Names as Social Signals: Quantifying Name-Based Appearance and Socioeconomic Biases in Text-to-Image Models

Zifeng Liang[1] **Fangjian Shen**[1] **Yingpeng Sang**[1‡] **Shuangjuan Li**[2]
[1]School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou, Guangdong, China
[2]College of Mathematics and Informatics (College of Software Engineering),
South China Agricultural University, Guangzhou, Guangdong, China
{liangzf5, shenfj3}@mail2.sysu.edu.cn sangyp@mail.sysu.edu.cn lishj2016@scau.edu.cn

Abstract

Human perception often relies on names when forming first impressions of unfamiliar people. As text-to-image (T2I) models are increasingly used in everyday content creation, names may also influence what these models generate under otherwise neutral prompts. We study name-based bias in mainstream T2I models along two dimensions: appearance patterns and socioeconomic status patterns. We curate name sets from China, America, Britain and France, generate over 6,280 images, and use automated VLM-based scoring to assess age appearance and styling, as well as occupation and living-environment attributes in work and home scenes. Across models, changing only the name yields reliable shifts in these attributes. The strength and consistency of the pattern vary by model, with Nano Banana and ChatGPT-5 showing more uniform name-linked defaults. We propose a quantitative framework to identify and compare name-based biases, helping advance fairer AI systems.

Keywords: Name-Based Bias; Text-to-Image Models; Social Stereotypes; Vision–Language Models; Socioeconomic Status

Introduction

Human social perception is fast and associative. A personal name can shape expectations about who someone is, including their age, ethnicity, and socioeconomic status, through patterns of impression formation (Allport, 1954; Devine, 1989). These regularities offer a useful lens on how social categories are organized in human cognition. In parallel, recent advances in text-to-image (T2I) models, such as Stable Diffusion (Rombach et al., 2022), FLUX.1 (Greenberg, 2025), Nano Banana (Raisinghani, 2025), and others, have enabled the generation of high-fidelity and contextually relevant images (Zhang et al., 2023). However, most current research focuses on technical issues or on explicit biases such as gender and race (Bianchi et al., 2023; Cho et al., 2023). The name-based bias dimension remains unexplored. In particular, how a name alone is associated with visual attributes such as apparent age, everyday settings, and material environments has not been examined.

Such patterns tied to names can matter because names are a common way to refer to people in everyday prompts. Under underspecified descriptions, models often infer missing details about the person and setting, such as apparent age, styling, and occupation. If these details shift in a consistent way with the name, the model may repeatedly produce similar

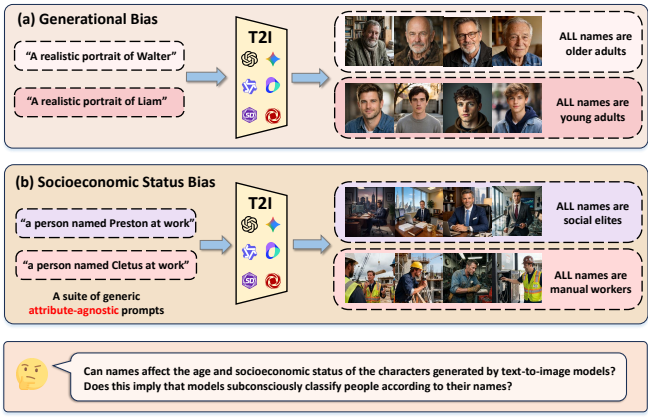

Figure 1: Illustration of name-based bias in T2I models.

portrayals for the same kinds of names. For example, names that are widely perceived as older generation names may more often yield subjects who appear older, along with more old fashioned styling, even without any age description. Likewise, names that prior work has linked to socioeconomic differences (Bertrand & Mullainathan, 2004; Gaddis, 2015) may more often be rendered in higher prestige workplaces and more affluent interiors, whereas names with different socioeconomic associations in those sources may more often appear in service or manual work settings and more modest living spaces (Mandal & Leavy, 2023). As T2I models are used in advertising, illustration, and character design, repeated defaults tied to names may reinforce familiar stereotypes about who is expected to look a certain age or belong in particular kinds of jobs and environments (Qadri et al., 2023; Shumailov et al., 2023; Weidinger et al., 2021). This may not only amplify the spread of bias, but also cause harm to specific groups (Fraser et al., 2023; Friedrich et al., 2023). Studying name-based bias is therefore not only a matter of model assessment, but also a necessary step toward understanding and, ultimately, reducing its broader social impact. To address this gap, we introduce a framework to analyze two recurring forms of name-based bias: Appearance Bias and Socioeconomic Status (SES) Bias. We focus on two research questions:

- **RQ1:** Do personal names systematically influence the apparent age and styling of generated people? If so, to what extent are these effects stable across contexts?

‡Corresponding author.

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

- **RQ2:** Do personal names systematically influence portrayals of socioeconomic status (SES), such as occupation and home settings? If so, to what extent do these patterns differ between workplace contexts and domestic contexts?

Our study uses a large scale design with testing across multiple national contexts for Appearance Bias. We examine whether defaults tied to names for age and styling persist across different naming systems rather than reflecting a property of any single context. For Appearance Bias, we test how name popularity across decades relates to the apparent age and era styling of generated subjects in four contexts: Chinese, American, British, and French. For SES Bias, we conduct a parallel analysis in these same contexts, examining how names commonly used by different socioeconomic groups shift occupational prestige and living environments in two settings (work vs. home). To quantify these patterns, we define indicators that capture how these biases appear in the outputs. Finally, we develop an automated evaluation procedure using a vision-language model (VLM) to annotate images along these indicators, enabling large scale and consistent measurement of name-based bias in T2I systems.

Contribution. The contribution can be summarized as:

- To our knowledge, we are among the first to systematically investigate Name-Based Bias in T2I generative models, dissecting it through two critical dimensions: Appearance and Socioeconomic Status (SES) bias.
- We construct a name dataset spanning multiple countries and social groups, apply it to several mainstream T2I models, and generate over 6,280 images to support analysis.
- We propose a VLM-based metric enabling the automated and quantitative assessment of these biases.

Related Work

Names as Social Signals in Cognition. In human social perception, names can shape first impressions and guide expectations (Allport, 1954; Devine, 1989). A large body of work shows that names influence judgments with real-world consequences; for example, racially distinctive names significantly affect hiring callbacks (Bertrand & Mullainathan, 2004). Beyond race and gender, names carry layered meanings that blend ethnicity with class indicators, suggesting a name's perceived demographic affiliation often implies socioeconomic background (Gaddis, 2015; Garg et al., 2018). Furthermore, names influence age perception, where temporal popularity trends guide perceivers to categorize individuals as elderly or young (Johfre, 2020a). This literature establishes names as cognitively potent signals, providing a foundation for comparing human cognitive patterns with model behaviors.

Biases in Text-to-Image Models. Text-to-image systems frequently reproduce biases from web-scale training data, ranging from cultural stereotypes to brand obsessions (Naik & Nushi, 2023; Shen et al., 2025; Zeng et al., 2025). Beyond documenting representational disparities, recent evaluations show that models often resolve underspecified descriptions by converging on stereotyped prototypes rather than maintaining diversity (Cho et al., 2022). However, most existing work measures bias through *explicit* demographic descriptors or task phrases (Friedrich et al., 2023). Recent studies in broader generative systems indicate that fundamental cultural narratives deeply shape model outputs (Shen et al., 2026), and names themselves serve as powerful implicit signals. In large language models, specific names can shift generated portrayals and modulate bias patterns (Berlincioni et al., 2024; Cheng et al., 2023; Sheng et al., 2019). Given that text encoders underpin image generation architectures, these linguistic associations risk propagating into the visual domain. Although initial audits exist (Fraser et al., 2023), systematic research is lacking on how name-based bias manifests across appearance and socioeconomic dimensions in mainstream T2I models. This study aims to fill this gap across multiple cultural contexts, contributing to fairer generative AI.

Defining Name-Based Bias in Models

This section defines Name-Based Bias in T2I models. Names are widely understood as social signals that help form quick impressions (Allport, 1954). We apply this idea to generative AI to investigate how name-based bias appears in model outputs and how it is measured.

Demographic research on naming trends indicates that names function as historical markers, carrying implicit signals about birth cohorts (Johfre, 2020b). The "Face-Name Matching Effect" further suggests that social expectations linked to names can shape perceived physical appearance (Zwebner et al., 2017), while classic psychological studies demonstrate that names alone can evoke distinct impressions of success and morality (Mehrabian, 2001). Sociological studies on stratification demonstrate that names act as indicators of social class, influencing judgments about employability and status (Bertrand & Mullainathan, 2004), while theories of distinction argue that social class is visibly manifested through material environments and tastes (Bourdieu, 1984). Furthermore, the Stereotype Content Model suggests that status indicators drive quick social categorization (Fiske et al., 2002), and research in the psychology of social class confirms that observers accurately infer status from subtle environmental remnants and artifacts (Kraus et al., 2017). Building on these foundations, we analyze this bias through two specific lenses: **Appearance Bias** and **Socioeconomic-Status (SES) Bias**.

Appearance Bias

Appearance Bias occurs when a name serves as a predictive signal for the biological age and historical era of the generated subject. As illustrated in Figure 2, models tend to generate elderly individuals for names common in older generations, and young adults for names popular among younger groups.

Age Bias. Models frequently map names to specific life stages. When prompted with names historically popular in earlier decades (e.g., "Gertrude" or "Walter"), models tend to generate elderly subjects with visible signs of aging. Con-

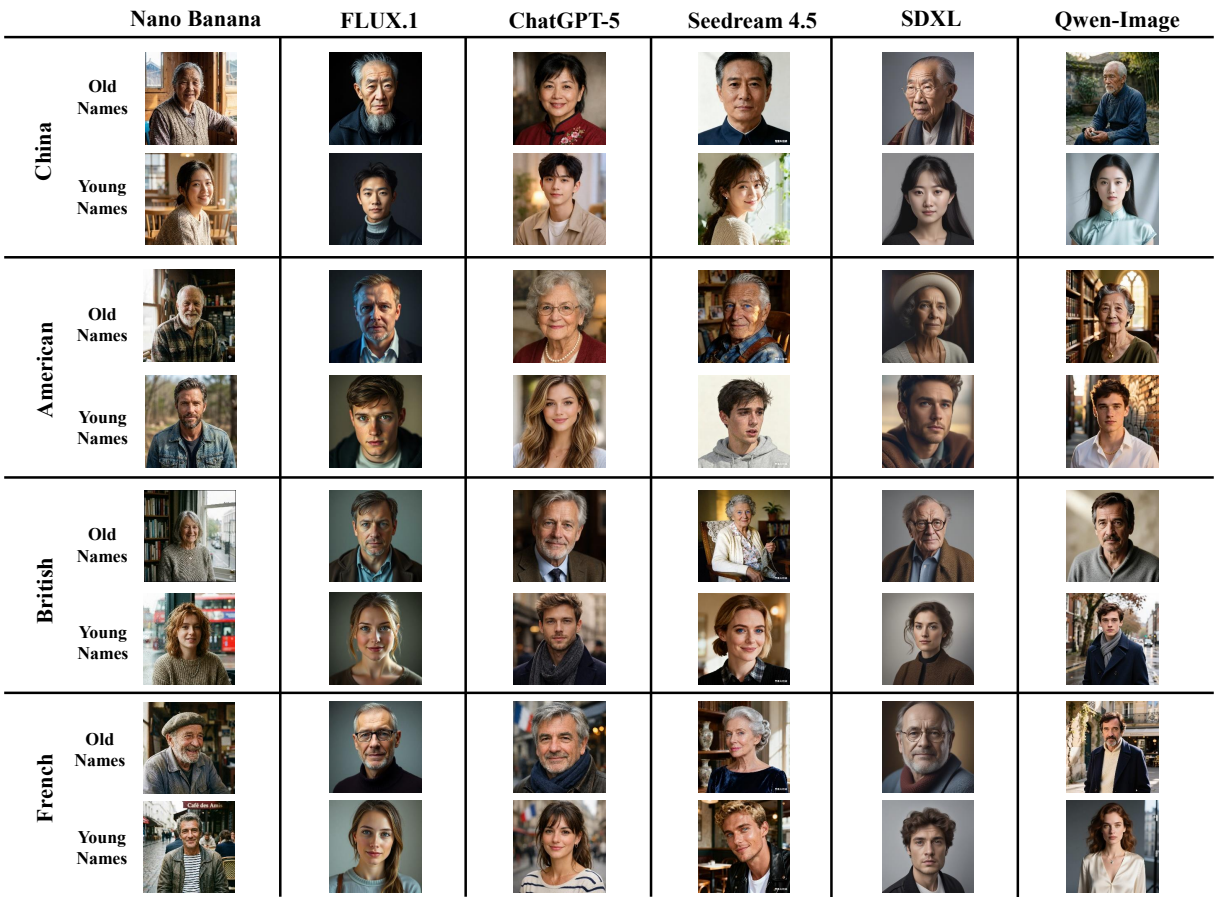

Figure 2: Visualizing Appearance Bias. Even without age prompts, names associated with different eras (e.g., 1950s vs. 2000s) frequently trigger distinct age and styling patterns.

versely, names popular in the last ten years (e.g., "Liam" or "Olivia") predominantly yield young adults or teenagers. Crucially, this often occurs even when the prompt does not specify age, suggesting the name alone may override the neutrality of the prompt and guide the subject into a specific age bracket.

Styling Bias. Names tend to influence the aesthetic style and era of styling. Beyond physiological age, names often trigger distinct sartorial norms. Subjects generated with names associated with older cohorts are frequently styled in conservative or dated attire, such as cardigans, pearl necklaces, or mid-20th-century formal suits. In contrast, subjects with names associated with younger cohorts are more commonly rendered in modern, casual streetwear, such as hoodies or graphic t-shirts. This link suggests that the model often treats the name as a style reference, aligning the subject's overall presentation with the peak popularity era of their name.

Socioeconomic-Status (SES) Bias

SES bias refers to the systematic skewing of a subject's perceived social class manifested through **workplace settings, living environment, and material possessions** based on the name provided. We emphasize that our categorization draws on statistical associations identified in sociological research (Clark, 2014; Coulmont, 2011; Gaddis, 2015), which link specific names to demographic data such as parental education or historical persistence of status. Consequently, the descriptors 'names associated with higher (or lower) socioeconomic status' used in our analysis are adopted directly from these external empirical findings. While the content may reflect certain stereotypes or biases, we emphasize that these do not represent our views. As shown in Figure 3, changing only the name in a prompt can shift the visual narrative from one of affluence to one of modesty or manual labor.

Occupation Bias. Names often bias the selection of occupational roles. Under neutral professional prompts, names historically linked to higher socioeconomic status tend to result in images depicting white-collar or knowledge-work roles, such as executives or doctors. Conversely, names associated with lower socioeconomic status are more frequently associated with blue-collar or service-oriented roles, often depicted in uniforms or engaging in manual tasks. The model effectively sorts names into implicit labour categories, narrowing the range of probable careers for specific demographic groups.

Lifestyle Bias. Names often influence the depiction of daily habits and leisure activities within the home. For names associated with higher socioeconomic status, models tend to

Figure 3: Visualizing SES Bias. Names associated with different socioeconomic statuses systematically skew the depicted occupation, environment, and material objects.

generate subjects engaged in active or cultural leisure, such as reading, dining, or socializing. Conversely, names associated with lower socioeconomic status are more frequently depicted performing domestic chores or in states of exhaustion and passive rest. This pattern links specific names to fixed behavioral routines, reinforcing stereotypes about who can enjoy leisure and who needs to work.

Environment Bias. Names often shape the inferred quality and atmosphere of the background setting. Whether in a home or office context, names associated with higher socioeconomic status are likely to appear in more modern, spacious, and bright environments characterized by contemporary architecture and orderly layouts. In contrast, names associated with lower socioeconomic status may more often trigger cramped, dimmer, or more rustic settings, such as crowded offices or homes with dated interior design. The name thus sets the background context, implicitly signaling the likely economic tier of the space the character inhabits.

Entity Bias. Names may influence the material value of surrounding objects. This bias manifests in the specific artifacts present in the scene. For names associated with higher socioeconomic status, office equipment implies sleek, high-tech devices (e.g., multiple monitors, MacBooks), while furniture often implies luxury materials like mahogany or leather. For names associated with lower socioeconomic status, the same prompts tend to yield older, bulkier technology or utilitarian, worn furniture. These material tendencies indicate that specific names influence the economic value of objects appearing around the character.

Experimental Setup

Model Choices. We survey six representative T2I models: Nano Banana (Raisinghani, 2025), ChatGPT-5 (OpenAI, 2025), SDXL (Podell et al., 2023), FLUX.1 (Greenberg, 2025), Seedream 4.5 (Seedream et al., 2025), and Qwen3-VL-235B-A22B (Wu et al., 2025). We aimed to include widely deployed and recent state-of-the-art models to maximize coverage. To strengthen external validity across regions, the set spans both Western and Chinese model families.

Dataset Design. We construct a controlled dataset spanning four cultural contexts: China, America, Britain, and France. This selection is strategic: it includes representative countries from both Eastern and Western cultural spheres. This diversity allows us to test whether name-based bias is a universal model artifact or a culture-specific phenomenon.

For **Appearance Bias (RQ1)**, we partition names into three groups: *Young*, *Middle-aged*, and *Old*. This design allows us to detect age-related differences and to examine how models map name cues to perceived age variation, i.e., whether the outputs systematically shift toward appearing younger or older appearances. We select names using historical popularity statistics from national statistical bureaus (e.g., the SSA for America and INSEE for France). The intuition is that a name peaking in earlier decades is more likely to be interpreted as

Table 1: Overview of name datasets and prompting strategies. Representative name cohorts are categorized by generation and socioeconomic status (SES) across four cultural contexts. Names listed are examples from the full dataset (Given Names only).

RQ	Group	China	America	Britain	France
RQ1	Old Names	Jianguo, Guiying	Walter, Linda	Colin, Susan	Philippe, Martine
	Middle-aged Names	Junjie, Youqing	Michael, Jennifer	Paul, Sarah	Laurent, Nathalie
	Young Names	Zixuan, Yutong	Liam, Olivia	Archie, Isla	Lucas, Léa
RQ2	High SES Names	Zhuoran, Yuwei	Maxwell, Victoria	Rupert, Philippa	Charles, Appoline
	Neutral SES Names	Weiming, Minjie	David, Jennifer	Paul, Claire	Thomas, Marie
	Low SES Names	Fuhai, Guizhen	Waylon, Misty	Gareth, Sharon	Kevin, Cindy

belonging to an older cohort, whereas a name with more recent popularity is more likely to signal a younger cohort. This selection strategy follows prior demographic audits (Garg et al., 2018; Johfre, 2020b) and ensures that each chosen name is strongly associated with a specific birth decade.

For **SES Bias (RQ2)**, we employ a tiered classification: *High SES*, *Neutral*, and *Low SES*. This stratification draws on established sociological research regarding name signaling and social mobility, which identifies names historically associated with specific socioeconomic backgrounds (Bertrand & Mullainathan, 2004; Clark, 2014; Coulmont, 2011; Gaddis, 2015). **We emphasize that we adopt these categories strictly to audit model behavior. The generated images are produced solely for analysis and do not represent our views.** Our goal is to explore robust, widely-shared name signals rather than to exhaustively map every possible name. Consequently, we select representative names for each group, as detailed in Table 1. We probe SES bias across two distinct settings, *Work* and *Home*, to assess how these impressions manifest differently in professional versus domestic contexts.

Configurations. All experiments were conducted with standardised configurations to ensure consistency and reproducibility. For each prompt, we generated $N = 10$ images and repeated this for both the appearance and SES settings. This repeated sampling reduces sensitivity to stochastic variation in the generation process and allows us to estimate the stability of each model's outputs beyond single-run idiosyncrasies. In total, this procedure yielded over 6,280 images across all models and across all names in the full dataset.

Evaluation. To examine Appearance Bias (RQ1), we utilized the neutral template "A realistic portrait of [NAME]". To examine SES Bias (RQ2), we employed context-specific templates: "A [Nationality] person named [NAME] at [work / home]". We quantify name-based bias by evaluating the extent of visual uniformity and stereotype consistency in the generated outputs. Using a fixed scoring rubric, we assign a score ranging from 0 to 10 for each bias dimension, where higher scores indicate a stronger fixation on stereotyped visual traits. We implement this rubric via a VLM-based evaluation procedure (Gemini 3.0 Pro), which generates structured annotations and computes quantitative scores. Crucially, each bias dimension is assessed through multiple sub-factors. The final aggregate score is computed as the mean of the relevant dimension scores, excluding any non-applicable categories.

Experimental Results

Main Results

Table 2 presents the quantitative results across six models. Nano Banana demonstrates the strongest bias, recording the highest overall scores for Appearance (8.12) and a strong Socioeconomic (8.01) dimension, driven by Environment (8.53) and Occupation (8.45). ChatGPT-5 records the highest overall SES score (8.09) with low variance (Occup. $\sigma = 0.88$), indicating it produces consistent stereotyped outputs regardless of the specific prompt category. Conversely, Qwen3-VL-235B yields the lowest overall scores (7.02 and 6.89) but exhibits high variance in the SES category ($\sigma = 1.37$), suggesting its outputs vary significantly across different cultural contexts. Notably, FLUX.1-schnell also shows strong bias in specific areas, such as Age (8.21), despite having a lower overall score than Nano Banana.

RQ1: Appearance Bias

Building on the overall trends, we first examine the appearance dimension in detail. Figure 4 breaks down the Appearance Bias scores by age cohort and cultural context. The data indicates that names associated with the Older adults cohort trigger the strongest visual stereotypes across all models, followed by Young adults, while the Middle-aged group shows the most variation. In the Older category, Nano Banana demonstrates high consistency, peaking at 9.28 in the Chinese context and 9.19 in the American context, suggesting a strong association between these names and elderly features. Regional differences are notable; for instance, Qwen3-VL scores 8.12 in the American context but drops to 5.82 in the French context, indicating lower sensitivity to French naming conventions for the elderly. For the Middle-aged group, ChatGPT-5 shows a significant divergence: it scores 7.13 in the Chinese context but drops to 3.55 in the American context, implying it lacks a consistent visual prototype for middle-aged Western names.

RQ2: Socioeconomic Status (SES) Bias

Next, we shift our focus to the socioeconomic dimension. Figure 4 reveals that names associated with High or Low SES

Table 2: Quantitative Comparison of Name-Based Bias across six T2I models. We report **Mean Score** and **Standard Deviation** (σ). Scores are aggregated across all tested cultural contexts and social groups. Higher means indicate stronger stereotype adherence. Within each column, the highest value is marked in **bold**, and the second highest is underlined.

Model	Appearance Bias						Socioeconomic Status (SES) Bias									
	Overall		Age		Styling		Overall		Occup.		Life.		Env.		Entity	
	Score	σ	Score	σ	Score	σ	Score	σ	Score	σ	Score	σ	Score	σ	Score	σ
Nano Banana	**8.12**	1.14	**8.28**	1.09	7.96	1.18	8.01	0.83	**8.45**	0.76	7.16	0.92	**8.53**	0.81	7.91	0.84
Qwen3-VL-235B	7.02	1.02	7.17	1.08	6.87	0.97	6.89	**1.37**	7.07	**1.39**	6.70	1.21	7.01	**1.42**	6.78	**1.45**
SDXL	7.34	1.29	7.75	1.23	6.93	1.34	7.10	1.35	7.44	1.24	6.58	**1.56**	7.28	1.33	7.10	1.29
Seedream 4.5	7.19	1.35	7.27	**1.41**	7.11	1.29	7.26	1.30	7.51	1.28	7.02	1.43	7.23	1.16	7.28	1.34
FLUX.1-schnell	7.66	1.21	8.21	1.16	7.13	1.25	7.80	1.16	7.78	1.07	7.80	1.14	7.48	1.37	8.14	1.06
ChatGPT-5	8.03	**1.43**	8.00	1.38	**8.06**	**1.47**	**8.09**	0.93	8.12	0.88	**7.83**	1.03	8.26	0.91	**8.15**	0.89

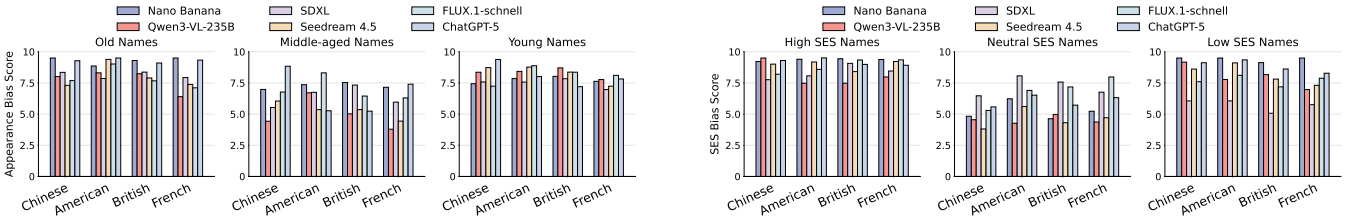

Figure 4: **Experimental results across countries and demographic groups.** (left) Mean overall Appearance Bias scores for each T2I model, broken down by country and age group. (right) Mean SES bias scores for each text-to-image model across four countries, broken down by name SES group.

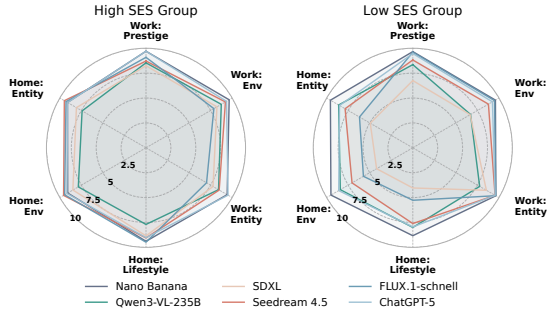

Figure 5: Dimension-specific SES bias scores across six models for High-SES and Lower-SES name groups. Higher scores indicate stronger bias.

groups result in consistently higher bias scores compared to Neutral names. In the High SES category, the American context drives the widest model divergence, where Nano Banana and ChatGPT-5 reach peak scores of 9.21 and 9.08 respectively, strongly linking these names to elite workplace settings. Conversely, for Neutral SES names, while most models drop to scores between 3.0 and 4.0, SDXL and FLUX.1 remain outliers with scores of 7.82 and 6.57 in the American setting, defaulting to stylized imagery even without explicit cues.

To understand the specific visual components driving these scores, Figure 5 clarifies the structure of this bias. For the Low SES group, we observe a distinct split in model behavior regarding domestic life. While Nano Banana maintains high bias across all dimensions (e.g., 8.41 on Home Environment), SDXL shows a notable reduction in bias for Home Lifestyle (score 3.88) and Home Environment (score 4.06). This suggests that while SDXL stereotypes occupational roles, it allows for more visual diversity when depicting the domestic settings of individuals with lower socioeconomic status, whereas models like Nano Banana and ChatGPT-5 apply stereotypes uniformly across both public and private contexts.

Conclusion

This study applies social perception research to T2I models and proposes a new framework for evaluating name-based bias. Our findings show that personal names consistently shape how mainstream models depict appearance attributes and influence scene portrayals related to socioeconomic status. We further quantify these models' bias in producing stereotyped impressions tied to names. By introducing a large-scale experimental setup and integrating VLM-based evaluation, we provide an effective method for identifying and comparing name-based bias in generative image models. This work strengthens the understanding of name-based bias in generative vision models and supports the development of fairer and more robust generative AI.

Acknowledgments

This work was supported in part by Guangdong S&T Programme (Grant No. 2024B0101040007).

References

Allport, G. W. (1954). *The nature of prejudice*. Addison-Wesley.

Berlincioni, L., Cultrera, L., Becattini, F., Bertini, M., & Del Bimbo, A. (2024). Prompt and prejudice. *arXiv preprint arXiv:2408.04671*.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, *94*(4), 991–1013.

Bianchi, F., Kalluri, P., Durmus, E., Ladhak, F., Cheng, M., Nozza, D., & Jurafsky, D. (2023). Easily accessible text-to-image generation amplifies demographic stereotypes at large scale. *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 1493–1504.

Bourdieu, P. (1984). *Distinction: A social critique of the judgement of taste*. Harvard University Press.

Cheng, M., Durmus, E., & Jurafsky, D. (2023). Marked personas: Using natural language prompts to measure stereotypes in language models. *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1504–1532.

Cho, J., Zala, A., & Bansal, M. (2022). DALL-Eval: Probing the reasoning skills and social biases of text-to-image generative models. *arXiv preprint arXiv:2202.04053*.

Cho, J., Zala, A., & Bansal, M. (2023). Dall-eval: Probing the reasoning skills and social biases of text-to-image generation models. *arXiv preprint arXiv:2302.06544*.

Clark, G. (2014). *The son also rises: Surnames and the history of social mobility*. Princeton University Press.

Coulmont, B. (2011). *Sociologie des prénoms*. La Découverte.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5–18.

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878.

Fraser, K. C., Nejadgholi, I., & Kiritchenko, S. (2023). A fairer face: Auditing and mitigating name-based biases in text-to-image generative models. *Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 10839–10856.

Friedrich, F., Schramowski, P., Brack, M., Struppek, L., Hintersdorf, D., Luccioni, S., & Kersting, K. (2023). Fair diffusion: Instructing text-to-image generation models on fairness. *arXiv preprint arXiv:2302.10893*.

Gaddis, S. M. (2015). Discrimination in the credential society: An audit study of race and college selectivity in the labor market. *Social Forces*, *93*(4), 1451–1479.

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, *115*(16), E3635–E3644.

Greenberg, O. (2025). Demystifying flux architecture. *arXiv preprint arXiv:2507.09595*.

Johfre, S. S. (2020a). What age is in a name? A test of names as signals of age. *Sociological Science*, *7*, 367–390.

Johfre, S. S. (2020b). What age is in a name? cohort-based name trends and the social perception of age. *Sociological Science*, *7*, 348–373.

Kraus, M. W., Park, J. W., & Tan, J. J. (2017). Signs of social class: The experience of economic inequality in everyday life. *Perspectives on Psychological Science*, *12*(3), 422–435.

Mandal, S., & Leavy, S. (2023). Vibecheck: Contextual bias in text-to-image generation. *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, 432–442.

Mehrabian, A. (2001). Characteristics attributed to individuals on the basis of their first names. *Genetic, Social, and General Psychology Monographs*, *127*(1), 59–88.

Naik, R., & Nushi, B. (2023). Social biases through the text-to-image generation lens. In F. Rossi, S. Das, J. Davis, K. Firth-Butterfield, & A. John (Eds.), *Proceedings of the 2023 AAAI/ACM conference on AI, ethics, and society, AIES 2023, montréal, qc, canada, august 8-10, 2023* (pp. 786–808). ACM.

OpenAI. (2025, August). GPT-5 System Card [Accessed: 2025-12-12].

Podell, D., English, Z., Lacey, K., Blattmann, A., Dockhorn, T., Müller, J., Penna, J., & Rombach, R. (2023). Sdxl: Improving latent diffusion models for high-resolution image synthesis. *arXiv preprint arXiv:2307.01952*.

Qadri, R., Gokul, M., Raz, R., & Srikumar, V. (2023). Ai's regimes of representation: A community-centered study of text-to-image models in south asia. *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency (FAccT)*, 506–517.

Raisinghani, N. (2025, November). Introducing nano banana pro [Google Blog (The Keyword). Accessed: 2025-12-12].

Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-resolution image synthesis with latent diffusion models. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 10684–10695.

Seedream, T., Chen, Y., Gao, Y., Gong, L., Guo, M., Guo, Q., Guo, Z., Hou, X., Huang, W., Huang, Y., et al. (2025). Seedream 4.0: Toward next-generation multimodal image generation. *arXiv preprint arXiv:2509.20427*.

Shen, F., Liang, Z., Wang, C., & Wen, W. (2025). Cider: A causal cure for brand-obsessed text-to-image models. *arXiv preprint arXiv:2509.15803*.

Shen, F., Zeng, Y., Lyu, D., & Wen, W. (2026). From words to worlds: Measuring cultural narrative bias in llms via a structural–value pipeline. *Proceedings of the ACM Web Conference 2026*, 8961–8972.

Sheng, E., Chang, K.-W., Natarajan, P., & Peng, N. (2019). The woman worked as a babysitter: On biases in language generation. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3405–3410.

Shumailov, I., Shumaylov, Z., Kazhita, Y., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2023). The curse of recursion: Training on generated data makes models forget. *arXiv preprint arXiv:2305.17493*.

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P.-S., Cheng, M., Glaese, M., Balle, B., Kasirzadeh, A., et al. (2021). Ethical and social risks of harm from language models. *arXiv preprint arXiv:2112.04359*.

Wu, C., Li, J., Zhou, J., Lin, J., Gao, K., Yan, K., Yin, S.-m., Bai, S., Xu, X., Chen, Y., et al. (2025). Qwen-image technical report. *arXiv preprint arXiv:2508.02324*.

Zeng, Y., Dong, F., Zhao, J., Zheng, P., Li, J., & Zhou, H. (2025). Towards culturally fair multimodal generation: Quantifying and mitigating orientalist biases in text-to-visual models. *Proceedings of the 33rd ACM International Conference on Multimedia*.

Zhang, C., Zhang, C., Zhang, M., & Kang, I. S. (2023). A survey on text-to-image generation. *arXiv preprint arXiv:2303.07909*.

Zwebner, Y., Sellier, A.-L., Rosenfeld, N., Goldenberg, J., & Mayo, R. (2017). We look like our names: The manifestation of name stereotypes in facial appearance. *Journal of Personality and Social Psychology*, *112*(4), 527–554.