Gender Bias and CC

Juliana Shihadeh
Computer Science and Engineering
Santa Clara University
jshihadeh@scu.edu

Abstract

Many creative systems depend more on data nowadays than before. This data is our own making. As a result, data dependent creative machines inherit our biases. My PhD research focuses on the presence and impact of gender bias in such machines, including non-binary gender. I specifically study brilliance bias, a type of gender bias associating high levels of intellect to men. Part of my work on it also includes exploring default bias, which assumes a person is a man. In this paper I discuss my research. First, I talk about research I've done to understand the presence of brilliance and default bias in creative systems. I then present computational creative system ideas from my research to instead reduce gender bias, followed by conclusions and future work.

Understanding the Presence of Gender Bias in Creative Systems

Psychologists have proven that people more often associate high levels of intellect to men than women. This phenomenon is known as "brilliance bias," developing around age 6 (Bian, Leslie, and Cimpian 2018). As a result of it, women are less likely to apply to jobs (Bian, Leslie, and Cimpian 2018) and pursue PhDs (Leslie et al. 2015). Due to its presence it holds back many women from reaching their full potential. Fields where brilliance bias has a dominant presence include computer science, economics, and philosophy (Leslie et al. 2015). Even seemingly positive statements, such as "girls are as good as boys at math," (Shashkevich 2019) hides brilliance bias. Due to its implicit nature, it's more challenging to counter.

Both language and images influence how people think (Hibbing and Rankin-Erickson 2003; Shashkevich 2019). Furthermore, the more a person is exposed to content the more it influences their mind based on the cultivation theory (Potter 1993; Shrum 1995). Given that the arts include both text and visual artifacts, it is essential to explore both within computational creativity to help mitigate gender bias. Human-machine creative collaborations depend on models developed with large amounts of data that carry our biases (Loughran 2022). Studies I have completed demonstrate there is brilliance bias in data-dependent machines for creative writing tasks and visual interpretations of people.

One study I did revealed a significant presence of brilliance bias in the generative language model GPT (Shihadeh et al. 2022). Stories generated about highly intellectual men and women showed a difference in language choice. While men were described with more intellect, innovative ability, and excellence, women were described with a focus on their personal lives and learning. An adjective and verb analysis I ran on the stories showed men were associated to greater levels of influence and authority compared to women. This indicates the language choice describes men's intellect having more impact on others than women.

Additionally, I studied brilliance bias in generative visual models (Shihadeh and Ackerman 2023). I tested four text-to-image models, Dall-E, Midjourney, Stable Diffusion and Craiyon, on brilliance prompts. Results show images of men seen as geniuses more often than women. Furthermore, women are more often shown with objects around their head compared to men. This could indicate models add objects around a woman's head to help point out she is thinking. "Brilliant", a word defined as "full of light, shining, or bright in color" [1] in addition to high intellect, results in more images of women than men with makeup, jewels, and smiling. It was notable some images cut off women's faces too.

As a part of researching brilliance bias, I compare it to default bias in creative systems. The default bias is the assumption a person is a man (Bailey, Williams, and Cimpian 2022; Lindsay 2023). Furthermore, people with default bias assume masculinity is norm (Cheryan and Markus 2020). For example, in the study I did on brilliance bias in text-to-image models, results showed that Craiyon generated more images of men than women for "person" compared to when a brilliance trait is added, such as "Genius person" (Shihadeh and Ackerman 2023). This highlights the importance of testing if machines assume a person is male by default before testing specific biases, such as brilliance bias.

Creating Computational Creative Systems to Overcome Gender Bias

In another recent publication (Shihadeh and Ackerman), I contributed to initiating a direction of research on how creative machines can help improve gender inequality. I explore applications of CC to improve gender inequalities and guidelines for them (Shihadeh and Ackerman).

[1]https://dictionary.cambridge.org/us/dictionary/english/brilliant

Women can explore their potential through human-machine collaborations. This can help overcome concepts society influences on them of what they can or cannot be. For instance, a CC system can help a young girl imagine what kind of future self she can be, such as her occupation, and explore alternate paths. Additionally CC systems could help give writers suggestions in their works that set an example of a world with gender equality or give examples of images of more inclusive environments to include in books.

Additionally, collaborating with an artistic machine can help guide a person through an expressive and reflective process about their encounter with gender discrimination. This includes human-machine collaborations to write songs, poetry and making art on sexism they experienced. Furthermore, women could explore alternate paths they could have lived if not for gender stereotypes. This is an opportunity to provide women the space to process what they endured. Anonymous submissions of gender bias women encounter could be converted into a story or film to share.

Lastly, CC systems can create immersive experiences that cultivate empathy to help others feel what women face as a result of gender bias. For instance, a man could select a profession a woman is a minority in along with additional prompted questions to guide them toward generating art that shows what it would be like to be one of few women. For example, they can experience being one of few women amongst CEOs, engineers, or music composers. A CC system can also help demonstrate what it would be like if the world uses female pronouns as default rather than male pronouns, and help generate stories with gender bias inverted, such as having more female leaders.

As a part of my research, I worked on developing guidelines to help enrich human-machine interactions focused on overcoming gender bias. They are to 1) evoke personal connection and reflection, 2) facilitate creative engagement and 3) be constructive and non-judgemental (Shihadeh and Ackerman). Personal connection and reflection can help bring into perspective how we're each impacted by bias. Facilitating experiences allowing users to be creative in exploring gender bias can increase their engagement and in turn what they learn from the experience (Immordino-Yang, Darling-Hammond, and Krone 2019). Additionally, it is important to aim to cultivate a constructive, non-judgmental creative environment for the sensitivity of the topic.

Conclusions & Future Work

When we develop creative machine collaborators, we can unintentionally pass on our biases to them. In particular, recent advances of data dependent creative machines increase the spread of societal biases. A strong presence of gender bias is one example.

Computational creativity provides a unique opportunity to make a difference in society because of it's multiple discipline focus, including psychology, art, and technology. Gender bias and computational creativity are co-dependent. Machines represent bias in a creative way, such as showing objects around a woman's head to imply society needs proof women think. The visual and language content in computational creative artifacts will influence our minds. Due to the cultivation theory, the more people will see content that has gender bias the more they'll have gender bias too. However, if we focus on generating creative systems that help guide people toward thinking in a more inclusive way we can instead motivate more gender equality in our world.

Future work in my research involves further understanding the depth of default bias and how it impacts brilliance bias. Additionally, I'll be expanding research on non-binary gender. Lastly, I'll be exploring how to effectively implement constructive creative systems that help influence gender inclusivity.

References

Bailey, A. H.; Williams, A.; and Cimpian, A. 2022. Based on billions of words on the internet, people= men. *Science Advances* 8(13):eabm2463.

Bian, L.; Leslie, S.-J.; and Cimpian, A. 2018. Evidence of bias against girls and women in contexts that emphasize intellectual ability. *American Psychologist* 73(9):1139.

Cheryan, S., and Markus, H. R. 2020. Masculine defaults: Identifying and mitigating hidden cultural biases. *Psychological Review* 127(6):1022.

Hibbing, A. N., and Rankin-Erickson, J. L. 2003. A picture is worth a thousand words: Using visual images to improve comprehension for middle school struggling readers. *The reading teacher* 56(8):758–770.

Immordino-Yang, M. H.; Darling-Hammond, L.; and Krone, C. R. 2019. Nurturing nature: How brain development is inherently social and emotional, and what this means for education. *Educational Psychologist* 54(3):185–204.

Leslie, S.-J.; Cimpian, A.; Meyer, M.; and Freeland, E. 2015. Expectations of brilliance underlie gender distributions across academic disciplines. *Science* 347(6219):262–265.

Lindsay, J. 2023. Male default: When images of neutral characters from picture books are not so neutral. *Gender Issues* 40(1):119–143.

Loughran, R. 2022. Bias and creativity. In *ICCC*, 354–358.

Potter, W. J. 1993. Cultivation theory and research: A conceptual critique. *Human communication research* 19(4):564–601.

Shashkevich, A. 2019. The power of language: How words shape people, culture. *Stanford News*.

Shihadeh, J., and Ackerman, M. Shattering bias: A path to bridging the gender divide with creative machines.

Shihadeh, J., and Ackerman, M. 2023. What does genius Look like? An analysis of Brilliance Bias in text-to-Image models. In *Proceedings of the International Conference on Computational Creativity*.

Shihadeh, J.; Ackerman, M.; Troske, A.; Lawson, N.; and Gonzalez, E. 2022. Brilliance bias in gpt-3. In *2022 IEEE Global Humanitarian Technology Conference (GHTC)*, 62–69. IEEE.

Shrum, L. J. 1995. Assessing the social influence of television: A social cognition perspective on cultivation effects. *Communication Research* 22(4):402–429.