Gender and Sex in the Computer Graphics research literature

Ana Dodik* Meta Platforms
Silvia Sellán* University of Toronto
Theodore Kim Yale University
Amanda Phillips Georgetown University

ABSTRACT

We survey the treatment of sex and gender in the Computer Graphics research literature and its scientific and real-world consequences. We conclude current trends on the use of gender in our research community are scientifically incorrect and constitute a form of algorithmic bias with harmful effects. We propose ways for correcting these trends and pose novel research questions.

1 INTRODUCTION

References to sex and gender can be found all throughout the Computer Graphics research literature: a dataset is said to contain images of men and women, user study participants are reported to have a certain male/female ratio, a body modeling algorithm trains two different gendered models, a voice modification method is said to work on male and female voices, etc.

The scientific consensus around the concepts of sex and gender has greatly evolved in the past decades (see, e.g., [Nature Editorial Board 2018]). As surveyed by Fausto-Sterling [2012], *sex* is not one but a combination of many biological classifications (*chromosomal sex, hormonal sex, reproductive sex, ...*) which cannot be unambiguously assigned in a binary way to as much as one in 50 people [Blackless et al. 2000]. *Gender*, on the other hand, is used to refer to an individual's self-identity [Money and Ehrhardt 1972], their performance of certain acts [Butler 2003] or arbitrary social organizational structures that segregate people in different public bathrooms and even decide who can access education or participate in public life [Lorber 1994]. By all these contemporary definitions, gender is non-binary, fluid and culturally-specific. Furthermore, assuming outdated binary definitions of sex and gender is not just scientifically incorrect, but can also be shown to be harmful to those who conform the least to this artificial binary [UNHCHR 2015].

Despite this, we observe that the treatment of sex and gender in Computer Graphics research still answers to a traditional binary understanding of it that excludes intersex and many transgender and gender non-conforming people. We argue that our community's current use of gender is imprecise, contradictory and detrimental to our scientific integrity. We examine the harmful real-world consequences of the algorithmic bias introduced by our modeling choices with respect to gender on how gender non-conforming people interact with our technology in their daily lives. We advocate for reexamining our treatment of gender and show that this will not only correct worrying trends in our community, but also open the door to whole new avenues of research.

2 SURVEY

Ana: So a small suggestion on how to make this more formal, we could indicate which type of algorithmic unfairness the different papers introduce, e.g. [Friedman and Nissenbaum 1996].

*Joint First Authors

Inspired by the work of Keyes [2018], we conducted a survey of all technical papers presented at SIGGRAPH North America and SIGGRAPH Asia since 2015 (see supplemental material). We observed references to gender routinely throughout, varying in nature from demographic information reported about user study participants to gender-specific algorithms. Whenever gender is used explicitly as a variable, it is always as a binary one. Despite its prominence, gender is never given a precise definition in all the reviewed Computer Graphics literature, and appears to be used implicitly as a proxy for anything from body proportions to facial expression to voice inflection in speech. **Ana: 100% this. When it comes to body modeling, from an algorithmic stand-point, you could want to cluster based on commonly co-occuring body features, but (binary) gender is just a bad (socially-constructed) proxy for that. Also, proxies are a thing in algo. fairness research [Barocas and Selbst 2016]. I am beig told that this is apparently similar to Judith Butler's view on binary bodies in Gender Trouble?**

An analysis of the above reveals worrying trends about the current use of gender as a variable in Computer Graphics, both scientifically and ethically. As we mention examples of works that perpetuate these trends, we stress that we do not associate any malicious intent to any. **Ana: Missing word?** Rather, we wish to show how seemingly neutral, well-established practices in our community can lead to us unwittingly perpetuating forms of algorithmic bias.

Ana: There is something to be said about the entire "It's just the data thats biased" discussion we keep having in ML ad nauseam. In our examples, it's not just the data that's the root of algorithmic unfairness, it's the active decisions of algorithm designers. For practitioners this means that, even if you don't have access to, e.g. body scans of non-binary people, you can still try to remove a part of that bias that is due to algorithm design decisions.

3 SCIENTIFIC CRITIQUE

Ana: I would refrain from using the word critique. While yes, these papers are worthy of critique, it would be better to present it as an analysis?

In our strive towards producing precise, high-quality reproducible research, we should be careful about using **Ana: to use** only clearly defined variables of study **Ana: "of study" confuses me**. However, our survey shows gender to be widely used yet undefined in our literature.

In doing this, we are effectively asking our readers and fellow researchers to project their common-knowledge understanding of gender to be able to read and reproduce our results. As centuries of social science teaches us (**Silvia: citations (Amanda?)**), this understanding can vary heavily from person to person and culture to culture. Thus, different researchers will interpret and implement our algorithms differently, impeding the advance of our science.

Ana: Hmm, I honestly think the problem is that researchers share the same, but very biased view of gender.

Ana: The use of "defined" and "undefined variable" will lead to anyone who is reading this having to Google "what's an undefined variable again?" This use of gender as an undefined variable is even more detrimental when different understandings of gender are conflated, thereby biasing algorithms towards grouping certain independent attributes together. To use a very simplified example, a human parametric model may use *female* to refer to a group of people who generally are shorter *and* have longer hair, and *male* to refer to those taller *and* with shorter hair. The use of this poorly-defined variable means that the statistical distributions of hair length and height are being artificially linked together, biasing the algorithm against including shorter humans with short hair, and viceversa. **Ana: Great point. However, human parametric models don't really include hair, so we should find something similar, e.g. hip/shoulder width or the presence or absence of breasts?**

We found examples of this bias throughout the literature: voice modification algorithms may conflate voice pitch with culturally-acquired speech inflections under the umbrella of gender, virtual garment try-on methods can merge a person's body proportions with their preference in attire, a proposed new conversational agent might join a person's visual appearance with traits in their non-verbal communication.

Our field's scientific advancement is damaged further when these biases go unreported and unstudied, as we found is the case in all our reviewed literature. If a human parametric model by design cannot replicate a certain sizeable class of humans (e.g, many transgender people), or if a virtual try-on algorithm cannot allow a person with certain body proportions visualize themselves wearing a skirt, these are *scientifically* limited algorithms. Thus, these limitations should be discussed as such so that they do not unwittingly permeate through the literature and so that others can work on eliminating them.

Ana: I think, at its core, this is an algorithmic fairness paper, and I think we will make a stronger point if we look at it through that lens. For example, this paragraph makes decent points, but I think it is very helpful do disambiguate between the different types of biases that exist.

Ana: When defining a task, a paper can introduce certain problems in feature selection and target variable definition:

- **Usage of proxies, both as features and as target variables—e.g. some papers use "gender" to mean "comonly co-occuring bodily characteristics", or "gender expression".**
- **Discretization of variables—some (all? really?!) papers will split all people up in exactly two boxes.**
- **Omitted variable bias—the success of using "gender" as a feature is overemphasized because it correlates with some other important feature which has been left out of the modeling stage (e.g. long hair).**

.

Ana: Different papers also introduce problems related to data collection:

- **Sample selection bias—not including any gender-diverse people in your dataset.**
- **Encoding existing prejudice—in our society, gender expression is often conflated with gender, and an algorithm might just learn that "long hair" means woman, and "really tall" means man.**
- **Measurement bias—e.g. getting third parties to "label" people's gender from images.**

Ana: Different papers also introduce problems in the algorithm design:

- **E.g. the choice of objective function—using a regularizer that makes sure to regress to an "average" body can create problems.**

Ana: Lastly, different methods can also unknowingly introduce "impact" unfairness, i.e. the unfairness that is introduced by designers as a response to the model being tested and deployed in the real world:

- **Feedback loops—e.g. if gender-diverse people have poor experiences clothing recommender systems, they are less likely to use them. As a consequence, companies offering those systems will have less data on how their systems impact gender-diverse people. Therefore, they will continue to improve their systems for the population using their systems, possibly at the further detrement of gender-diverse people.**
- **Nudging people in different directions—e.g. a trans person might need to artificially change their voice pitch in order to not get misgendered by a voice recognition system.**

4 ETHICAL CRITIQUE

Ana: I wonder if a section title such as "Real World Impact" wouldn't be better?

As scientific researchers, we must be aware of the effect that our arbitrary modelling decisions have in the real world as our algorithms are used by governments and private companies.

Since many people's gender experiences fall outside the male/female binary, our research's insistence on it can contribute to frustration (at best) and discrimination (at worst) when they interact with technology. A researcher's seemingly inocuous decision to use different search spaces for fitting male and female body proportions leads to airport body scanners that routinely subject transgender passangers to humilliation (see [Beauchamp 2019]). A modelling choice to conflate body proportions with choices in attire ironically excludes precisely the people with non-normative bodies who are the most in danger in traditional physical changing rooms (see e.g., [Silver 2017]). **Ana: lol please explain this to my tech lead...**

These negative effects are compounded even further as our algorithms are being used to generate synthetic datasets on which to train Machine Learning algorithms outside of our research area. If we do not examine and properly report our algorithm's limitations in representing people outside of the gender binary, these can later be used to train autonomous vehicles to detect pedestrians ([Behzadi 2021]), medical diagnosing tools ([Chen et al. 2021]) and even security threat detection ([Brewer 2020]).

Furthermore, as Computer Graphics researchers, we must consider our role in shaping whose stories get to be told and who gets to seem themselves represented in the entertainment culture. By conflating different attributes under the umbrella of gender, we exclude gender non-conforming individuals from every videogame and movie created using our tool, further invisibilizing already-invisible and marginalized communities.

It bears mentioning that our research community's entrenchment in the traditional gender binary is a rare example of Computer Graphics research lagging behind the needs of our partner industries. *Metahuman*, the latest photorrealistic character modeller by Unreal Engine [2021] has no mention of gender; Google [2020] removed all gender references from its Cloud Vision API; video games as diverse as *Animal Crossing: New Horizons*, *Cyberpunk 2077* and *Forza Horizon 5* completely decouple attributes like hairstyle, body proportions, voice pitch and prononouns from one another. **Ana: I am *very* against including Cyberpunk as a positive example of anything gender-related.**

Finally, the current use of gender in the Computer Graphics literature creates a hostile environment for gender non-conforming members of our research community, which goes against ACM SIGGRAPH's goal to be *a model of inclusion, equity, access and diversity for all*: **Ana: I would end the sentence here. Also, the previous sentence might be good to include in the abstract or as one of the very first sentences.** by seeing colleagues and collaborators consistently exclude us **Ana: us → gender-diverse people** from their own research work, we are (willingly or not) sent the message that we do not belong in this research community, encouraging us to look for jobs elsewhere.

5 WHERE DO WE GO FROM HERE?

We believe the reasons above to be enough to make us reevaluate the role of gender in our community's scientific literature.

For example, the reporting of gender among other demographic information in user study participants and dataset collection subjects answer to a scientifically positive goal (experimental transparency) as well as an ethical one, to safeguard against the "male default" that plagues science and has plagued it since its infancy. However, we found instances in our survey of participants being reported as of "unknown gender", which may indicate that their gender is being assumed post facto by researchers as opposed to self reported, leading to the potential misidentification and exclusion of gender non-conforming individuals or of those from certain ethnicities (see e.g., [Buolamwini and Gebru 2018; Santamaría and Mihaljević 2018]). Therefore, we would argue it is still advisable to include this kind of data, as long as it is self reported by participants who are given a breadth of gender options not restricted to the traditional binary ones.

On the other hand, the scientific and ethical harm caused by gender-segregated algorithms is likely too significant to offset any possible benefits. At the very least, these choices should be justified and their consequences in terms of excluding gender non-conforming individuals should be examined and clearly stated. Eventually, we hope that our field evolves to address these limitations and move beyond the outdated gender binary. We trust that our fellow researchers share our scientific excitement in this new frame of reference and the potential novel research directions it opens; for example:

- What is a complete parametric model for the human body that is decoupled from gender and accurately represents the diverse bodies of all humans, regardless of whether they conform to traditional gender norms?
- How can our research inform or contrast more modern understandings of gender? Can data-based methods be used to evaluate cultural differences in gender presentation?
- How can we evaluate our algorithms for bias towards the gender binary? What tools are needed to obtain or synthesize data that covers more diverse experiences of gender?

We acknowledge that our proposed break with tradition may bring with it effort and difficult conversations, but these are challenges worth facing in the interest of scientific advancement as well as producing a fairer, more inclusive future.

REFERENCES

Solon Barocas and Andrew D. Selbst. 2016. Big Data's Disparate Impact. *California Law Review* 104, 3 (2016), 671–732. http://www.jstor.org/stable/24758720

Toby Beauchamp. 2019. *Going Stealth: Transgender Politics and U.S. Surveillance Practices.* Duke University Press, Chapel Hill, NC.

Yashar Behzadi. 2021. Synthetic data to play a real role in enabling ADAS and autonomy. *Automotive World* (2021).

Melanie Blackless, Anthony Charuvastra, Amanda Derryck, Anne Fausto-Sterling, Karl Lauzanne, and Ellen Lee. 2000. How sexually dimorphic are we? Review and synthesis. *American Journal of Human Biology: The Official Journal of the Human Biology Association* 12, 2 (2000), 151–166.

Tim Brewer. 2020. DHS Awards $1 Million to Support Machine Learning Development for Airport Security. *Synthetik Applied Technologies Blog* (2020).

Joy Buolamwini and Timnit Gebru. 2018. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency.*

Judith Butler. 2003. Gender trouble. *Continental feminism reader* (2003), 29–56.

Richard J Chen, Ming Y Lu, Tiffany Y Chen, Drew FK Williamson, and Faisal Mahmood. 2021. Synthetic data in machine learning for medicine and healthcare. *Nature Biomedical Engineering* (2021), 1–5.

Anne Fausto-Sterling. 2012. *Sex/gender: Biology in a social world.* Routledge.

Batya Friedman and Helen Nissenbaum. 1996. Bias in Computer Systems. *ACM Trans. Inf. Syst.* 14, 3 (jul 1996), 330–347. https://doi.org/10.1145/230538.230561

Google. 2020. Ethics in Action: Removing Gender Labels from Cloud's Vision API. https://diversity.google/story/ethics-in-action-removing-gender-labels-from-clouds-vision-api/. Online; accessed 20 January 2022.

Os Keyes. 2018. The misgendering machines: Trans/HCI implications of automatic gender recognition. *Proceedings of the ACM on human-computer interaction* (2018).

Judith Lorber. 1994. *Paradoxes of gender.* Yale University Press.

John Money and Anke A Ehrhardt. 1972. Man and woman, boy and girl: Differentiation and dimorphism of gender identity from conception to maturity. (1972).

A Nature Editorial Board. 2018. US proposal for defining gender has no basis in science. *Nature* 563, 7729 (11 2018), 5.

Lucía Santamaría and Helena Mihaljević. 2018. Comparison and benchmark of name-to-gender inference services. *PeerJ Computer Science* (2018).

Laura Silver. 2017. Topshop Refused To Let A Trans Person Into An All-Gender Changing Room. *BuzzFeed News* (2017).

UNHCHR. 2015. Discrimination and violence against individuals based on their sexual orientation and gender identity. (2015).

Unreal Engine. 2021. Digital Humans | Metahuman Creator. https://www.unrealengine.com/en-US/digital-humans/. Online; accessed 20 January 2022.