


Complex faces and naïve machines

A commentary on facial perceptions of age, gender, and leader preferences in the age of AI

Brian R. Spisak , *Harvard University; HSC Analytics*

ABSTRACT. Tasks driven by artificial intelligence (AI), such as evaluating video job interviews, rely on facial recognition systems for decision-making. Therefore, it is extremely important that the science behind this technology is continually advancing. If not, visual stereotypes, such as those associated with facial age and gender, will lead to dangerous misapplications of AI.

Keywords: face perception, age, gender, leadership, voting, facial recognition systems

Face perception research offers broad and compelling insights, with many useful (yet potentially worrying) applications. From inferences about personality (Walker & Vetter, 2016) to the growing use of facial recognition systems in crowd monitoring (Veltmeijer et al., 2021) and the hiring process (Suen et al., 2019), this work is here to stay. Therefore, it is extremely important that face perception research guards against harmful biases through theoretical rigor, responsible development, and prudent implementation. To serve as an example, I comment on (1) the continued (and outdated) use of gender as a binary factor in face perception research (see Shen & Shoda, 2021, as a recent example), (2) the need to further explore the value of age diversity (see Tuncdogan et al., 2017, for a discussion of existing and future work), and (3) the importance of expanding theoretical horizons (rather than the excessive “gap-spotting” outlined in Sandberg & Alvesson, 2011). In short, facial recognition systems need to be based on state-of-the-art social science research, not antiquated (and discrimination-inducing) output.

Here are three opportunities for safeguarding against misguided research and practice:

1. Moving past binary gender

When I decided to use faces, it was not because I was necessarily interested in face perception research. Instead, my colleagues and I realized it was a clever

way to investigate the implicit biases that followers hold regarding nonbinary aspects of a leader’s gender (e.g., facial masculinity versus femininity; see Little et al., 2007; Spisak, Homan et al., 2012). My hypothesis then, and now, is that follower decision-making regarding gender is more sensitive than men versus women. We clearly vary beyond assigned biological sex, and there are important nonbinary signals (or at least cues) influencing leader selection that should not go overlooked—especially given the growing appreciation that a binary understanding of gender in psychological research is an outdated concept (Hyde et al., 2019). Instead, face perception research should explore new spaces, such as how voting preferences are changing now that society is starting to appreciate the realities of gender. A new world is emerging that is far more sensitive to differences in assigned sex, gender, and gender identities. Social science research needs to help facial recognition systems keep up with this liberating time (see Scheurman et al., 2019).

2. Considering the value of age diversity

Though modern research finds fluctuations in leader preferences based on age and life span, much of the foundation is atheoretical or not entirely representative of reality. Shen and Shoda (2021), for instance, found that the likelihood of voting for a candidate decreases as the candidate’s perceived age increases—noting that this negative relationship starts for male candidates around a perceived age of 45. However, this finding stands in stark contrast with almost all congressional and parliamentary profiles around the world (e.g., the average age of the current U.S. House of Representatives is 58.4 and the

doi: [10.1017/pls.2021.30](https://doi.org/10.1017/pls.2021.30)

Correspondence: Brian R. Spisak, Harvard University, Cambridge, MA, USA. Email: bspisak@hsph.harvard.edu

U.S. Senate is 64.3; see Manning, 2021). In reality, younger candidates have the hardest time gaining status and leadership roles (Inter-Parliamentary Union, 2021). It is, therefore, important to explore this youth barrier for insights and opportunities—for example, previous work has found that younger leaders are preferred for establishing and maintaining peace and cooperation (Spisak, 2012) as well as leading change (Spisak et al., 2014).

The next generation of face perception research needs to build on such findings to better understand and activate the value of age diversity in leadership and society. Social science, for example, can inform facial recognition systems using age to identify leaders who are more likely to encourage the reduction of fossil fuels and the push toward renewable alternatives (see the “green leadership” example in Spisak et al., 2014).

3. Contributing to better practice and policy

The use of facial recognition technology is developing rapidly, even though its reliability and validity is a perennial topic of debate. From the U.S. Transportation Security Administration’s application of “observation techniques” based on the work of Ekman and colleagues (see Heaven, 2020) to recent publications suggesting that emerging facial recognition technology is based on outdated psychological science (e.g., Barrett et al., 2019), the foundation is shaky. Yet, despite this uncertainty (and explicit flaws in some cases), the public and private sectors are pushing forward with automated systems for many tasks, such as assessing personality, monitoring (crowd) emotions, and evaluating job applicants (e.g., Suen et al., 2019; Veltmeijer et al., 2021). Vendors often use terms like “scientifically validated” to reassure investors, clients, and the public that any costs associated with bias, discrimination, and inequality are minimal (Harlan & Schnuck, 2021). Also, facial recognition systems are advancing (whether society likes it or not) because they are relevant in many domains and profitable in practice. Therefore, these AI systems must be monitored and shaped so they can become human centered and uplifting, not biased and draconian.

Conclusion

The ubiquity of facial recognition calls for rigorous, multidisciplinary research to ensure the implementation of robust face perception insights. Researchers and

practitioners need to thoroughly explore relevant literature, move beyond the binary, and investigate the value of diversity. Otherwise, “scientifically validated” technology will continue to incorporate false assumptions. As social scientists, we must safeguard against this unreliable and invalid future.

References

- Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest*, 20(1), 1–68. <https://doi.org/10.1177/1529100619832930>
- Harlan, E., & Schnuck, O. (2021). *Objective or biased: On the questionable use of Artificial Intelligence for job applications*. Bayerischer Rundfunk & Report München. Retrieved November 2, 2021, from <https://web.br.de/interaktiv/ki-bewerbung/en/>
- Heaven, D. (2020). Why faces don’t always tell the truth about feelings. *Nature*, 578(7796), 502–505. <https://doi.org/10.1038/d41586-020-00507-5>
- Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *American Psychologist*, 74(2), 171–193. <https://doi.org/10.1037/amp0000307>
- Inter-Parliamentary Union. (2021). *Youth participation in national parliaments*. Retrieved November 2, 2021, from <https://www.ipu.org/youth2021>
- Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. (2007). Facial appearance affects voting decisions. *Evolution and Human Behavior*, 28(1), 18–27. <https://doi.org/10.1016/j.evolhumbehav.2006.09.002>
- Manning, J. E. (2021, August 5). *Membership of the 117th Congress: A Profile*. Congressional Research Service. <https://crsreports.congress.gov/product/pdf/R/R46705>
- Sandberg, J., & Alvesson, M. (2011). Ways of constructing research questions: Gap-spotting or problematization? *Organization*, 18(1), 23–44. <https://doi.org/10.1177/1350508410372151>
- Scheurman, M. K., Paul, J. M., & Brubaker, J. R. (2019). How computers see gender: An evaluation of gender classification in commercial facial analysis services. *Proceedings of the ACM on Human-Computer Interaction*, 3, 1–33. <https://doi.org/10.1145/3359246>
- Shen, Y. A., & Shoda, Y. (2021). How candidates’ age and gender predict voter preference in a hypothetical election.

Psychological Science, 32(6), 934–943. <https://doi.org/10.1177/0956797620977518>

Spisak, B. R. (2012). The general age of leadership: Older-looking presidential candidates win elections during war. *PLOS ONE*, 7(5), e36945. <https://doi.org/10.1371/journal.pone.0036945>

Spisak, B. R., Grabo, A. E., Arvey, R. D., & Van Vugt, M. (2014). The age of exploration and exploitation: Younger-looking leaders endorsed for change and older-looking leaders endorsed for stability. *The Leadership Quarterly*, 25(5), 805–816. <https://doi.org/10.1016/j.leaqua.2014.06.001>

Spisak, B. R., Homan, A. C., Grabo, A., & Van Vugt, M. (2012). Facing the situation: Testing a biosocial contingency model of leadership in intergroup relations using masculine and feminine faces. *The Leadership Quarterly*, 23(2), 273–280. <https://doi.org/10.1016/j.leaqua.2011.08.006>

Suen, H. Y., Hung, K. E., & Lin, C. L. (2019). TensorFlow-based automatic personality recognition used in asynchronous video interviews. *IEEE Access*, 7, 61018–61023. <https://doi.org/10.1109/ACCESS.2019.2902863>

Tuncdogan, A., Acar, O. A., & Stam, D. (2017). Individual differences as antecedents of leader behavior: Towards an understanding of multi-level outcomes. *The Leadership Quarterly*, 28(1), 40–64. <https://doi.org/10.1016/j.leaqua.2016.10.011>

Veltmeijer, E. A., Gerritsen, C., & Hindriks, K. (2021). Automatic emotion recognition for groups: A review. *IEEE Transactions on Affective Computing*. <https://doi.org/10.1109/TAFFC.2021.3065726>

Walker, M., & Vetter, T. (2016). Changing the personality of a face: Perceived Big Two and Big Five personality factors modeled in real photographs. *Journal of Personality and Social Psychology*, 110(4), 609–634. <https://doi.org/10.1037/pspp0000064>