching in the Middle School*, *10*(2), 100–103.

Boaler, J. and Levitt, S.D. (2019) *Opinion: Modern high school math should be about data science — not Algebra 2*. LA Times. https://www.youcubed.org/wp-content/uploads/2019/10/LA-times-op-ed.pdf .

Boylan, M. (2016). Ethical dimensions of mathematics education. *Educational Studies in Mathematics*, *92*, 395–409. https://doi.org/10.1007/s10649-015-9678-z .

Chonko, L. (n.d). Ethical theories. Direct Selling Education Foundation (DSEF). https://www.dsef.org/wp-content/uploads/2012/07/EthicalTheories.pdf .

Cobb, P. (1999). Individual and collective mathematical development: the case of statistical data analysis. *Mathematical Thinking and Learning*, *1*(1), 5–43. https://doi.org/10.1207/s15327833mtl0101_1

Cobb, P., Confrey, J., diSessa, A. A., Lehrer, R., & Schauble, L. (2003). Design experiments in education research. *Educational Researcher*, *32*(1), 9–13.

Corbett-Davies, S., Pierson, E., Feller, A., Goel, S., & Huq, A. (2017). Algorithmic decision making and the cost of fairness. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, (pp. 797–806). ACM.

org/10.2307/1229039
D'Ambrosio, U. (1998). Mathematics and peace: Our responsibilities. *Zentralblatt für Didaktik der Mathematik*, *30*, 67. https://doi.org/10.1007/BF02653170
D'Ignazio, C., & Klein, L. F. (2020). *Data feminism.* Cambridge, MA: The MIT Press.
Darragh, L. (2021). Innovative learning environments and the digital era: Finding space for mathematics identity. In D. Kollosche (Ed.), *Exploring new ways to connect: Proceedings of the Eleventh International Mathematics Education and Society Conference* (Vol. 1, pp. 7–28). Tredition. https://doi.org/10.5281/zenodo.5469648 .
Ernest, P. (2018). The ethics of mathematics: Is mathematics harmful? In P. Ernest (Ed.), *ICME-13 Monographs* (pp. 187–216). https://doi.org/10.1007/978-3-319-77760-3_12
Esmonde, I. (2014). "Nobody's Rich and Nobody's Poor … It Sounds Good, but It's Actually Not": Affluent Students Learning Mathematics and Social Justice. *Journal of the Learning Sciences*, *23*(3), 348–391. https://doi.org/10.1080/10508406.2013.847371
Freire, P. (1970/2018) *Pedagogy of the Oppressed* Penguin Books.
Gould, R., Machado, S., Ong, C., Johnson, T., Molyneux, J., Nolen, S., Tangmunariunkit, H., Trusela, L., Zanontian, L., (2016). Teaching data science to secondary students: The mobilize introduction to data science curriculum. In J.Engel (Ed.), *Promoting understanding of statistics about society. Proceedings of the Roundtable Conference of the International Association of Statistics Education (IASE).* Berlin. https://iase-web.org/Conference_Proceedings.php .
Gutiérrez, R. (2013). The Sociopolitical turn in mathematics education. *[Equity special issue]. Journal for Research in Mathematics Education*, *44*(1), 37–68.
Gutstein, E. (2006). *Reading and writing the world with mathematics.* New York: Routledge.
Heaven, W.D. (2020, July 17). Predictive policing algorithms are racist. They need to be dismantled. Lack of transparency and biased training data mean these tools are not fit for purpose. If we can't fix them, we should ditch them. MIT Technology Review. https://www.technologyreview.com/2020/07/17/1005396/predictive-policing-algorithms-racist-dismantled-machine-learning-bias-criminal-justice/ .
Heinemann, B., Opel, S., Budde, L., Schulte, C., Frischemeier, D., Biehler, R., Podworny, S., & Wassong, T. (2018). *Drafting a data science curriculum for secondary schools.* In Proceedings of the 18th Koli Calling International Conference on Computing Education Research (Koli Calling '18). Article 17, 1–5. Association for Computing Machinery, NY. https://doi.org/10.1145/3279720.3279737 .
Hooks, B. (2014). *Teaching to Transgress.* Routledge.
IEP (n.d.). Internet encyclopedia of philosophy. Retrieved March 24, 2022, from https://iep.utm.edu/ .
King, J. (1991). Dysconscious racism: Ideology, identity, and the miseducation of teachers. *The Journal of Negro Education*, *60*(2), 133–146. https://doi.org/10.2307/2295605
Koh, Y. (2020). The movement to modernize math class. *The Wall Street Journal.* https://www.wsj.com/articles/the-movement-to-modernize-math-class-11605110629 .
Kokka, K. (2020). Social justice pedagogy for whom? Developing privileged students' critical mathematics consciousness. *The Urban Review*, *51*(1), 101–122. https://doi.org/10.1007/s11256-020-00578-8
Lainpelto, K. (2019). Ett dömande eller ett fördömande rättsväsende?–En fråga om integration (2019). In K. Helander, & P. Leviner (Eds.), *Barn, migration och integration i en utmanande tid* (pp. 127–158). Ragulka Press (2019).
Lengnink, K. (2005). Reflecting mathematics: An approach to achieve mathematical literacy. *ZDM*, *37*(3).
Matthews, D. (2019). *The importance of maths education to the future of Big Data.* Data Floq. https://datafloq.com/read/importance-maths-education-future-big-data/6108 .
Meijer, A., & Wessels, M. (2019). Predictive policing: Review of benefits and drawbacks. *International Journal of Public Administration*, *42*(12), 1031–1039. https://doi.org/10.1080/01900692.2019.1575664
Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism.* NYU Press.
Noddings, N. (1988). An ethic of caring and its implications for instructional arrangements. *American Journal of Education*, *96*(2), 215–230.
Norlock, K. (2019). Feminist Ethics. Stanford Encyclopedia of philosophy. https://plato.stanford.edu/entries/feminism-ethics/ .
O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy.* Penguin.
Pearsall, B. (2010). Predictive policing: The future of law enforcement? *NIJ Journal*, *266*, 16–19.
PISA (2020). *PISA 2021: Mathematics Framework.* https://pisa2021-maths.oecd.org/ .
Register, J., Stephan, M. ([_])#38; Pugalee, D. (in press). Ethical Mathematics Consciousness: An analytic and design framework for developing students' ethical reasoning in mathematics and data science. 2023 AMTE Conference, New Orleans.
Register, J., Stephan, M., & Pugalee, D. (2021). Ethical reasoning in mathematics: New directions for didactics in US mathematics education. *Mathematics*, *9*(8), 799. https://doi.org/10.3390/math9080799
Renfors, C.,. Justitieombudsmannen. (2015) Allvarlig kritik mot dåvarande Polismyndigheten i Skåne för bristerna vid behandlingen av personuppgifter inom myndighetens kriminalunderrättelseverksamhet. (Dnr 5205–2013). www.jo.se http://www.jo.se/PageFiles/6353/5205–2013.pdf .
Rubel, L. (2017). Equity-directed instructional practices: Beyond the dominant perspective. *Journal of Urban Mathematics Education*, *10*(2), 66–105.
Rubel, L., Hall-Wieckert, M., & Lim, V. (2016). Teaching mathematics for spatial justice in urban schools: Beyond a victory narrative. *Harvard Educational Review*, *86*(4), 556–579. https://doi.org/10.17763/1943-5045-86.4.556
Rubel, L., Peralta, L., Herbel-Eisenman, B., Jiang, S., Kahn, J., Lim, V. (2021) Theorizing data science education: an intersectional feminst perspective on data, power, and "playing the gate" [Paper presentation]. *Proceedings of the forty-third annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, 217–221.
Sabbagh, D. (2020, September 28) Trump 2016 campaign 'targeted 3.5m black Americans to deter them from voting' *The Guardian.* https://www.theguardian.com/us-news/ 2020/ sep/28/trump-2016-campaign-targeted-35m-black-americans-to-deter-them-from-voting.
Shor, I. (1993). Education is politics: Paulo Freire's critical pedagogy. In P. Leonard, & P. McLaren (Eds.), *Paulo Freire: A Critical Encounter* (pp. 25–35). New York, NY, USA: Routledge.
Skovsmose, O. (1990). Mathematical education and democracy. *Educational Studies in Mathematics*, *21*(2), 109–128.
Skovsmose, O. (1994). Towards a critical mathematics education. *Educational Studies in Mathematics*, *27*(1), 35–57.
Steffensen, L., Herheim, R., & Rangnes, T.E. (2018). Wicked problems in school mathematics. In *Proceedings of the 42nd Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 227–234). PME.
Stephan, M., Register, J., Reinke, L., Robinson, C., Pugalenthi, P., & Pugalee, D. (2021). People use math as a weapon: Critical mathematics consciousness in the time of COVID-19. *Educational Studies in Mathematics.*
Straehler-Pohl, H. (2007). De|mathematisation and ideology in times of capitalism: recovering critical distance. In H. Straehler-Pohl, N. Bohlmann, & A. Pais (Eds.), *The disorder of mathematics education: challenging the sociopolitical dimensions of research* (pp. 35–52). Springer.
Sumpter, D. (2018). *Outnumbered From Facebook and Google to Fake News and Filter-bubbles - The Algorithms That Control Our Lives.* Bloomsbury Sigma,.
Tong, P. & Yong, F. (2015). Implementing and developing Big Data Analytics in the K-12 curriculum: A preliminary stage. Big Data and Analytics Education Conference, Puerto Rico. https://www.researchgate.net/publication/281643480_Implementing_and_Developing_Big_Data_Analytics_in_the_K-12_Curriculum-_A_Preliminary_Stage .
Villani, C., Bonnet, Y., & Rondepierre, B. (2018). For a meaningful artificial intelligence: towards a French and European strategy. *Conseil National Délután numérique*.
Wiggan, G. (2012). *Education in a strange land: Globalization, urbanization and urban schools: the social and educational implications of the geopolitical economy.* Nova Publishers,.
Williams, B., A Brooks, C., & Shmargad, Y. (2018). How Algorithms Discriminate Based on Data They Lack. In *Journal of Information Policy*, *8* pp. 78–115). Penn State University Press,. https://www.jstor.org/stable/10.5325/jinfopoli.8.2018.0078 .
Zuboff, S. (2019). *The age of surveillance capitalism.* Public Affairs, Hachette Book Group.