scientific reports

OPEN

Humor as a window into generative AI bias

Roger Saumure[1✉], Julian De Freitas[2] & Stefano Puntoni[1]

A preregistered audit of 600 images by generative AI across 150 different prompts explores the link between humor and discrimination in consumer-facing AI solutions. When ChatGPT updates images to make them "funnier", the prevalence of stereotyped groups changes. While stereotyped groups for politically sensitive traits (i.e., race and gender) are less likely to be represented after making an image funnier, stereotyped groups for less politically sensitive traits (i.e., older, visually impaired, and people with high body weight groups) are more likely to be represented.

Recent years have witnessed two major trends. The first is the increasing use and availability of generative artificial intelligence (AI) commercial systems (e.g., ChatGPT, Gemini). By early 2023, for example, it was estimated that OpenAI's ChatGPT had accumulated over 100 million monthly users[1]. The second trend is the enhanced interoperability between generative AI models, whereby large language models (LLMs) and image generators can readily communicate to generate or modify images. For example, a user might type a prompt into ChatGPT like: "*Create an image of someone reading a book*". The LLM GPT4 processes this prompt and expands it into a much more detailed prompt: "*An image of a person sitting in a cozy, well-lit room, deeply engrossed in reading a book. They are comfortably seated in an armchair with a warm blanket draped over their legs. The room is filled with soft, ambient lighting, creating a serene atmosphere…*" that is used to guide the text-to-image model DALL-E3 to generate an image.

How do AI models communicate and what are the potential consequences of these interactions? In this research, we explore these questions by examining how text-to-image generators and LLMs interact to create images. Specifically, through a preregistered audit of 600 images created by a commercial generative AI, we find that using ChatGPT to update images by making them "funnier" increases the prevalence of some stereotyped groups: people with high body weight[2], older, and visually impaired. Our results suggest that these observed biases are primarily attributable to the text-to-image model rather than the language model.

We study the intersections between AI image generation, bias, and humor for both theoretical and substantive reasons. First, to our knowledge, no research has examined humor and bias in the context of AI-generated images. Research on the relationship between humor and prejudice is a large area of study in psychology, sociology, and communication[3–5]. A common finding of this literature is that humor can reinforce prejudice and stereotypes. Indeed, prior work has shown that humor can perpetuate prejudiced attitudes by normalizing derogatory stereotypes and facilitating their expression in socially acceptable ways[6–9]. For instance, Bill and Naus found that humorously framing sexist incidents increased the acceptability of sexism[10]. Similarly, Katz and Wing-Paul found that humorously insulting people with intellectual disabilities increased the acceptability of ableism and decreased participants' likelihood of confronting the insulter[11]. As such, humor could act as a window into whether generative AI can exacerbate prejudice and stereotypes, carrying both theoretical and practical implications.

While it is true that some types of humor, like satirical humor, can also undermine stereotypes by exposing the absurdity of prejudiced beliefs and encouraging audiences to question and reject harmful stereotypes[12–14], here we explore a form of humor that is more likely to exacerbate prejudices. Specifically, we ask whether asking generative AI to "make images funnier" is more likely to give rise to groups that are typically the target of prejudice, e.g., the initial image is a man with average weight, whereas the 'funnier' image is a man with high body weight. Because the target of the humor is the traditionally prejudiced group, this instance of humor can reinforce harmful stereotypes about, say, high body weight individuals. This, in turn, can perpetuate prejudice against such people, such as fatphobia, contributing toward real-world discrimination and social exclusion. In short, the humor is most likely to be interpreted as "punching down", making fun of a group that already faces prejudice, rather than "punching up", targeting systems of power or shared human experiences that marginalize these groups[15,16].

Second, whereas considerable efforts have been made to mitigate bias from algorithms in two dimensions (race and gender)[17–21], we find evidence of bias in other dimensions that might have been overlooked (i.e.,

[1]Department of Marketing, The Wharton School, University of Pennsylvania, Philadelphia, PA, USA. [2]Department of Marketing, Harvard Business School, Harvard University, Boston, MA, USA. ✉email: saumure@wharton.upenn.edu

people with high body weight, older, and visually impaired groups). As such, our research is one of the first to raise awareness of new forms of bias that arise from the interactions between LLMs and text-to-image models.

Method

The audit study was preregistered (https://aspredicted.org/WZB_WG9) and all data and code are publicly available (https://osf.io/MR5SU/). We carried out an audit of 600 images created by a commercial generative AI model. We collected data over two stages, with the help of four research assistants (RAs) blind to our research question.

In the first stage, two RAs were given the same set of 150 text prompts (describing a human carrying out an action) to manually input into the generative AI (each RA had a unique account), resulting in 150 images per RA. After generating the first set of images, the RAs modified the images by asking the model to make them "funnier"—Fig. 1. This resulted in a second set of 150 images, resulting in 600 images total.

Importantly, to determine whether potential biases arose due to the LLM or text-to-image model, we also collected the detailed textual descriptors internally generated by GPT4 to guide DALL-E3 for image creation, given that, as we note above, it generates its own more detailed prompt for image generation. This was done by asking GPT4 "What prompt did you use to generate this image?" for each pair of images, resulting in a total of 600 internal prompts.
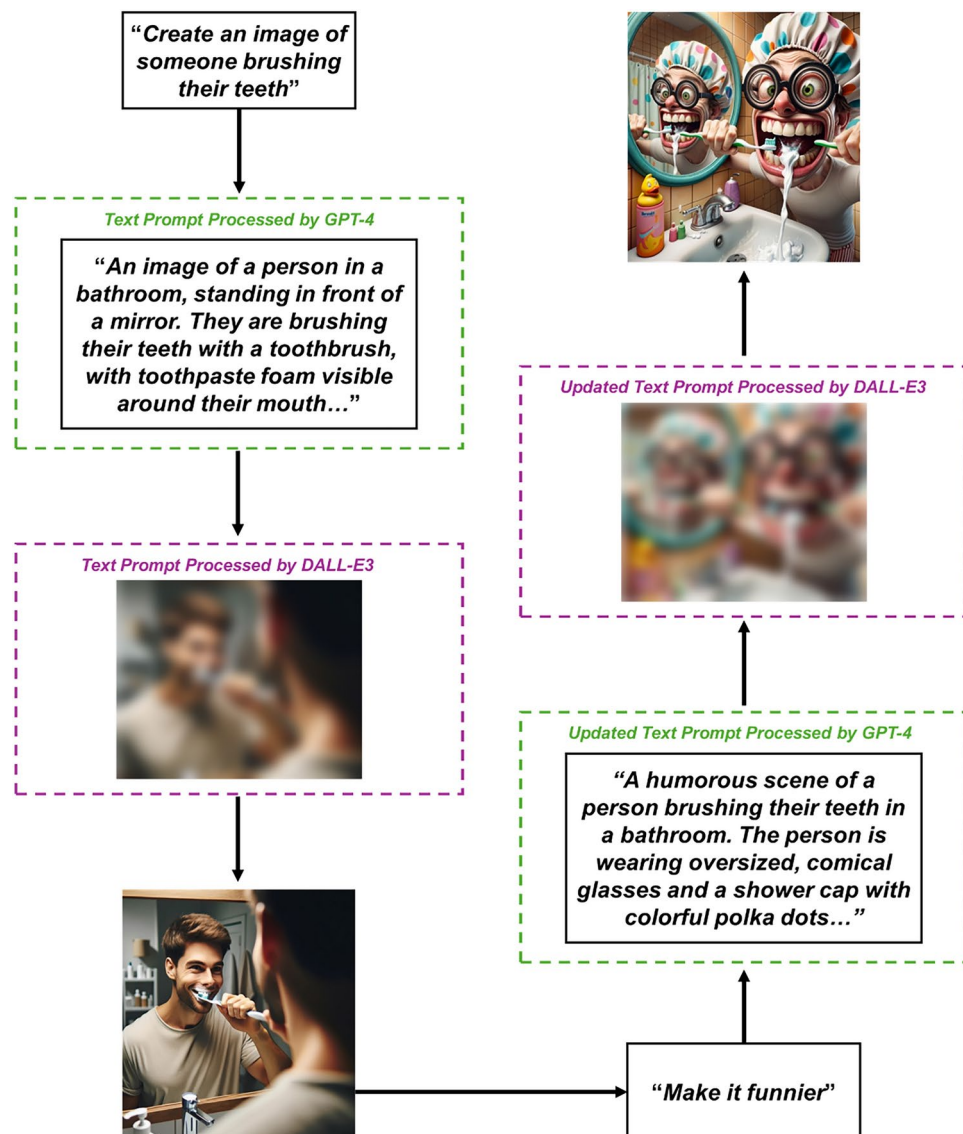

Fig. 1. Generative AI Image Creation. *The images were generated using ChatGPT.* A textual prompt is processed by the LLM GPT-4 to create an internal representation of the image that needs to be generated. This representation includes details about the layout, color scheme, and other visual elements that align with the textual description. Once the internal representation is ready, the image generator DALL-E3 uses an image-to-text model to produce the image.

Per our preregistration, we excluded images that met our exclusion criteria to ensure the reliability of our coding. We excluded images that featured non-human entities (e.g., animals) or more than one human. If a trait(s) was not discernible from an image, we excluded the image only for the measure of bias that could not be determined from the image. For example, if an image depicted a female but their race could not be determined, we excluded the image for the race measure but retained it for the gender measure.

In the second stage, two separate, hypothesis-blind research assistants participated in a qualitative coding task. They dummy coded (1, 0) the set of images across five dimensions: race (White, Black, Asian, Hispanic), gender (male or female), eyesight (no glasses or glasses), bodyweight (low body weight or high body weight), and age (young/middle- or old-aged). After completing the qualitative coding task, RAs met to resolve any disagreements.

Results

We analyzed the data using three analyses: (1) an equal-weighted omnibus bias analysis, (2) a severity-weighted omnibus bias analysis, and (3) trait-specific bias analyses. The omnibus and severity-weighted analyses allowed us to capture the overall presence and magnitude of bias across all dimensions in generative AI output, whereas the trait-specific bias analyses allowed us to capture the presence and magnitude of bias in specific dimensions (e.g., race, gender, bodyweight).

The equal-weighted omnibus bias estimate was computed using a pre-established coding scheme. For each dimension, we determined if there was bias toward stereotyping across the two images (coded as 1, e.g., White → minority), away from it (coded as -1, e.g., minority → White), or no bias (coded as 0, e.g., White → White, or minority → minority). Results of the equal-weighted omnibus bias analysis revealed a mean bias toward stereotyping in the AI-generated images ($M = 0.39$, $SE = 0.05$, $t(264) = 7.25$, 95% CI [0.286, 0.499], $p < 0.001$, $d = 0.45$).

The severity-weighted omnibus bias estimate was similar but computed based on a pretest (N = 300, recruited from Prolific; $M_{age} = 39$; 50% female; 67% White). The pretest asked participants how much they thought policy makers should be concerned about bias toward the groups we were examining: "In your opinion, how much should policy makers be concerned about discrimination towards [group]" (0 = "not at all", and 100 = "a great deal"). We created this estimate because recent polls have shown that Americans weigh certain biases (e.g., racial bias towards Black vs. Asian people) differently[22]. These weights were used to compute a weighted bias measure across the five dimensions. For instance, if participants rated concern about discrimination towards women an 80 out of 100 and towards men a 20 out of 100, then the weights would be 0.8 and 0.2, respectively. The bias towards gender stereotyping (female → male) would then be coded as 0.6 (0.8–0.2). We acknowledge that alternative methods (e.g., prior research, expert opinion) could be used to capture the severity of different biases; we chose the current approach to gain a sense of the general public's concern, since members of the public are users of the technology and a key stakeholder for both generative AI companies and policy makers. Importantly, our main findings remain consistent even when not applying these weights. Indeed, just like the equal-weighted analysis, we found a mean bias toward stereotyping in AI-generated images in the severity-weighted omnibus bias analysis ($M = 0.09$, $SE = 0.01$, $t(264) = 8.65$, 95% CI [0.0729, 0.116], $p < 0.001$, $d = 0.531$).

To identify the presence of bias across specific dimensions, we estimated four preregistered multilevel logistic regression models. Specifically, we regressed the presence of each trait (coded as a binary variable) on the image version (original vs. funnier), including prompt as a random intercept (Table 1). Due to complete separation in the multilevel logistic regression for the bodyweight dimension, we used an OLS regression with robust standard errors clustered by prompt. We found significant differences between the two sets of images (first and funnier version) for all dimensions. The bias was strong and in the predicted direction for the dimensions of age ($\beta = 2.93$, $SE = 0.62$, $z = 4.74$, 95% CI [1.72, 4.11], $p < 0.001$), bodyweight ($\beta = 0.095$, $SE = 0.018$, $t(523) = 5.19$, 95% CI [0.059, 0.13], $p < 0.001$), and eyesight ($\beta = 2.83$, $SE = 0.42$, $z = 6.69$, 95% CI [2.00, 3.65], $p < 0.001$) dimensions. However, we found the opposite effect for the race ($\beta = -1.32$, $SE = 0.52$, $z = -2.53$, 95% CI [−2.34, −0.30], $p = 0.011$) and gender ($\beta = -0.88$, $SE = 0.38$, $z = -2.28$, 95% CI [−1.63, −0.12], $p = 0.022$) dimensions—Fig. 2a. So, prompting the model to make the image funnier led to older, heavier, and more visually impaired subjects, but fewer minority race and female subjects.

We speculate that this surprising trend might be related to relative differences in the political sensitivity of biases across different dimensions. To provide initial evidence for this argument, we created a "political sensitivity" dummy variable based on a separate group of participants' responses in a preregistered study (N = 100, recruited from Prolific; $M_{age} = 40$; 55% female; 67% White; https://aspredicted.org/WR4_X6H). Specifically, we asked participants to indicate how concerned they believed businesses would be if they were accused of bias

	Race	Gender	Bodyweight	Eyesight	Age
Funnier	− 1.32*	− 0.88*	0.095***	2.83***	2.93***
	(0.52)	(0.38)	(0.018)	(0.42)	(0.62)
Intercept	− 2.61***	− 2.91***	0.00	− 3.63***	− 4.77***
	(0.26)	(0.44)	(0.00)	(0.45)	(0.69)
Observations	491	515	525	462	509

Table 1. Regression results examining the effect of making an image funnier on the presence of several traits (race, gender, eyesight, bodyweight, age). Standard errors in parentheses. Significance levels: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.
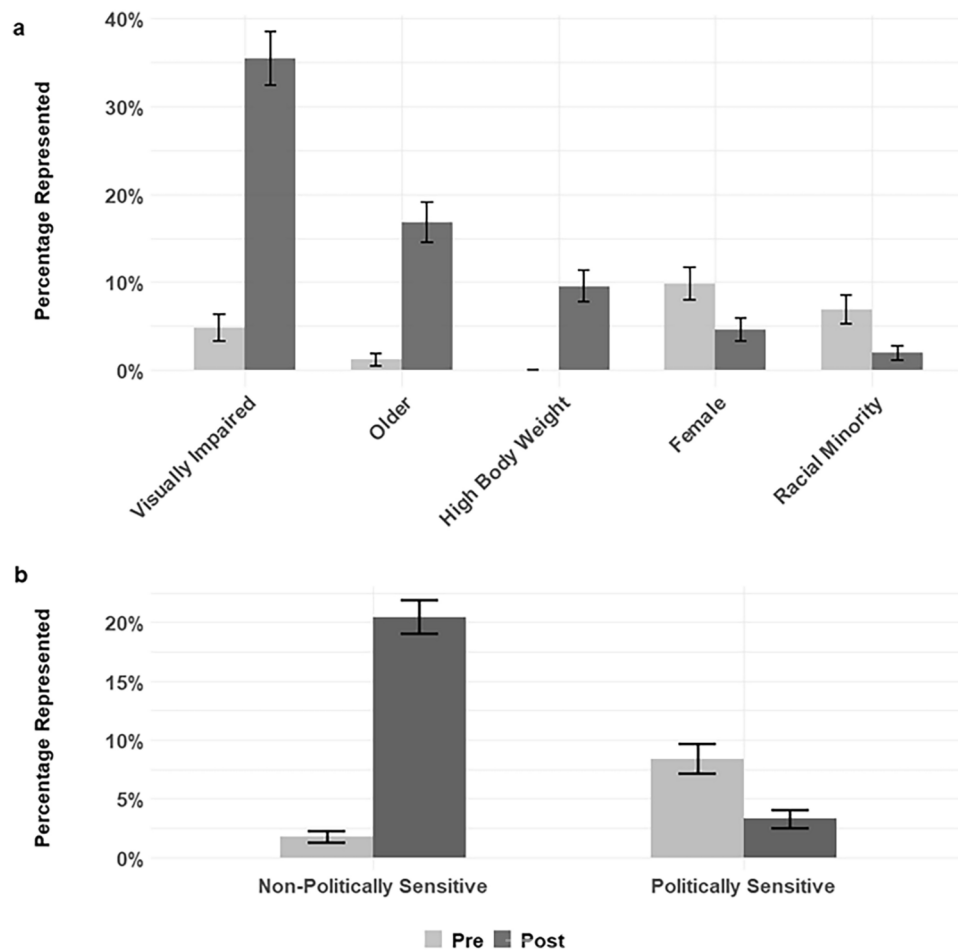
Fig. 2. Percentage of Minority Groups Represented. (**a**) The percentage of minority groups (visually impaired, older, people with high body weight, racial minorities, female) represented before and after making the images funnier. Error bars represent standard errors of proportions. (**b**). The percentage of politically sensitive (racial minorities and female) and non-politically sensitive (visually impaired, people with high body weight, and older) groups represented after making the images funnier. Error bars represent standard errors of proportions. Politically sensitive groups were less likely to be represented after making the images funnier, whereas non-politically sensitive groups were more likely to be represented after making the images funnier.

against different kinds of groups (e.g., based on racial or bodyweight): "In your opinion, how concerned would businesses be if they were accused of [group] bias?" (0 = "not at all concerned" and 100 = "very concerned").

Consistent with prior research on prejudice and stereotyping[23–26], we predicted that participants would indicate businesses are more concerned about accusations of racial and gender bias than biases related to age, bodyweight, or eyesight. We focused on participants' perceptions of how concerned businesses would be, rather than on their own beliefs or that of society, under the assumption that companies offering generative AI are most likely to respond to consumers' perceptions of the company. As such, companies may prioritize correcting politically sensitive biases (race and gender) due to higher perceived risks of public backlash or legal repercussions, while inadvertently neglecting less politically sensitive biases (age, bodyweight, and eyesight).

As predicted, we found that bias based on race and gender (M = 80.00) was rated as much more politically sensitive than bias based on age, bodyweight, or eyesight (M = 61.20; $\beta = 18.80$, SE = 1.84, $t(399) = 10.25$, 95% CI [15.20, 22.40], $p < 0.001$). This effect remained significant when controlling for participants' self-reported race, gender, and age ($\beta = 18.80$, SE = 1.83, $t(399) = 10.25$, 95% CI [15.20, 22.40], $p < 0.001$). Based on these results, we then conducted an additional (not preregistered) analysis on the image data. In this analysis, the race and gender dimensions were coded as "politically sensitive," and the bodyweight, eyesight, and age dimensions were coded as "not politically sensitive." We modeled the interaction between political sensitivity and our focal independent variable (making the image funnier) by estimating a multilevel logistic regression. Specifically, we regressed the presence of each trait on the image version (original vs. funnier), political sensitivity (politically sensitive vs. not politically sensitive), and their interaction. We included prompt as a random intercept. As predicted, we found a significant interaction between political sensitivity and image version on minority group representation ($\beta = 3.67$, SE = 0.42, $z = 8.79$, 95% CI [2.85, 4.49], $p < 0.001$). Whereas politically sensitive groups were less likely to be represented after making an image funnier ($\beta = -1.00$, SE = 0.30, $z = -3.38$, 95% CI [−1.58, −0.42],

$p < 0.001$), non-politically sensitive groups were more likely to be represented after making an image funnier ($\beta = 2.67$, SE $= 0.29$, $z = 9.07$, 95% CI [2.09, 3.24], $p < 0.001$)—Fig. 2b.

Finally, to determine whether the bias originated from the language model or text-to-image model, we carried out an additional analysis that examined the presence of textual descriptors along the five bias dimensions in the textual prompts that GPT4 sent to DALL-E3 (obtained by asking GPT4, "What prompt did you use to generate this image?"). We did not find evidence of bias in the language model in this additional analysis except for in the eyesight dimension ($\chi^2(1) = 34.04$, $p < 0.001$), whereby the proportion of the word "glasses" was higher in the updated textual descriptors (17.74% vs. 2.26%). By a process of elimination, this suggests that most of the bias in the images originates from the image model, not the language model. This pattern is consistent with researchers' and commentators' remarks about the greater difficulty of assessing bias in image as opposed to language models[27,28].

Discussion

Taken together, our research shows that using ChatGPT to update images by making them "funnier" increases the prevalence of certain stereotyped groups, while decreasing the prevalence of others. Specifically, we find evidence of bias against minorities for less politically sensitive dimensions (age, bodyweight, and eyesight) and evidence against majorities for more politically sensitive dimensions (race and gender). The presence of bias against older, heavier and visually impaired people is concerning since past work in psychology and sociology has shown that such forms of 'punching down' humor can exacerbate stereotypes[5,29] and the downstream prejudices of fat shaming, ageism, and ableism. Whereas practitioners have made considerable efforts to mitigate algorithmic biases related to race and gender, biases across other dimensions—such as age, body weight, and disability—have received less attention. With that said, a growing body of research has begun exploring these areas[30–33]. For instance, Herold et al. assessed disability bias in pre-trained natural language processing (NLP) models underlying AI-based assistive technologies, finding significant associations of disability with lower warmth and competence[32]. Similarly, Arseniev-Koehler and Foster investigated how machine learning models internalize societal schemas related to body weight, revealing biases against individuals with high body weight[30]. Our work contributes to this emerging literature by identifying similar patterns of biases in AI-generated images arising from interactions between LLMs and text-to-image models.

It is challenging to determine the cause of these findings. Since humans are known to exhibit bias on all dimensions, one possibility is that app makers, through their efforts (e.g., dataset curation, fine-tuning, and filtering), have selectively corrected for racial and gender bias. These particular minority groups might garner more attention and advocacy because they are more salient in Western society, for historical, cultural, and political reasons[34]. If this interpretation is correct, then LLMs trained on corpora in other languages may exhibit different patterns of bias. For example, we should expect LLMs that generate images based on prompts in Hindi to be more corrected against biases against Muslims (the largest religious minority in India), given the more salient tension between Hindus and Muslims in India.

Although we document a new pattern of bias in generative AI, our results raise several implications and directions for future research. Notably, we observed significant underrepresentation of discriminated groups in the original images before any modifications were made. For example, the baseline proportions of high body weight and female individuals featured in the original images were 0% and 9.80%, respectively—a severe underestimation of the national averages of 73.60%[35] and 50.50%[36]. Such underrepresentation is itself a bias in AI-generated images that could be just as problematic as the biases introduced when making the images "funnier". This lack of representation may perpetuate stereotypes by reinforcing default assumptions about what is normal or representative of society (e.g., white men of low body weight), making other traits seem less normal and representative[37].Thus, we encourage future research to investigate whether such underrepresentation is observed across different AI models and to address whether the absence of certain groups leads to increased stereotyping or perpetuation of default assumptions.

We also encourage future research to directly examine whether specific image features (e.g., complexity, spatial composition, lighting) systematically change after modifying the images. For instance, the modified image in Fig. 1 is more colorful, cartoon-like, and complex than the original. Are such image transformations systematic and, if so, do they further perpetuate stereotypes—say, by enhancing the negative impact of the humorized depiction of a stereotyped group? Are attitudes toward certain groups more affected by such image transformations than others? While these questions are not the primary focus of the current research, we believe that they merit a formal investigation.

Finally, our findings open avenues for future research to explore the downstream consequences of our findings. As we noted, while prior studies suggest that humor can exacerbate stereotypes and prejudices[7–9], other work indicates the opposite—that humor can challenge and even undermine stereotypes[12–14]. For instance, Zimbardo posits that humor can challenge stereotypes toward Muslim communities by highlighting the absurdity of these stereotypes[14]. Similarly, Borgella et al. find that humor can mitigate prejudice by reducing anxiety in interracial interactions and fostering dialogue between different racial groups[12]. We have pointed out why in this case the absurd, cartoon-like images produced when making an image "funnier" are more likely to reinforce than undermine stereotypes, although this remains an empirical question.

Ultimately, since these generative models are widely utilized for many purposes, and because their potential to perpetuate prejudice is extended through increasing interoperability, we believe all dimensions of prejudice deserve attention and advocacy from the public, policymakers and corporations. Stakeholders can strive for balance in all dimensions, rather than bias in any direction.

Data availability

All data and code are publicly available (https://osf.io/MR5SU/).

References

1. Porter, J. ChatGPT continues to be one of the fastest-growing services ever. *The Verge* https://www.theverge.com/2023/11/6/23948386/chatgpt-active-user-count-openai-developer-conference (2023).
2. Puhl, R. M. Weight stigma, policy initiatives, and harnessing social media to elevate activism. *Body Image* **40**, 131–137 (2022).
3. Keith-Spiegel, P. Early conceptions of humor: Varieties and issues. In *The Psychology of Humor* (eds Goldstein, J. H. & McGhee, P. E.) 3–39 (Academic Press, 1972).
4. Mickes, L., Walker, D. E., Parris, J. L., Mankoff, R. & Christenfeld, N. J. S. Who's funny: Gender stereotypes, humor production, and memory bias. *Psychonomic Bull. Rev.* **19**, 108–112 (2011).
5. Zillmann, D. Disparagement humor. In *Handbook of Humor Research* (eds McGhee, P. E. & Goldstein, J. H.) 85–107 (Springer, 1983).
6. Argüello Gutiérrez, C., Carretero-Dios, H., Willis, G. B. & Moya, M. Joking about ourselves: Effects of disparaging humor on ingroup stereotyping. *Group Process. Intergroup Relat.* **21**, 568–583 (2018).
7. Ferguson, M. A. & Ford, T. E. Disparagement humor: A theoretical and empirical review of psychoanalytic, superiority, and social identity theories. *HUMOR* **21**, 283–312 (2008).
8. Ford, T. E. & Ferguson, M. A. Social consequences of disparagement humor: A prejudiced norm theory. *Pers. Soc. Psychol. Rev.* **8**, 79–94 (2004).
9. Kirk, M. R. The role of humor in the social construction of gendered and ethnic stereotypes. *Race Gender Class* **9**, 76–95 (2002).
10. Bill, B. & Naus, P. The role of humor in the interpretation of sexist incidents. *Sex Roles* **27**, 645–664 (1992).
11. Katz, J. & Wing-Paul, D. Taking a joke seriously: When does humor affect responses to the slurring of people with intellectual disabilities. *HUMOR* **33**, 563–579 (2020).
12. Borgella, A. M., Howard, S. & Maddox, K. B. Cracking wise to break the ice: The potential for racial humor to ease interracial anxiety. *HUMOR* **33**, 105–135 (2020).
13. Strain, M. L., Martens, A. L. & Saucier, D. A. Rape is the new black: Humor's potential for reinforcing and subverting rape culture. *Transl. Issues Psychol. Sci.* **2**, 86–95 (2016).
14. Zimbardo, Z. Cultural politics of humor in (de)normalizing Islamophobic stereotypes. *Islamophobia Stud. J.* **2**, 59–81 (2014).
15. Ford, T. E., Richardson, K. & Petit, W. E. Disparagement humor and prejudice: Contemporary theory and research. *HUMOR* **28**, 171–186 (2015).
16. Hodson, G., Rush, J. & MacInnis, C. C. A joke is just a joke (except when it isn't): Cavalier humor beliefs facilitate the expression of group dominance motives. *J. Pers. Soc. Psychol.* **99**, 660–682 (2010).
17. Ananya. AI image generators often give racist and sexist results: Can they be fixed. *Nature* **627**, 722–725 (2024).
18. Chen, Z. Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Hum. Soc. Sci. Commun.* **10**, 567 (2023).
19. Gupta, M., Parra, C. M. & Dennehy, D. Questioning racial and gender bias in AI-based recommendations: Do espoused national cultural values matter?. *Inf. Syst. Front.* **24**, 1465–1481 (2022).
20. O'Connor, S. & Liu, H. Gender bias perpetuation and mitigation in AI technologies: Challenges and opportunities. *AI Soc.* **39**, 2045–2057 (2024).
21. Varsha, P. S. How can we manage biases in artificial intelligence systems–a systematic literature review. *Int. J. Inf. Manag. Data Insights* **3**, 100165 (2023).
22. Daniller A. Majorities of Americans see at least some discrimination against Black, Hispanic and Asian people in the U.S. *Pew Research Center* https://www.pewresearch.org/short-reads/2021/03/18/majorities-of-americans-see-at-least-some-discrimination-against-black-hispanic-and-asian-people-in-the-u-s/ (2021).
23. Crandall, C. S., Eshleman, A. & O'Brien, L. Social norms and the expression and suppression of prejudice: The struggle for internalization. *J. Pers. Soc. Psychol.* **82**, 359–378 (2002).
24. Fiske, S. T. Stereotyping, prejudice, and discrimination. In *The Handbook of Social Psychology* (eds Gilbert, D. T. et al.) 357–411 (McGraw-Hill, 1998).
25. North, M. S. & Fiske, S. T. An inconvenienced youth? Ageism and its potential intergenerational roots. *Psychol. Bull.* **138**, 982–997 (2012).
26. Puhl, R. M. & Heuer, C. A. The stigma of obesity: A review and update. *Obesity* **17**, 941–964 (2009).
27. Cevik, J. et al. Assessment of the bias of artificial intelligence generated images and large language models on their depiction of a surgeon. *ANZ J. Surg.* **94**, 287–294 (2024).
28. Wang W, Bai H, Huang J-t, Wan Y, Yuan Y, Qiu H, Peng N & Lyu MR. New job, new gender? Measuring the social bias in image generation models. *arXiv preprint* arXiv:2401.00763 (2024).
29. Zenner, W. P. Joking and ethnic stereotyping. *Anthropol. Quart.* **2**, 93–113 (1970).
30. Arseniev-Koehler, A. & Foster, J. G. Machine learning as a model for cultural learning: Teaching an algorithm what it means to be fat. *Sociol. Methods Res.* **51**, 1484–1539 (2022).
31. Charlesworth, T. E. S., Sanjeev, N., Hatzenbuehler, M. L. & Banaji, M. R. Identifying and predicting stereotype change in large language corpora: 72 groups, 115 years (1900–2015), and four text sources. *J. Pers. Soc. Psychol.* **125**, 969–990 (2023).
32. Herold, B., Waller, J. & Kushalnagar, R. S. Applying the Stereotype Content Model to assess disability bias in popular pre-trained NLP models underlying AI-based assistive technologies. in *Proceedings of the 9th Workshop on Speech and Language Processing for Assistive Technologies (SLPAT 2022)* 58–65 (2022).
33. Zorrilla-Muñoz, V., Moyano, D. L., Marcos Carvajal, C. & Agulló-Tomás, M. S. Towards equitable representations of ageing: Evaluation of gender, territories, aids and artificial intelligence. *Land* **13**, 1304 (2024).
34. Kardosh, R., Sklar, A. Y., Goldstein, A., Pertzov, Y. & Hassin, R. R. Minority salience and the overestimation of individuals from minority groups in perception and memory. *Proc. Natl. Acad. Sci. USA* **119**, e2116884119 (2022).
35. Centers for Disease Control and Prevention. FastStats - Overweight Prevalence. *Centers for Disease Control and Prevention* https://www.cdc.gov/nchs/fastats/obesity-overweight.htm (2023).
36. U.S. Census Bureau. QuickFacts: United States. *Census.gov* https://www.census.gov/quickfacts/fact/table/US/LFE046218 (2023).
37. Cheryan, S. & Markus, H. R. Masculine defaults: Identifying and mitigating hidden cultural biases. *Psychol. Rev.* **127**, 1022–1052 (2020).

Author contributions

R.S., J.D.F. and S.P. conceived the project and designed the research. R.S. supervised data collection, analyzed the data, wrote the original draft, and prepared Figs. 1–2 and table 1. All authors reviewed and edited the manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to R.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.