8-2024

Predicting personality or prejudice? Facial inference in the age of artificial intelligence

Shilpa MADAN

Gayoung PARK

Predicting Personality or Prejudice? Facial Inference in the Age of Artificial Intelligence

Shilpa Madan[a]
Singapore Management University
50 Stamford Road
Singapore, 178889
shilpamadan@smu.edu.sg

Gayoung Park
Virginia Tech
880 West Campus Drive, Blacksburg, Virginia
USA 24061
gayoung@vt.edu

[a] Corresponding author

Predicting Personality or Prejudice? Facial Inference in the Age of Artificial Intelligence

Abstract

Facial inference, a cornerstone of person perception, has traditionally been studied through human judgments about personality traits and abilities based on people's faces. Recent advances in artificial intelligence (AI) have introduced new dimensions to this field, employing machine learning algorithms to reveal people's character, capabilities, and social outcomes based just on their faces. This review examines recent research on human and AI-based facial inference across psychology, business, computer science, legal, and policy studies to highlight the need for scientific consensus on whether or not people's faces can reveal their inner traits, and urges researchers to address the critical concerns around epistemic validity, practical relevance, and societal welfare before recommending AI-based facial inference for consequential uses.

Keywords: facial inference, facial profiling, artificial intelligence, machine learning, bias, privacy

Predicting Personality or Prejudice? Facial Inference in the Age of Artificial Intelligence

"Human face is a slate where every line has a meaning;
a slight cramp can look like mockery and a scratch like deceitfulness."

Lichtenberg (1778)

1 Introduction

People's faces play a key role in how others perceive and respond to them—affecting everything from snap judgments to behavioral outcomes [1, 2]. Given the proximal and evident nature of facial stimuli, it is not surprising that people use others' faces to infer their personalities and character traits that are not so apparent [3-5]. The idea that people's outer appearance can reveal something deeper about their character is intuitively appealing. Since the time of Aristotle, humans have believed in and practiced physiognomy—seeking to infer others' character based on their appearance [6]. Johann Caspar Lavater (1878), a Swiss theologian, published a series of popular articles in which he argued that imperfections on the face were a mark of personal imperfections. Even though folk wisdom and popular culture are replete with reminders that others can be "wolves in sheep's clothing," and "don't judge a book by its cover," people spontaneously infer others' traits such as sociability [7], competence [8, 9], trustworthiness [10, 11], and even sexual unfaithfulness [12] from their physical appearance, often in split seconds. Advances in machine learning have allowed researchers to model these facial inferences and generate facial stimuli that are perceived to vary across a wide variety of character traits, including trustworthiness and dominance [13].

1.1 Consequences of Facial Inferences

Extensive evidence shows that such inferences can profoundly affect critical outcomes in legal, financial, interpersonal, and political domains. For instance, partners perceived to be more trustworthy based on their faces received greater rewards from children in an economic trust

game [14]. Beyond children's games, adults are as likely to rely on facial inferences for consequential decisions—defendants perceived to be less agreeable based on their faces were more likely to receive a guilty verdict [15], and those perceived to be less trustworthy were more likely to receive the death penalty [16]. Entrepreneurs who were perceived to be more competent, confident, trustworthy, and resilient based on their faces were more likely to receive investor funding [17]. Business leaders judged to be more powerful based on their faces enjoyed better popularity-based outcomes (e.g., vote share) and performance-based outcomes (e.g., company profits) [18]. Similarly, country leaders judged to be more competent based on their faces received more foreign direct investment [19].

1.2 Accuracy of facial inferences

Such judgments may be common, spontaneous, and consequential, but they are not very accurate [20]. Specifically, most academic research has focused on whether people can predict others' personalities (e.g., extraversion, openness, agreeableness, emotional stability, conscientiousness, trustworthiness, dominance, etc.) at *above-chance* levels (i.e., more than 50%). Studies show that people can infer others' level of openness, extraversion, neuroticism, agreeableness, self-esteem, religiosity, competence, and sexual orientation based on their appearance at above-chance levels [21-24]. Even though significant, the correlations between others' ratings and self-ratings on these personality traits are small to moderate (~ .14 to .27), limiting the practical usefulness of these ratings [23, 25]. Therefore, statements about being able to accurately predict others' character from their appearance necessitate distinguishing between significance testing and the magnitude of the effect [22].

In addition to this modest accuracy for some traits, the evidence is equivocal for other traits, such as political orientation [26, 27] and intelligence [28]. Finally, these inferences are

driven by both perceiver and target characteristics. Importantly, in a study investigating 700,000 ratings of faces, trait inferences (e.g., creative, intelligent, competent, etc.) were affected more by perceivers' characteristics than the target stimuli [29, 30]. Facial inferences have also been shown to be influenced by culture [31], learned stereotypes about social groups [32], target gender [9], and individual differences in social learning [11].

1.3 (Over) Confidence in Facial Inferences

Interestingly, people not only infer others' personalities from their faces even when more diagnostic information is available, but they are also exceedingly confident in their judgments [5, 33, 34], indicating that people lack metacognition about their facial inferences. Importantly, the reliance on face-based judgments persists even when people are made aware of the inaccuracy and far-reaching implications of such judgments [35], unless they are confronted with negative information about the target that is both extreme and reliable [10]. Research shows that people's belief in physiognomy [36, 37], or the idea that an individual's appearance reveals their character [34], predicts their confidence in their face-based trait judgments. Importantly, the more people believe that individuals' appearance reveals their character, the more they support AI-based facial inference (facial profiling) for consequential uses such as law enforcement, employment, education, and business [34].

2 AI-based facial inference

With the advent of machine learning, facial inference has evolved beyond the capabilities of human observers. Computer algorithms have been developed to assess facial features and predict or infer people's traits and tendencies—also known as facial profiling. Unlike facial recognition, which simply matches a face to an identity in a database, facial profiling seeks to draw conclusions about a person's character traits, or even predict future actions based on

physical appearance. Facial profiling has gained traction in popular culture as well as in the business and governance sectors due to its purported ability to revolutionize decision-making by enhancing efficiency, convenience, and customization, while reducing costs. For instance, facial profiling algorithms that can infer criminal intent could be invaluable for public safety agencies by helping them arrest potential criminals before they act. Similarly, facial profiling algorithms that can infer trustworthiness could help banks and insurance firms reduce the incidence of delinquent loans. Facial profiling tools that can evaluate competence or other traits could aid employers in hiring more suitable candidates, and those that assess consumers' preferences could enable retail stores and advertisers to offer customized recommendations. Indeed, there are facial profiling apps that promise to provide comprehensive personality and compatibility assessments based solely on a selfie (e.g., Facemetrics). For instance, an Israeli startup, Faception, claims that its facial profiling algorithm based on deep learning methods can predict whether someone is a terrorist, or a pedophile, among other things, simply by analyzing their face. Facial profiling technology also promises to streamline the insurance claims process by predicting people's likelihood of submitting fraudulent insurance claims through a mere selfie [38].

2.1 Rationale behind AI-based facial inference

In addition to these commercial applications, extensive academic research has developed algorithms that claim to reliably infer underlying traits including, but not limited to, extraversion, openness, conscientiousness, neuroticism, and agreeableness [39-43], political orientation [44-46], criminality [47], and sexual orientation [20] from people's pictures or short video clips. The correlations between actual and inferred attributes range from modest to high (r ~ .13 - .80) [20, 40, 45, 47]. This work is based on the argument that the low accuracy of humans when judging character traits from others' faces does not necessarily mean that relevant cues are not

prominently displayed. Instead, people may simply be unable to detect or interpret them [20].

For example, researchers have used a machine learning algorithm to determine the

trustworthiness of CFOs through facial features such as the angle of the inner eyebrow ridge, the

roundness of the face, and the angle of the chin [48], and find that auditors charge a lower fee to

firms whose CFOs were identified as more trustworthy-looking by the machine learning

algorithm. Similarly, financial analysts perceived to be more trustworthy based on their faces by

a machine learning algorithm had better forecast accuracy, presumably because firm managers

were more inclined to share relevant information with analysts who appeared more trustworthy

[49].

2.2 Concerns related to AI-based facial inference

Despite its growing popularity in both academic research and commercial applications,

scholars across multiple disciplines have criticized AI-based facial inference as resurrecting the

pseudoscience of physiognomy by prioritizing predictive power at the data level without

questioning the epistemic foundations [50, 51]. AI-based facial inference is fundamentally built

on the assumption that specific facial attributes can reliably predict personal traits— echoing

outdated beliefs that lack scientific consensus [52]. Given the consequential use cases of AI-

based facial inference systems highlighted above, the dangers of such assumptions are not

merely theoretical. While human social psychology research has focused on the ability to predict

character traits based on facial stimuli with above-chance accuracy, recent academic research and

commercial AI-based facial inference systems promise near-perfect accuracy. Casting doubts on

these claims, commercial gender detection tools [53] and job candidate selection software [54,

55] have been shown to make erroneous gender and race classifications.

Furthermore, these AI-based systems inherit the biases of the data they are trained on, casting doubts on the claims to enhance fairness. For example, driven by their over-representation in criminal databases, African American individuals were more likely to be wrongly classified as suspects [56], underscoring that AI-based facial inference systems risk perpetuating further discrimination against marginalized communities [57]. These concerns are exacerbated by AI-based facial inference systems often operating as 'black boxes,' where the algorithms' internal workings and decision-making processes remain opaque. This lack of transparency may mean that the inferences are simply based on irrelevant or contextual data, or worse, systematic human biases and stereotypes reflected in the data on which it was trained [50, 58]. Apart from these issues that are specific to facial inference, the use of these systems without clear consent, especially in surveillance and law enforcement contexts, raises significant concerns about privacy, mass surveillance, and the erosion of civil liberties, leading law researchers to urge policymakers to stall the proliferation of AI-based facial inference systems and declare them deceptive, unfair, unjust, and oppressive [55].

Importantly, these myriad concerns about AI-based facial inferences are shared by laypeople. Recent research shows a remarkable correlation between the judgments of AI-based facial inferences by AI-competent experts and laypeople [51], with both AI experts and laypeople unwilling to trust character trait inferences made by AI [51, 59]. These judgments were driven by people's beliefs about the irrelevance of facial images for inferring personality traits, potential over-reliance on social stereotypes to form these judgments, lack of evidence that AI can reliably infer character traits from pictures, and ethical and discriminatory concerns.

2.3 Accurate, Fair, and Transparent AI-based Facial Inference

As such, these AI-based facial inference systems lack epistemic validity, and there is no scientific consensus regarding their claims of predictive accuracy [50-52]. As the first step, the field needs to form a scientific consensus about whether or not the face reveals specific traits with sufficiently high accuracy that it can be used to make consequential decisions, and communicate this consensus to the public and to policy experts [34]. Next, rigorous ethical frameworks are needed to ascertain the societal value of AI-based facial inference systems. In conjunction, the research community needs to make efforts to diversify and make more representative the datasets that these systems are trained on to reduce biases. The algorithms and code underlying AI-based facial inference systems must be openly available so that experts assess its validity. The development of such systems must be held to the highest empirical and disclosure standards about their limitations, including classification or prediction accuracy, area under the curve (AUC), sensitivity, specificity, precision, balanced accuracy, and the no information rate [60, Table 1]; that way, users and stakeholders have a full picture of all relevant accuracy metrics, not just hand-picked metrics that make the developers look attractive. Given the potentially severe consequences of mistaken AI-based facial inference systems, this work must be based on interdisciplinary scholarship including, but not limited to, social psychologists, computer scientists, ethicists, and legal and policy experts to ensure that any progress is scientifically valid, practically useful, and societally beneficial.

3 Conclusion

In summary, people tend to make quick judgments about others' character [3], capabilities [24], preferences [61], social outcomes [47], and even humanity [62] from others' faces. Such face-based inferences are hard to update [10, 35] and have the potential to profoundly shape crucial outcomes in life [15-18]. However, these inferences are not very

accurate [5, 22-23]. Despite the limited accuracy, people are over-confident in their ability to do so [33] and tend to generalize this ability to AI-based facial inference systems [34], which, in turn, increases their support for the use of AI-based facial inference for important decisions in diverse domains including business, government, and education. While intuitively appealing, scholars have criticized AI-based facial inference as propagating the pseudoscience of physiognomy and harming marginalized communities [50, 52, 55]. Without the necessary interventions outlined above, AI-based facial inference, as it currently stands, not only holds little value but poses huge risks to individuals and the society.

Conflicts of Interest

Acknowledgments

References

References of particular interest have been highlighted as:
* of special interest
* * of outstanding interest

1. Todorov A: *Face value: the irresistible influence of first impressions*. Princeton University Press; 2017, https://doi.org/10.1515/9781400885725.

2. Zebrowitz LA: **First impressions from faces**. *Curr Dir Psychol Sci* 2017, **26:**237–242, https://doi.org/10.1177/0963721416683996.

3. Naumann LP, Vazire S, Rentfrow PJ, Gosling SD: **Personality judgments based on physical appearance**. *Personal Soc Psychol Bull* 2009, **35**:1661-1671, https://doi.org/10.1177/0146167209346309.

4. Giacomin M, Mulligan A, Rule NO: **Dictators differ from democratically elected leaders in facial warmth**. *Soc Psychol Personal Sci* 2021, **12**:1216-1224, https://doi.org/10.1177/1948550621991368.

5. ** Hong S, Suk HW, Choi Y, Na J: **Face-based judgments: accuracy, validity, and a potential underlying mechanism**. *Psychol Sci* 2021, **32**:1452–1462, https://doi.org/10.1177/09567976211000308.

6. Re DE, Rule NO: **Appearance and physiognomy**. In the *APA handbook of nonverbal communication*. Edited by Matsumoto DE, Hwang HC, Frank MG, American Psychological Association; 2016:221–256, https://doi.org/10.1037/14669-009.

7. Jaeger B, Evans AM, Stel M, van Beest I: **Understanding the role of faces in person perception: Increased reliance on facial appearance when judging sociability**. *J Exp Soc Psychol* 2022, **100**:104288, https://doi.org/10.1016/j.jesp.2022.104288.

8. Lin C, Keles U, Adolphs R: **Four dimensions characterize attributions from faces using a representative set of English trait words**. *Nat Commun* 2021, **12**:5168, https://doi.org/10.1038/s41467-021-25500-y.

9. Oh D, Buck EA, Todorov A: **Revealing hidden gender biases in competence impressions of faces**. *Psychol Sci* 2019, **30**:65–79, https://doi.org/10.1177/0956797618813092.

10. * Shen X, Ferguson MJ: **How resistant are implicit impressions of facial trustworthiness? When new evidence leads to durable updating**. *J Exp Soc Psychol* 2021, **97**:104219, https://doi.org/10.1016/j.jesp.2021.104219.

11. Sutherland CA, Burton NS, Wilmer JB, Blokland GA, Germine L, Palermo R, Collova JR, Rhodes G: **Individual differences in trust evaluations are shaped mostly by environments, not genes**. *Proc Natl Acad Sci USA 2020,* **117**:10218–10224, https://doi.org/10.1073/pnas.1920131117.

12. Foo YZ, Loncarevic A, Simmons LW, Sutherland CA, Rhodes G: **Sexual unfaithfulness can be judged with some accuracy from men's but not women's faces**. *R Soc Open Sci* 2019, **6**:181552, http://dx.doi.org/10.1098/rsos.181552.

13. **Peterson JC, Uddenberg S, Griffiths TL, Todorov A, Suchow JW: **Deep models of superficial face judgments**. *Proc Natl Acad Sci USA* 2022, **119**: e2115228119, https://doi.org/10.1073/pnas.2115228119.

14. Siddique S, Jeffery L, Palermo R, Collova JR, Sutherland CA: **Children's dynamic use of face-and behavior-based cues in an economic trust game**. *Dev Psychol* 2022, **58**: 2275–2286, https://doi.org/10.1037/dev0001438.

15. Ford KD, Penton-Voak IS, Pound N: **Justice is (not so) blind: Effects of facial masculinity and agreeableness on perceptions of criminal guilt**. *Evol Behav Sci* 2022, **16**:67, https://doi.org/10.1037/ebs0000246.

16. Wilson JP, Rule NO: **Facial trustworthiness predicts extreme criminal-sentencing outcomes**. *Psychol Sci* 2015, **26**:1325-1331, https://doi.org/10.1177/0956797615590992.

17. Huang X, Ivković Z, Jiang JX, Wang IY: **Angel investment and first impressions**. *J Financ Econ* 2023, **149**:161–178, https://doi.org/10.1016/j.jfineco.2023.05.001.

18. Giacomin M, Rule NO: **How static facial cues relate to real-world leaders' success: a review and meta-analysis**. *Eur Rev Soc Psychol* 2020, **31**:120-148, https://doi.org/10.1080/10463283.2020.1771935.

19. François A, Panel S, Weill L: **Dictators' facial characteristics and foreign direct investment**. *Leadersh Q* 2023, **34**:101644, https://doi.org/10.1016/j.leaqua.2022.101644.

20. Wang Y, Kosinski M: **Deep neural networks are more accurate than humans at detecting sexual orientation from facial images**. *J Personal Soc Psychol* 2018, **114**: 246–257, https://doi.org/10.1037/pspa0000098.

21. Borkenau P, Brecke S, Möttig C, Paelecke M: **Extraversion is accurately perceived after a 50-ms exposure to a face**. *J Res Personal* 2009, **43**:703–706, https://doi.org/10.1016/j.jrp.2009.03.007.

22. Carney DR, Colvin CR, Hall JA: **A thin slice perspective on the accuracy of first impressions.** *J Res Personal* 2007, **41**:1054-1072, https://doi.org/10.1016/j.jrp.2007.01.004.

23. **Foo YZ, Sutherland CA, Burton NS, Nakagawa S, Rhodes G: **Accuracy in facial trustworthiness impressions: Kernel of truth or modern physiognomy? A meta-

analysis. *Personal Soc Psychol Bull* 2022, **48**:1580–1596,

https://doi.org/10.1177/01461672211048110.

24. Todorov A, Mandisodza AN, Goren A, Hall CC: **Inferences of competence from faces predict election outcomes**. *Sci* 2005, **308**:1623–1626,

https://doi.org/10.1126/science.1110589.

25. Penton-Voak IS, Pound N, Little AC, Perrett DI: **Personality judgments from natural and composite facial images: More evidence for a "kernel of truth" in social perception**. *Soc Cogn* 2006, **24**:607–640, https://doi.org/10.1521/soco.2006.24.5.607.

26. Samochowiec J, Wänke M, Fiedler K: **Political ideology at face value**. *Soc Psychol Personal Sci* 2010, **1**:206-213, https://doi.org/10.1177/1948550610372145.

27. Todorov A, Olivola CY, Dotsch R, Mende-Siedlecki P: **Social attributions from faces: Determinants, consequences, accuracy, and functional significance**. *Annu Rev Psychol* 2015, **66**:519-45, https://doi.org/10.1146/annurev-psych-113011-143831.

28. Zebrowitz LA, Rhodes G: **Sensitivity to "bad genes" and the anomalous face overgeneralization effect: Cue validity, cue utilization, and accuracy in judging intelligence and health**. *J Nonverbal Behav* 2004, **28**:167–185,

https://doi.org/10.1023/B:JONB.0000039648.30935.1b.

29. Hehman E, Sutherland CA, Flake JK, Slepian ML: **The unique contributions of perceiver and target characteristics in person perception**. *J Personal Soc Psychol* 2017, **113**:513, http://dx.doi.org/10.1037/pspa0000090.

30. Xie SY, Flake JK, Hehman E: **Perceiver and target characteristics contribute to impression formation differently across race and gender**. *J Personal Soc Psychol* 2019, **117**:364–385, https://doi.org/10.1037/pspi0000160.

31. Jones BC, DeBruine LM, Flake JK, Liuzza MT, Antfolk J, Arinze NC, Ndukaihe IL, Bloxsom NG, Lewis SC, Foroni F, Willis ML: **To which world regions does the valence–dominance model of social perception apply?** *Nat Hum Behav* 2021, **5**:159–169, https://doi.org/10.1038/s41562-020-01007-2.

32. Xie SY, Flake JK, Stolier RM, Freeman JB, Hehman E: **Facial impressions are predicted by the structure of group stereotypes**. *Psychol Sci* 2021, **32**: 1979–1993, https://doi.org/10.1177/09567976211024259.

33. Hassin R, Trope Y: **Facing faces: studies on the cognitive aspects of physiognomy.** *J Personal Soc Psychol* 2000, **78**:837–852, https://doi.org/10.1037//0022-3514.78.5.837.

34. **Madan S, Savani K, Johar GV: **How you look is who you are: The appearance reveals character lay theory increases support for facial profiling**. *J Personal Soc Psychol* 2022, **123**:1223–1242, https://doi.org/10.1037/pspa0000307.

35. Jaeger B, Todorov AT, Evans AM, van Beest I: **Can we reduce facial biases? Persistent effects of facial trustworthiness on sentencing decisions**. *J Exp Soc Psychol* 2020, **90**:104004, https://doi.org/10.1016/j.jesp.2020.104004.

36. Jaeger B, Evans A, Stel M, van Beest I: Lay beliefs in physiognomy explain overreliance on facial impression. Preprint, https://doi.org/10.31234/osf.io/8dq4x.

37. Suzuki A, Tsukamoto S, Takahashi Y: **Faces tell everything in a just and biologically determined world: Lay theories behind face reading**. *Soc Psychol Personal Sci* 2019, **10**:62-72, https://doi.org/10.1177/1948550617734616.

38. Ricanek K: *Using facial analytics in underwriting*. ITL. 2023. https://www.insurancethoughtleadership.com/life-health/using-facial-analytics-underwriting. Accessed 28 April 2024.

39. Hickman L, Bosch N, Ng V, Saef R, Tay L, Woo SE: **Automated video interview personality assessments: Reliability, validity, and generalizability investigations**. *J Appl Psychol* 2022, **107**:1323–1351, https://doi.org/10.1037/apl0000695.

40. Kachur A, Osin E, Davydov D, Shutilov K, Novokshonov A: **Assessing the Big Five personality traits using real-life static facial images**. *Sci Rep* 2020, **22**:8487, https://doi.org/10.1038/s41598-020-65358-6.

41. Moreno-Armendáriz MA, Martínez CA, Calvo H, Moreno-Sotelo M: **Estimation of personality traits from portrait pictures using the five-factor model**. *IEEE Access* 2020, **8**:201649-201665, https://doi.org/10.1109/ACCESS.2020.3034639.

42. Suen LW, Lunn MR, Katuzny K, Finn S, Duncan L, Sevelius J, Flentje A, Capriotti MR, Lubensky ME, Hunt C, Weber S: **What sexual and gender minority people want researchers to know about sexual orientation and gender identity questions: A qualitative study**. *Arch Sex Behav* 2020, **49**:2301–2318, https://doi.org/10.1007/s10508-020-01810-y.

43. Xu J, Tian W, Lv G, Liu S, Fan Y: **Prediction of the big five personality traits using static facial images of college students with different academic backgrounds.** *IEEE Access* 2021, **9**:76822-76832, https://doi.org/10.1109/ACCESS.2021.3076989.

44. Kosinski M: **Facial recognition technology can expose political orientation from naturalistic facial images**. *Sci Rep* 2021, **11**:100, https://doi.org/10.1038/s41598-020-79310-1.

45. **Kosinski M, Khambatta P, Wang Y: **Facial recognition technology and human raters can predict political orientation from images of expressionless faces even when

controlling for demographics and self-presentation. *Am Psychol* 2024, https://doi.org/10.1037/amp0001295. Advance online publication.

46. Rasmussen SH, Ludeke SG, Klemmensen R: **Using deep learning to predict ideology from facial photographs: expressions, beauty, and extra-facial information**. *Sci Rep* 2023, **13**:5257, https://doi.org/10.1038/s41598-023-31796-1.

47. Wu X, Zhang X: Responses to critiques on machine learning of criminality perceptions. Preprint, https://doi.org/10.48550/arXiv.1611.04135.

48. Hsieh TS, Kim JB, Wang RR, Wang Z: **Seeing is believing? Executives' facial trustworthiness, auditor tenure, and audit fees**. *J Account Econ* 2020, **69**:101260, https://doi.org/10.1016/j.jacceco.2019.101260.

49. Peng L, Teoh SH, Wang Y, Yan J: **Face value: Trait impressions, performance characteristics, and market outcomes for financial analysts**. *J Account Res* 2022, **60**: 653-705, https://doi.org/10.1111/1475-679X.12428.

50. **y Arcas BA, Mitchell M, Todorov A: **Physiognomy in the Age of AI.** In *Feminist AI: Critical Perspectives on Algorithms, Data, and Intelligent Machines*. Edited by Browne J, Cave S, Drage E, McInerney K, Oxford Academic; 2023:208–236, https://doi.org/10.1093/oso/9780192889898.003.0013.

51. ** Ullstein C, Engelmann S, Papakyriakopoulos O, Hohendanner M, Grossklags J: **AI-competent individuals and laypeople tend to oppose facial analysis AI**. In *Proc of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*. 2022:1-12, https://doi.org/10.1145/3551624.3555294.

52. Stark L: **Artificial intelligence and the conjectural sciences**. *BJHS Themes* 2023, **8**:35-49, https://doi.org/10.1017/bjt.2023.3.

53. Buolamwini J, Gebru T: **Gender shades: Intersectional accuracy disparities in commercial gender classification**. In *conference on fairness, accountability and transparency*. Proc Mach Learn Res; 2018:77-91, https://proceedings.mlr.press/v81/buolamwini18a.html.

54. Raghavan M, Barocas S, Kleinberg J, Levy K: **Mitigating bias in algorithmic hiring: Evaluating claims and practices**. In *Proc of the 2020 conference on fairness, accountability, and transparency.* 2020:469-481, https://doi.org/10.1145/3351095.3372828.

55. **Stark L, Hutson J: **Physiognomic artificial intelligence**. *Fordham Intell Prop Media & Ent LJ* 2022, **32**:922, https://ir.lawnet.fordham.edu/iplj/vol32/iss4/2.

56. Garvie C, Frankle J: *Facial-recognition software might have a racial bias problem. The Atlantic.* 2016. https://www.theatlantic.com/technology/archive/2016/04/the-underlying-bias-of-facial-recognition-systems/476991/. Accessed 28 April 2024.

57. Santow E: **Can artificial intelligence be trusted with our human rights?** *AQ: Aust Q* 2020, **91**:10–17, https://www.jstor.org/stable/26931483.

58. **Wang D: **Presentation in self-posted facial images can expose sexual orientation: Implications for research and privacy**. *J Personal Soc Psychol* 2022, **122**:806–824, https://doi.org/10.1037/pspa0000294.

59. Engelmann S, Ullstein C, Papakyriakopoulos O, Grossklags J: **What people think AI should infer from faces.** In *Proc of the 2022 ACM Conference on Fairness, Accountability, and Transparency.* 2022:128-141, https://doi.org/10.1145/3531146.3533080.

60. Sheetal, A, Savani, K: **A machine learning model of cultural change: Role of prosociality, political attitudes, and Protestant work ethic**. *Am Psychol* 2021, **76**: 997-1012. https://psycnet.apa.org/record/2022-13889-012

61. Tian L, Alaei R, Rule NO: **Appearance reveals music preferences**. *Personal Soc Psychol Bull* 2022, **48**:1635-1650, https://doi.org/10.1177/01461672211048291.

62. Alaei R, Deska JC, Hugenberg K, Rule NO: **People attribute humanness to men and women differently based on their facial appearance**. *J Personal Soc Psychol* 2022, **123**:400, https://doi.org/10.1037/pspi0000364.

Further Reading

Further information on references of particular interest

5** This research focuses on the validity of face-based inferences among human participants. While facial inferences accurately predicted behavior, perceivers' confidence in their judgments was unrelated to accuracy, underscoring that people lack meta-cognition about their facial inferences. Finally, the authors provide evidence that the accuracy of facial inferences may possibly be due to self-fulfilling effects—driven by perceivers' beliefs rather than by an actual link between face and character.

10** This research shows that people's face-based inferences are highly resistant to updating. People updated their implicit face-based inferences only when they encountered extreme and credible counter-information.

13** Using deep learning, this research models the perceptual bases of facial impressions of 34 perceived social and physical attributes such as age, attractiveness, political orientation, and trustworthiness. The model is, therefore, able to generate faces that vary on these

attributes, manipulate facial stimulus to be high or low in these attributes, and predict the impressions a specific face will evoke in a North American audience.

23** Across two meta-analyses, the results suggest that there might be a very modest kernel of truth (r = .14) in people's impressions of trustworthiness from faces. The accuracy is somewhat higher with traits such as aggressiveness and unfaithfulness (which are more clearly physiological) than traits that are not—such as honesty or cheating.

34** Across ten studies, this research investigates a novel antecedent for people's overconfidence in their appearance-based judgments—the lay theory that an individual's appearance reveals their character. This, in turn, increases people's support for AI-based facial inference for consequential uses such as security, employment, education, and financial services.

45** In this research, the authors employed an AI-based algorithm to predict people's political orientation from standardized images. The accuracy (r = .22) was comparable to human raters (r = .21). The algorithm also predicted the political orientation of 3,401 politicians from the U.S., U.K., and Canada with a correlation of approximately .13.

50** This chapter argues that AI-based facial inference is the modern reincarnation of physiognomy. Using two case studies, the authors explain how AI-based facial inference propagates harmful stereotypes and can lead to discrimination and injustice.

51** This research examines the ethical evaluations of facial inferences made by AI, focusing on how people with AI expertise and laypeople perceive AI's ability to make inferences from human faces. While both groups do not trust many AI-based facial inferences, they differ in the reasons for doing so.

55** This article provides a legal studies perspective on AI-based facial inference, which the authors refer to as "physiognomic AI." Advocating for urgent legislative intervention, the article recommends declaring such technologies as inherently unfair and deceptive, enhancing biometric privacy laws, and banning their use in public spaces.

58** This study investigated whether facial profiling algorithms' ability to discern sexual orientations is driven by the distinct self-presentation styles of people of different sexual orientations. Significant differences were found; for instance, gay individuals tended to wear glasses more frequently and upload brighter images compared to heterosexual individuals.