WEP Warwick Evans Publishing

Transactions on Social Science, Education and Humanities Research
ISSN: 2960-1770, eISSN: 2960-2262 | Volume 14
LHPPR 2025

From Epistemic Injustice to AI Ethics: A Frickerian Approach to Gender Discrimination

Zhiyuan Cheng *

School of Philosophy Fudan University, Shanghai, China

* Corresponding Author Email: 22300160037@m.fudan.edu.cn

Abstract. Testimonial Injustice occurs when the speaker is wrongfully undermined in his/her capacity as a knower due to the identity prejudice against his/her, while Hermeneutical Injustice occurs when some social groups cannot make sense of their social experiences due to their exclusion from hermeneutical participation. Both kinds of injustice result from structural social prejudice and can especially disadvantage some groups but not others. In the real world, these two types of injustice are embodied in the design and practice of AI. Based on gender-biased data, AI further exacerbates the Testimonial Injustice women suffered when processing and accessing information. And the low participation rate of women in occupations such as Science, Technology, Engineering, Mathematics results in the low diversity and inclusiveness of AI technologies and exacerbates the Hermeneutic Injustice faced by women. Therefore, in the era of AI, these two types of injustice are practiced by new technologies and manifested in the aggravation of existing gender biases against women.

Keywords: gender bias; artificial intelligence; Testimonial Injustice; Hermeneutical Injustice.

1. Introduction

In the contemporary era, artificial intelligence (AI) has permeated every corner of society, revolutionizing the way we live, work, and interact. From autonomous vehicles to large language models (LLMs), AI technologies are becoming transformative forces to reshape the future. However, beneath the surface of this technological utopia lies a complex web of challenges and ethical dilemmas.

Gender equality has long been a cornerstone of a just and progressive society. It is not only a matter of human rights but also a prerequisite for sustainable development, economic growth, and social harmony. Despite significant strides in the pursuit of gender parity over the past few decades, women around the world still face various forms of discrimination and barriers in multiple aspects of life, such as education, employment, political representation, and social participation. As AI continues to expand its reach and influence, emerging evidence suggests that AI systems are not immune to bias, and in particular, they may perpetuate and even exacerbate the existing gender biases against women. Therefore, it is of great significance to discuss how AI deepens gender biases against women and the answer to it will contribute improvement of technical ethical norms and ensure the healthy development of AI technology on the premise of respecting human rights and safeguarding women‘s rights and interests.

In recent years, papers around this topic have concentrated on „Gender by AI“, discussing how specific AI technologies can optimize gender equality (such as AI devices that can detect dangerous situations for women) [1]. However, few papers have probed into how AI discriminates against women, especially women from marginalized groups. What’s more, some scholars who have noticed the discrimination that AI brings to women focus on the fact that women have few opportunities to use artificial intelligence [2]. Other scholars have also pointed out that algorithmic bias will further worsen the unjust treatment of women from marginalized groups [3, 4]. However, they rarely use a theoretical framework to explain this, so their arguments lack persuasiveness. In Epistemic Injustice——Power and the Ethics of Knowing, by bringing light to certain ethical aspects of two of our most basic everyday epistemic practices: conveying knowledge to others by telling them, and

making sense of our own social experiences, Miranda Fricker sheds light on how epistemic interactions might have an irrepressible connection with social power, and how social disadvantage can produce unjust epistemic disadvantage, which provides new approach to understand the injustice posed by AI on women [5].

The aim of this article is to use the term of Testimonial Injustice and Hermeneutical Injustice given by Miranda Fricker to discuss why AI exacerbates existing gender biases against women. Firstly, the paper will explain the meaning of Testimonial Injustice and Hermeneutical Injustice and discuss why they are formed and how they bring harm to women. Then, this article will respectively explain how these two types of injustice are practiced by AI and manifested the aggravation of existing gender biases against women. Finally, the paper will give the conclusion and propose some ways to mitigate the problem.

2. Testimonial Injustice and Hermeneutical Injustice

Testimonial Injustice and Hermeneutical Injustice both result from structural identity prejudice and can disadvantage some groups but not others.

2.1. Testimonial In justice

Fricker thinks both two types of injustice stem from identity power. Fricker understands power as "a practically socially situated capacity to control others' actions" [5] and she thinks it leads to social coordination. Among all types of power, she identifies identity power as an operation of power that depends on universal understanding of social identity, which is formed by collective social imagination that govern, for example, what being a woman (maybe obedient) or a man (maybe powerful) means. From this understanding, gender identity power then comes in a way that women can be made deferred to a man's word due to his identity as a man. According to Fricker, this kind of power can control our actions even when those stereotypes are not truthful [5]. A woman will still do as the stereotype that women should obey their husband indicates though she doesn't agree with it. Simply put, identity power runs through shared social imagination and influences our actions even in spite of our beliefs.

Identity power entails Testimonial Injustice. According to Fricker, identity power is involved in knowledge exchange from speaker to hearer, which is called testimonial exchange, because hearers need to use social stereotypes as assessments of the speaker's credibility [5]. And the use of stereotypes may be misleading if it poses a prejudice against the speaker, because it might cause Testimonial Injustice. Fricker thinks the speaker sustains such a Testimonial Injustice if and only if she receives a credibility deficit injustice, which means she is wrongfully undermined in her capacity as a knower, due to identity prejudice in the hearer [5]. Therefore, Testimonial Injustice runs as identity power and is ethically wrong because it is a matter of one or several social groups effectively controlling the behavior of another social group—preventing them from conveying knowledge. Fricker takes In Mockingbird as an example to illustrate the ethical problems drowned by Testimonial Injustice. Due to the shared prejudiced racial stereotypes that "all Negroes lie, that all Negroes are basically immoral beings, that all Negro men are not to be trusted around our women" [5], the innocent black defendant feared to defend for himself, and the members of the jury sentenced him to death even without tangible proofs.

2.2. Hermeneutical Injustice

Fricker introduced the other type of injustice——Hermeneutical Injustice. Some social groups experience this type of injustice when some parts of their social experience, which was greatly related to their interests, were not collectively understood and remained barely intelligible [5]. For example, Carmita Wood suffered sexual harassment in her workplace, but she didn't understand the nature of this mistreatment towards her and the word 'sexual harassment' also didn't come into term [5]. After leaving work, because she was unable to describe her suffering, her words were not taken seriously

and her claim for unemployment benefits was denied. Before understanding what sex harassment is, many women were left deeply troubled, confused, and vulnerable to continued harassment. Their hermeneutical disadvantage renders them unable to make sense of their ongoing mistreatment, and this thus prevents them from protesting it, letting alone taking effective measures to stop it.

Hermeneutical Injustice is also a kind of structural discrimination. Fricker owes the formation of Hermeneutical Injustice to persistent and wide-ranging hermeneutical marginalization, indicating some social groups are excluded from hermeneutical participation that would have value for them [5]. Hermeneutical marginalization can also result from identity power. The era Fricker was in (from the late 20th century to the early 21st century) coincided with the third wave of the global feminist movement (1990s – 2010s) [5]. There were prejudicial stereotypes that women underperformed men in occupations such as academia, politics and law, and their position was marked by social powerlessness in relation to men, which prevented them from participating on equal terms with men in those practices by which collective social meanings are generated. In the case of sexual harassment, it is because women experienced prejudicial exclusion from hermeneutical participation in the spread of knowledge that many women suffered without having any hermeneutical understanding of their experience. Only when the problem was unneglectable did the word 'sexual harassment' finally come into a term.

Hermeneutical Injustice might often be worsened by Testimonial Injustice. In a society where the Hermeneutical Injustice is systematic, when some social groups subject to identity prejudice try to convey a scantly understood experience to others, their words may be warranted low credibility and not be taken seriously, further preventing them from having access to the shared hermeneutical resource.

3. Testimonial Injustice and Hermeneutical Injustice Practiced by AI

In the real world, these two types of injustice are embodied in the design and practice of AI, resulting in the aggravation of existing gender biases against women.

3.1. Testimonial Injustice Embodied in the Design and Practice of AI

Due to Testimonial Injustice, a large amount of gender-biased data has been collected by AI. As AI developers do not believe that women's opinions can bring much value, nor do they consider the relevant data representing the female group to be of great importance, when they design AI data system, they overly tilt the focus towards the data of males. From basic physiological indicators to complex behavior pattern data, the collection of female data has long been marginalized. For example, researchers at the Viterbi School of Engineering at the University of Southern California used AI technology to analyze 3,000 English books in the Project Gutenberg corpus and found that the frequency of male characters in the recommended excellent literary works was four times that of female characters [6, 7]. From the case it is indicated that as society does not believe that female authors can produce valuable literary works or news reports, researchers did not fully consider the value and characteristics of female authors' works when building AI database. As a result, AI may be biased against women when collecting and analyzing literary works. At the same time, In the field of medical AI, it has been found that the proportion of female cardiovascular disease data is less than 18%, resulting in the accuracy of AI's risk prediction for female patients being 25% lower than that for male patients. The study suggests establishing „gender-sensitive" data collection standards and requiring that the proportion of women in training data be no less than 40% [8].

Furthermore, the female group may also be reluctant to actively provide their information to the AI database because their opinions and expressions are not valued. It has been reported that female career information is not as detailed as male career information in many professional databases. Take LinkedIn as an example, although it is a globally renowned professional social platform with certain advantages in the collection and collation of professional data, there are still gender differences. Research has found that women tend to fill in their work experience more briefly on LinkedIn and

their skill descriptions are not prominent enough, and it is partly because some women are worried that excessive display will be regarded as self-promotion [9,10]. At the same time, the social recognition of women‘s professional achievements is relatively low, resulting in women lacking sufficient motivation to provide professional information and show their ambition [10]. Therefore, the female professional data in many AI databases used for workplace recruitment is relatively single and unable to fully reflect the career plan and professional level of women in the workplace. So, Testimonial Injustice, as a social power, not only makes the developers in the AI era not attach importance to the opinions and data representing the female group, but also invisibly prevents women from providing information and knowledge highly related to their interests. A large number of AI databases have information containing gender biases.

Based on such biased data, AI further exacerbates the Testimonial Injustice women suffered when processing and accessing information. When women express their views and share their experiences, the algorithms of AI may unreasonably undermine the credibility of their words due to their gender. For example, the simplicity of the data of female in the professional database directly leads to discrimination against women in the actual scenario for recruitment. Since 2014, Amazon‘s machine learning team has developed AI that could review job seekers‘ resumes and automate the recruitment of top talents [11]. Due to the biased data, the model has been biased against women in a way that when a description such as „Captain of the Women‘s International Chess Club“ that clearly indicates a female identity appears in the resume, the job seeker would be automatically downgraded [11]. As a result, bias towards women‘s professional abilities and the credibility of their words has been further strengthened. At the same time, when women report workplace bullying or unequal treatment to the human resources department, the AI to analyze employee feedback information may, due to the bias against female complaints in the data, judge female complaints as being vexatious or emotional problems, causing women‘s reasonable demands to be neglected. Therefore, female employees will continue to suffer unfairness in the workplace. The same is true in the medical field, when the intelligent diagnosis system analyzes the words of female patients describing their symptoms, due to the biased data set, they may judge women as „too sensitive“ and think they „tend to exaggerate their conditions“, and thus ignore some key symptoms, leading to misdiagnosis. In the long run, due to the Testimonial Injustice strengthened by AI, women will find it increasingly difficult to make a powerful voice and achieve fair and just expression in social communication and rights protection.

3.2. Hermeneutical injustice embodied in the design and practice of AI

Few women participate in occupations such as STEM, which is greatly needed in the development of AI. With the fast-paced development of AI, there is a high demand for positions related to the STEM (Science, Technology, Engineering, and Mathematics) fields. However, this field has long been regarded as male-dominated industries and women‘s participation in these occupations is significantly low due to educational and social factors [12]. From the educational perspective, there has been a tendency to guide women away from mathematics and science subjects. It has been showed that half of Europeans are indifferent to the absence of women in scientific disciplines and science is universally considered a boring subject with few job opportunities and dominated by men [12]. From the social perspective, it is a biased gender stereotype that an engineer should be a man, and men are more often regarded as synonymous with traits such as intelligence and strong logical thinking, which are characteristic of science and engineering disciplines. As a result, that biased concept has consolidated and strengthened men‘s status in STEM (Science, Technology, Engineering, and Mathematics) disciplines and invisibly suppressed the abilities and untapped potential of women in these fields. At the same time, in many cultures, the traditional gender division of labor expects women to take on more family responsibilities, which poses greater obstacles for women when they pursue career development in STEM fields. For example, Mauritius is an island country located in the Indian Ocean, and its culture is a blend of Indian, African, European, and Chinese immigrant traditions. Among them, Indian culture has a particularly significant impact on gender roles that women are expected to take on major household and childcare responsibilities, while men are regarded as the breadwinners of the family [13,14]. This cultural norm has led to a long-term lower

participation rate of women in the STEM field than men. And when women enter the STEM fields, they inevitably have to bear more pressure and burden.

The low participation rate of women in occupations such as STEM leads to their hermeneutic marginalization. In the field of STEM, since men have a higher proportion in the technology field, they usually find it easier to master and develop the latest AI technologies. On the contrary, due to the relatively small number of female developers, their experiences, viewpoints, and insights are difficult to integrate into the mainstream knowledge system and industry discourse, resulting in their lack of voice in the development and decision-making processes of AI technologies. Therefore, female researchers always find it hard to get in hermeneutical participation to shape the core knowledge and concepts of AI (such as the Large Language Model and Blackbox), further deepening their degree of hermeneutic marginalization.

Hermeneutic marginalization affects the diversity and inclusiveness of AI technologies and exacerbates the Hermeneutic Injustice faced by women. Many women without a scientific and technological background don't know what AI is and how to use it [5,6]. It has been found that female researchers use AI tools 7% less frequently compared to male researchers [2]. What's more, among rural women in India, only 12% have been exposed to AI education [15], and it has been showed that women have 32% higher resistance to AI tools than men [16]. Without a comprehensive understanding of AI, women encounter great inconveniences and difficulties in their daily lives and have limited access to public knowledge. Moreover, many products of AI technologies do not fully consider the needs of female users, and women often fail to identify the ethical problems such as privacy, gender biases, and transparency behind these technologies, and even if they do, they are unable to change the situation. For example, some apps such as „Virtual Girlfriend" are designed to provide emotional companionship for male users. However, they excessively output vulgar images of women, ignoring women's dignity and rights. Some health-related AI products may overlook women's specific health problems during the design process, resulting in these products being unable to fully meet women's needs. Unable to recognize and point out the ethical issues existing in these technologies unfriendly to women, women fail to have their interests protected.

4. Conclusion

Testimonial Injustice occurs when the speaker is wrongfully undermined in his/her capacity as a knower due to the identity prejudice against his/her, while Hermeneutical Injustice occurs when some social groups cannot make sense of their social experiences due to their exclusion from hermeneutical participation. Both kinds of injustice result from structural social prejudice and can especially disadvantage some groups but not others. In the real world, these two types of injustice are embodied in the design and practice of AI. Based on gender-biased data, AI further exacerbates the Testimonial Injustice women suffered when processing and accessing information. And the low participation rate of women in occupations such as Science, Technology, Engineering, Mathematics results in the low diversity and inclusiveness of AI technologies and exacerbates the Hermeneutic Injustice faced by women. Therefore, in the era of AI, these two types of injustice are practiced by new technologies and manifested in the aggravation of existing gender biases against women.

All in all, reducing gender bias in AI technology for women requires systematic solutions that cover multi-dimensional collaborative improvements such as data, algorithms, teams, and social policies. For example, it is necessary to ensure that training data covers female samples from different cultural backgrounds and occupations to avoid the solidification of traditional gender roles. At the same time, tools are needed to detect and eliminate the associations of stereotypes in the data. What's more building an interdisciplinary team and introducing female developers and ethicists to participate in decision-making is also of great significance to mitigate the Hermeneutical Injustice suffered by women. There is still a long way to go to achieve gender fairness in the era of AI.

References

[1] Patón-Romero, J. David, et al. "State of gender equality in and by artificial intelligence." IADIS International Journal on Computer Science and Information Systems 17.2 (2022): 31-48.

[2] Dorta-González, Pablo, et al. "Generative artificial intelligence usage by researchers at work: Effects of gender, career stage, type of workplace, and perceived barriers." Telematics and Informatics 94 (2024): 102187.

[3] Huang, Linus Ta-Lun, et al. "Ameliorating algorithmic bias, or why explainable AI needs feminist philosophy." Feminist Philosophy Quarterly 8.3/4 (2022).

[4] Buolamwini, Joy, and Timnit Gebru. "Gender shades: Intersectional accuracy disparities in commercial gender classification." Conference on fairness, accountability and transparency. PMLR, 2018.

[5] Fricker, Miranda. Epistemic injustice: Power and the ethics of knowing, Oxford University Press, New York, 2007, pp.15-60, 147-175.

[6] Nagaraj, Akarsh, and Mayank Kejriwal. "Dataset for studying gender disparity in English literary texts." Data in Brief 41 (2022): 107905.

[7] Nagaraj, Akarsh, and Mayank Kejriwal. "Robust quantification of gender disparity in pre-modern english literature using natural language processing." arXiv preprint arXiv:2204.05872 (2022).

[8] Cirillo, Davide, et al. "Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare." NPJ digital medicine 3.1 (2020): 81.

[9] Altenburger, Kristen, et al. "Are there gender differences in professional self-promotion? an empirical case study of linkedin profiles among recent mba graduates." Proceedings of the international AAAI conference on web and social media. Vol. 11. No. 1. 2017.

[10] Moore, Sarah E. "The language of LinkedIn: Popular publications, the gender gap, and pedagogy." Business and Professional Communication Quarterly 82.4 (2019): 401-417.

[11] Chang, Xinyu. "Gender Bias in Hiring: An Analysis of the Impact of Amazon's Recruiting Algorithm." Advances in Economics, Management and Political Sciences 23 (2023): 134-140.

[12] VanHeuvelen, Tom, and Natasha Quadlin. "Gender inequality in STEM employment and earnings at career entry: Evidence from millennial birth cohorts." Socius 7 (2021): 23780231211064392.

[13] Tandrayen-Ragoobur, Verena, and Deepa Gokulsing. "Gender gap in STEM education and career choices: what matters?." Journal of Applied Research in Higher Education 14.3 (2022): 1021-1040

[14] Sachdev, Aditi Rabindra. "Gender disparity in STEM across cultures." Industrial and organizational psychology 11.2 (2018): 309-313.

[15] Kumar, Shailendra, and Sanghamitra Choudhury. "Gender and feminist considerations in artificial intelligence from a developing-world perspective, with India as a case study." Humanities and Social Sciences Communications 9.1 (2022): 1-9.

[16] Fauville, Geraldine, et al. "Video-conferencing usage dynamics and nonverbal mechanisms exacerbate Zoom Fatigue, particularly for women." Computers in Human Behavior Reports 10 (2023): 100271.