Exploring the influence of word definition on AI-generated facial representations

Vojtěch Fiala[*1], Przemyslaw Zywiczynski[1], Anna Szala[1], Marek Placiński[1], Slawomir Wacewicz[1]

[*]Corresponding Author: vfiala@umk.pl
[1]Center for Language Evolution Studies, Nicolaus Copernicus University, Toruń, Poland

One of the key questions in language evolution is the role of linguistic labels in driving cognitive evolution, esp. by influencing categorisation, including visual perceptual stimuli (Lupyan 2006, Lupyan & Casasanto 2015). Interestingly, this problem finds parallels in a large swathe of human-evolutionary research since a standard study format in evolutionary psychology is to rate visual stimuli – very often human faces – on characteristics presented as verbal labels (Langlois et al., 2000). Even closely related psychological concepts are defined by different words and cued by distinctive facial configurations (e.g., Mileva, 2016). For example, a face imagined as "socially attractive" (a face appealing in the social context) may possess different characteristics than a "sexually attractive face" (Kruger, 2006). Our study therefore explores the problem of the interface between facial characteristics and their linguistic descriptions.

Despite the impact that a linguistic formulation may exert on perception (Lupyan et al. 2020), studies on perceived facial characteristics, e.g., attractiveness, usually do not provide a definition of the focus characteristic they want the participants to rate: out of 65 on the topic from 2021–2023, we found only 1 paper that did this. While a majority of these articles concern mate choice, others focus on social psychology, economics, and political sciences. Given this broad diversity of contexts in which 'attractiveness' was studied, we decided to test whether a deep learning text-to-image model, Stable Diffusion XL, returns detectably different images when prompted for sexual vs social attractiveness. Trained on ~6 billion image-text pairs, the model presents a potent source of images accompanied by name descriptions. We used one prompt for sexual attractiveness (*prompt 1*: "attractive European man/woman, sexual mating and partnership context") and another for social attractiveness (*prompt 2*: "attractive prosocial, friendly, and cooperative European man/woman" with "sexy" and "good mate" as negative prompts).

For each of the two prompts, we generated 120 facial images (60 women, 60 men), and applied standard selection criteria in evolutionary psychology: full face visible, not horizontally or vertically tilted, closed mouth, neutral expression, and facial features not covered by hair/facial hair, resulting in 108 faces (54 W, 54 M). The faces were subject to geometric morphometrics analysis. We marked each face with 72 landmarks (see Kleisner et al., 2019) in the program tpsDig (Rohlf, 2015). We ran a generalised Procrustes analysis (separately for each sex) using package geomorph in R (Baken et al. 2021; Adams, 2023), and computed the average configuration for each category (sexual vs. social, Figure 1). We then calculated the distance of each facial configuration from its ingroup/outgroup mean configuration. A subsequent analysis with the function permudist of the package Morpho (Schlager, 2017) suggests that the faces tend to cluster around its groups' average configuration (Procrustes distance between social-sexual group means [PDM] = 0.00030, $p < 0.001$ for men, PDM = 0.00032, $p < 0.001$ for women). Although numerically small, the difference between groups is significant.

The results suggest that AI-generated representations of human faces (for this preliminary study limited to the faces of European origin) systematically differ depending on the particular phrasing of the prompt, to the degree that AI-generated socially vs. sexually attractive faces can be distinguished by their geometric-morphometrics properties. This novel approach to visualising the variance of human facial characteristics based on their linguistic description can be flexibly applied to the faces of people from other ethnic backgrounds and can be extended with the application of AI graphic tools (currently, we are using the text-to-image generator Dall-e, and faces of models from populations outside Europe). Our future research aims to include automated estimates of facial attractiveness, real human faces, human raters, and different facial characteristics to test whether both human-made and automated facial characterisations are sensitive to word definitions.

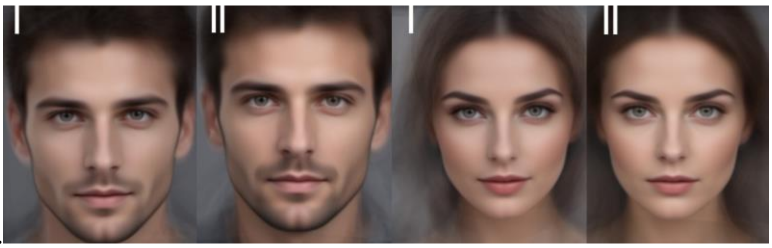

Figure 1. Representations of sexually (I) vs socially (II) attractive faces of 'men' (left) and 'women' (right), based on images created by Stable Diffusion.

Acknowledgements

We are thankful to prof. Karel Kleisner for introducing us to the methods of geometric morphometrics.

Short summary:
The role of linguistic labels in categorisation and determining the scope of a category may have large consequences for ascribing characteristics to visual stimuli. Despite that, studies on perceived facial characteristics, including attractiveness, usually do not provide the readers with a definition of the rated category. Using the deep learning text-to-image model Stable Difusion, we created average visual representations of two distinctive definitions of facial attractiveness: socially and sexually attractive faces. Subsequently, we used tools of landmark-based geometric morphometrics to explore if the visual representations differ across the definitions. The results suggest that the AI-generated representations of attractive human faces are context- and verbal description-dependent and that faces generated with social and sexual attractiveness primes can be distinguished by methods of geometric morphometrics. This result is relevant in the context of links between the perceptual salience of facial characteristics and linguistic expressions referring to these characteristics.

OSF link to the project: https://osf.io/qund8/?view_only=c5654a1ffe5c497e840702b00cf123ec

References

Adams, D., Collyer, M., Kaliontzopoulou, A., & Baken, E. (2023). Geomorph: Software for geometric morphometric analyses. R package version 4.0.6. https://cran.r-project.org/package=geomorph.

Baken, E. K., Collyer, M. L., Kaliontzopoulou, A., & Adams, D. C. (2021). geomorph v4. 0 and gmShiny: Enhanced analytics and a new graphical interface for a comprehensive morphometric experience. *Methods in Ecology and Evolution, 12*(12), 2355–2363.

Kleisner, K., Pokorný, Š., & Saribay, S. A. (2019). Toward a new approach to cross-cultural distinctiveness and typicality of human faces: The cross-group typicality/distinctiveness metric. *Frontiers in psychology, 10*, 124.

Kruger, D. J. (2006). Male facial masculinity influences attributions of personality and reproductive strategy. *Personal Relationships, 13*(4), 451–463.

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological bulletin, 126*(3), 390.

Lupyan, G. 2006. Labels Facilitate Learning of Novel Categories. In: Cangelosi, A., Smith, A.D.M, & Smith, A. (eds.) 2006. The Evolution of Language. Proceedings of the 6th International Conference (EVOLANG6). Singapore: World Scientific Publishing.

Lupyan, G., & Casasanto, D. (2015). Meaningless words promote meaningful categorization. *Language and Cognition*, 7(2), 167-193.

Lupyan, G., Rahman, R. A., Boroditsky, L., & Clark, A. (2020). Effects of language on visual perception. *Trends in cognitive sciences*, 24(11), 930-944.

Mileva, V. (2016). Social status in humans: Differentiating the Cues to dominance and prestige in men and women. Dissertation Thesis. University of Stirling

Rohlf, F. J. (2015). The tps series of software. *Hystrix, 26*(1), 9–12.

Schlager, S. (2017). Morpho and Rvcg–Shape Analysis in R: R-Packages for geometric morphometrics, shape analysis and surface manipulations. In Statistical shape and deformation analysis (pp. 217–256). *Academic Press*.