

Cultural and Visual Determinants of Avatar-Based Impression: A Cross-National Study on Relaxing, Trustworthiness, and Emergency Suitability

Liwen Zhang¹, Takashi Sakamoto¹, Toru Nakata^{1,2},
and Toshikazu Kato¹

¹Graduate School of Science and Engineering, Chuo University, Tokyo, Japan

²National Institute of Advanced Industrial Science and Technology, Tokyo, Japan

ABSTRACT

As AI-generated avatars become more common in public communication, understanding how facial features shape impressions across cultures is crucial. This study investigated how geometric facial traits influence perceived trustworthiness, relaxing impression, and emergency suitability among Japanese and Chinese participants. A total of 288 participants evaluated 24 AI-generated faces selected through Principal Component Analysis based on eye and mouth-related features. Eye-related features acted as universal predictors, whereas smile-related features varied culturally. These findings highlight the need for culturally adaptive avatar design, especially in high-stakes contexts such as emergency communication.

Keywords: AI-generated avatars, Facial impression formation, Cross-cultural perception, Trust and emergency communication, Principal component analysis

INTRODUCTION

AI-generated avatars are increasingly used in critical scenarios such as emergency guidance, digital counselling, and public announcements. Understanding how their facial features influence user impressions—especially Trustworthiness (*Tw*), Relaxing Impression (*RI*), and Emergency Suitability (*ES*), is essential to foster trust and responsiveness.

While psychological theories like Mayer's trust model (Mayer et al., 1995) have outlined the foundations of interpersonal trust, relatively few studies have examined how specific geometric facial traits in realistic avatars influence these three dimensions. Moreover, most existing work focuses on general trust perceptions rather than contexts that demand urgency and calmness, such as emergencies.

This study addresses this gap via a cross-cultural comparison (Japan vs. China) of impressions toward AI-generated avatars. By analyzing eye size, face proportions, and mouth shape, we identify universal and culture-specific predictors of avatar-based trust.

RELATED WORK

Trustworthiness and dominance are key facial impression dimensions across cultures (Oosterhof and Todorov, 2008). Features like large eyes and subtle smiles influence perceptions of competence and warmth (Hu et al., 2021; Bhattacharya et al., 2020), especially in mediated contexts where visual cues shape engagement.

Importantly, cross-cultural differences shape how facial signals are interpreted. Studies suggest that East Asians tend to process faces holistically and rely more on eye-region features, while Western viewers focus more on mouth-related cues when decoding emotions (Jack et al., 2009; Yuki et al., 2007). In collectivist societies like Japan and China, subtle emotional cues and neutral expressions are often emphasized to maintain social harmony (Markus and Kitayama, 1991), potentially affecting how users perceive trust or urgency.

Few studies have modeled how facial measurements relate to multiple impression dimensions in cross-cultural settings. Emergency Suitability (*ES*), in particular, remains underexplored in urgent-avatar contexts. Building on this foundation, the present study aims to explore how AI-generated facial structures influence three key impression dimensions: *RI*, *Tw*, and *ES*, across two East Asian cultures.

RESEARCH HYPOTHESES

This study investigates how trust toward AI-generated avatars is cognitively structured and shaped by cultural context. Based on prior psychological and cross-cultural theories, we propose the following hypotheses:

- **H1 (Three-Dimensional Trust Hypothesis):** Trust in avatars consists of three distinct but interrelated components: Relaxing Impression (*RI*), Trustworthiness (*Tw*), and Emergency Suitability (*ES*).
- **H2 (Cultural Moderation Hypothesis):** Cultural background modulates the interpretation of facial features in avatar evaluations. Specifically, Japanese participants, due to cultural emphasis on emergency responsiveness are expected to prioritize Emergency Suitability, whereas Chinese participants who may value interpersonal harmony are expected to place greater emphasis on Relaxing Impression.

Given prior research associating eye cues with perceived competence and dominance, and smile curvature with emotional warmth or tension, we further hypothesize that *PC1* (which is about eye-related traits) would be a primary predictor of Trustworthiness, while *PC2* (which concerns smiling mouth shape) may play a more important role in *Relaxing Impression* and *Emergency Suitability*.

METHOD

Participants

A total of 288 participants (144 from Japan and 144 from China, aged 18–65, 50% male/female) were recruited via university mailing lists and online platforms. All reported frequent internet use and regular exposure to virtual avatars, ensuring comparable familiarity across groups.

Context and Stimuli

We used 300 AI-generated face images (Generated Photos, Icons8) systematically varied by age, sex, and ethnicity to examine how facial features influence impressions.

In the experiment, each avatar appeared as the presenter of an emergency evacuation plan, providing a consistent context for evaluation. Since these avatars were generated from commercial web-trained models, potential demographic biases (e.g., underrepresentation or averaging effects) may affect cross-cultural interpretations.

Facial Feature Extraction and PCA-Based Stimulus Selection

We extracted five geometric facial features from each avatar using a consistent landmark-based method (Fig. 1):

- *Facial width-to-height ratio (fWHR)*: distance between zygions divided by upper face height.
- *Eye-to-face ratio (EFR)*: eye area divided by total face area.
- *Eye size index (ESI)*: eye height divided by eye width.
- *Mouth-Width*: horizontal distance between mouth corners.
- *Smile-Arc-Angle*: cosine of the angle formed by vectors from mouth corners to the philtrum.

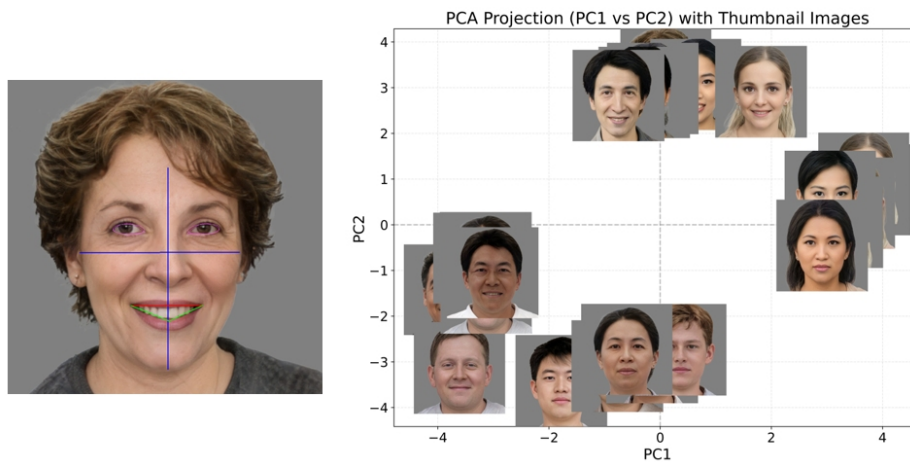


Figure 1: Facial landmark visualization (left) and PCA projection of avatars (right), highlighting *pc1*. (*Large and horizontally wide eyes*) and *PC2* (*Narrow smiling mouth*). Representative extreme avatars are shown as thumbnails.

These features were selected based on prior research linking facial geometry to impression formation (Oosterhof and Todorov, 2008; Hu et al., 2021; Song et al., 2021). Principal Component Analysis (PCA) with varimax rotation was conducted to identify latent structure patterns. Although the KMO value was marginal (0.539), Bartlett's test was significant ($\chi^2 = 807.30$, $p < .001$), indicating acceptable factorability. Based on eigenvalues > 1 and scree plot inspection, three components were retained (*PC1–PC3*), explaining 88.5% of total variance:

- *PC1*: “Large and horizontally wide eyes” (high *EFR*, low *ESI*).
- *PC2*: “Narrow smiling mouth” (high *Smile-Arc-Angle*, low *Mouth-Width*).
- *PC3*: “Wide and short face” (high *fWHR*).

To ensure variation in visual stimuli, avatars with extreme *PC1–PC3* scores (top and bottom 10%) were selected for the experiment.

Measurement of Impressions

Participants rated 24 avatars on three impression dimensions: Relaxing Impression, Trustworthiness, and Emergency Suitability, using a 7-point Likert scale (1 = “strongly disagree”, 7 = “strongly agree”). Each dimension comprised three items, including one reverse-coded statement. The full questionnaire is presented in Table 1.

Table 1: Impression evaluation items for Avatar rating (7-point Likert scale).

Aspect	Item	Questionnaire Statement
Relaxing Impression (<i>RI</i>)	R1	I feel that this avatar helps me relax.
	R2(Rev)	I feel tense when seeing this avatar (reverse-coded).
	R3	I feel a sense of safety when looking at this avatar.
Trustworthiness (<i>Tw</i>)	T1	I find the information conveyed by this avatar trustworthy.
	T2	I feel that the behavior of this avatar is consistent.
	T3(Rev)	The appearance of this avatar seems hard to trust (reverse-coded).
Emergency Suitability (<i>ES</i>)	E1	This avatar can effectively convey urgent information.
	E2(Rev)	This avatar does not seem suitable for emergency notification (reverse-coded).
	E3	I perceive a sense of urgency from this avatar.

RESULTS

Cross-Cultural Differences in Avatar Impression Scores

To examine cultural and visual effects on avatar impression ratings, we conducted two-way ANOVAs with Country (Japan vs. China) and Image (1–24) as between-subject factors across three dimensions: *RI*, *Tw* and *ES*.

Significant main effects of Country and Image were found for all dimensions (all p s <.001). Japanese participants consistently rated avatars higher than Chinese participants, especially in *Tw* and *ES*.

Significant Country $\times$ Image interactions emerged for *RI* (p <.001) and *ES* (p = .033), but not for *Tw*.

These results indicate that both cultural background and avatar appearance influence impression ratings, with Japanese participants generally providing more favorable evaluations (see Fig. 2).



Figure 2: Mean ratings of 24 avatars on three impression dimensions (*RI*, *Tw*, and *ES*). Red and blue lines represent responses from Chinese and Japanese participants, respectively. Shaded areas indicate 95% confidence intervals.

Relationship Among Latent Impression Factors and Participants' Answers

To examine the structural relationships among *RI*, *ES* and *Tw*, we conducted separate structural equation modelling (SEM) analyses for Chinese and Japanese participants.

For the Chinese group, results supported a full sequential mediation model (see Fig. 3, left panel):

- $RI \rightarrow ES$ ($\beta = 0.52, p <.001$)
- $ES \rightarrow Tw$ ($\beta = 0.32, p <.001$)
- $RI \rightarrow Tw$ ($\beta = 0.47, p <.001$)

For the Japanese group, a similar pattern emerged, with some variations:

- $RI \rightarrow ES$ ($\beta = 0.56, p <.001$)
- $ES \rightarrow Tw$ ($\beta = 0.46, p <.001$)
- $RI \rightarrow Tw$ ($\beta = 0.42, p <.001$)

All factor loadings were high and significant across both groups, supporting the validity of the measurement model. However, the indirect pathway from *ES* to *Tw* was notably stronger in the Japanese group, suggesting a cultural difference in how urgency-related traits influence perceptions of trust.

Model fit indices indicated good to excellent fit:

- Chinese model: $\chi^2(24) = 167.95, p <.001$; CFI = .992; RMSEA = .072
- Japanese model: $\chi^2(24) = 114.93, p <.001$; CFI = .996; RMSEA = .057

While both models achieved strong CFI values ($> .95$), the slightly higher RMSEA in the Chinese model may reflect greater intra-group variability in impression processing.

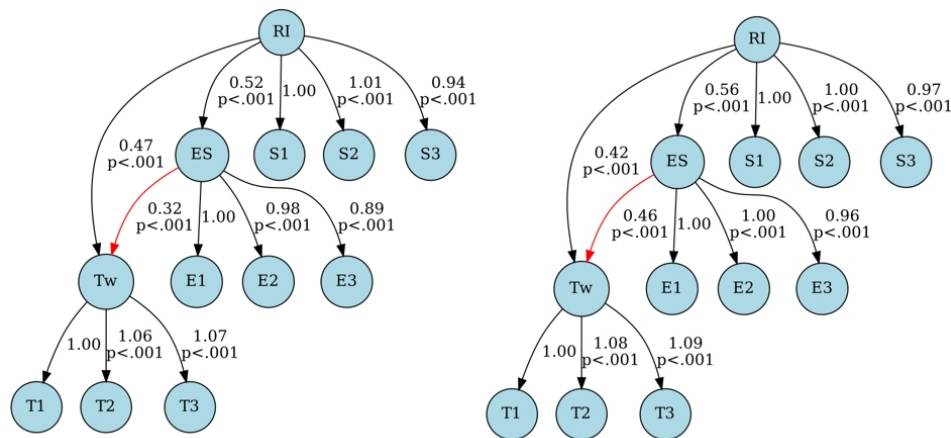


Figure 3: Illustrates the SEM models and highlights culturally significant differences in red. Left: Chinese group. Right: Japanese group.

These findings suggest a shared underlying structure across cultures, but with nuanced differences in pathway strength. In particular, *ES* plays a more prominent role in shaping trust judgments among Japanese participants.

Integrated Path Analysis of Facial Features and Impressions

To understand how demographic factors shape the facial appearance of avatars, we analysed the left portion of the path model (see Fig. 4), focusing on:

1. Demographic attributes (*Age, Sex, Race*),
2. Latent facial configurations (“*Eye Area Configuration*” and “*Mouth Smile*”).

The eye-related component, which loads positively on *Eye-to-Face Ratio* (*EFR*) and negatively on *Eye Size Index* (*ESI*), was positively associated with Age ($\beta = .12$), and negatively with Sex ($\beta = -.38$) and Race ($\beta = -.08$). This suggests that avatars perceived as older, male, and non-Asian tend to have larger eye area ratios but proportionally shorter vertical eyes. The smile-related component (“*Mouth Smile*”), represented by *Smile-Arc-Angle* and *Mouth-Width*, showed only weak or non-significant associations with demographic variables. These findings indicate that the demographic presets used in AI generation systematically influence facial structure, particularly eye-related geometry, which is known to shape trust-related impressions. This step serves as a foundational layer for understanding how impressions are indirectly influenced by facial geometry.

Cultural Comparison of PCA Components and Impression Ratings

To examine cultural differences in how PCA components relate to impression ratings, we conducted separate regression analyses for the Chinese and

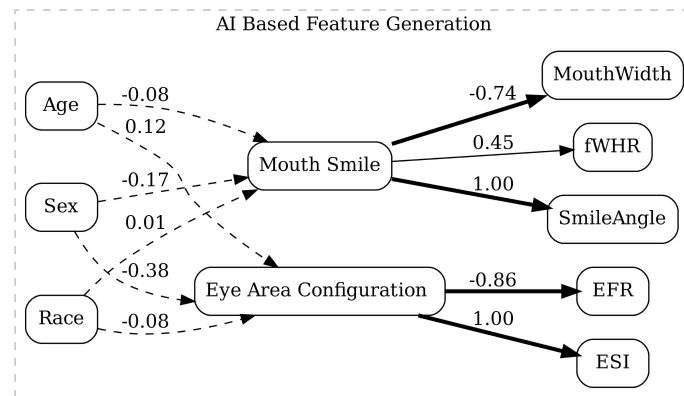


Figure 4: Partial path model showing how demographic attributes (Age, Sex, Race) influence two latent facial structure factors: “eye area configuration” and “mouth smile.” Solid lines indicate significant paths ($p < .05$), and dashed lines indicate weak or non-significant associations.

Japanese groups. Figure 5 presents these findings: the left panel shows standardized beta coefficients for the Chinese group, and the right panel visualizes the beta differences (JP–CN).

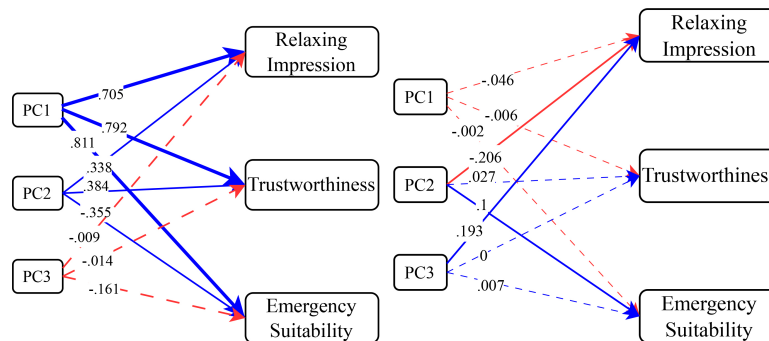


Figure 5: Left: Standardized regression paths from PCA components ($pc1$ – $pc3$) to impression dimensions in the Chinese group. Right: Beta differences (Japan – China). Blue/red = positive/negative difference; solid = $|\Delta\beta| \geq .10$; dashed = $|\Delta\beta| < .10$; line thickness reflects significance in the Chinese group.

In the Chinese group (Fig.5, left), PC1 was a strong and consistent predictor across all impression dimensions: *RI* ($\beta = .705$, $p < .001$), *Tw* ($\beta = .792$, $p < .001$), and *ES* ($\beta = .811$, $p < .001$). PC2 showed moderate effects on *Tw* ($\beta = .384$, $p = .012$), *ES* ($\beta = .355$, $p = .017$), and a marginal effect on *RI* ($\beta = .338$, $p = .046$). PC3 did not significantly predict any of the impression outcomes.

The right panel shows beta differences (Japan minus China). Line styles indicate the extent of difference: solid lines for $|\Delta\beta| \geq .10$, dashed for smaller differences; blue for positive differences (JP > CN), red for negative (JP < CN). Notably, $PC2 \rightarrow RI$ showed a substantial negative difference ($\Delta\beta = -.206$), suggesting that Japanese participants were less influenced by PC2

(*Narrow smiling mouth*) when judging relaxing impressions. $PC2 \rightarrow ES$ showed a modest positive difference ($\Delta\beta = +0.100$), indicating slightly greater sensitivity among Japanese participants. Interestingly, $PC3 \rightarrow RI$ also showed a notable positive difference ($\Delta\beta = +0.193$), implying a culture-specific role of facial width in evaluating relaxing impressions.

Overall, results support H2, highlighting cultural moderation: both groups rely strongly on $PC1$, but Chinese participants appear more sensitive to $PC2$ for evaluating impressions, whereas Japanese participants may attribute more meaning to facial width ($PC3$) when judging relaxation.

Three-Layer Path Model of Feature PCA Impression

To clarify how facial geometry contributes to impression formation, we constructed a three-layer path model (Fig. 6). The left portion visualizes the standardized loadings from the five geometric facial features to the three PCA components ($PC1$ – $PC3$), effectively representing how each principal component is composed of specific facial features. This graphical structure corresponds to the PCA results described in Section 4.3, and confirms that each component (e.g., $PC1$: large and horizontally wide eyes) reflects a unique combination of facial cues.

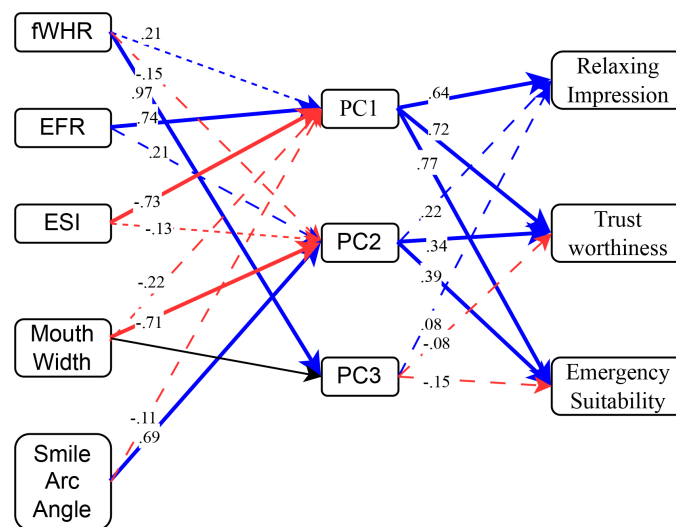


Figure 6: Three-layer path model showing relationships from geometric facial features to PCA components (left), and from PCA components to impression dimensions (right). Solid lines indicate significant relationships ($\beta \geq 0.30$), bold lines highlight strong predictors ($\beta \geq 0.50$), and dashed lines indicate weaker or non-significant paths.

The right portion of the figure shows the standardized regression coefficients from the PCA components to the three impression dimensions: RI , Tw , and ES . Strong beta values ($|\beta| \geq 0.50$) are represented with bold solid lines, moderate effects ($|\beta| \geq 0.30$) with thin solid lines, and weaker or non-significant effects with dashed lines.

This model illustrates a hierarchical structure: facial geometry first determines latent visual factors ($PC1$ – $PC3$), which subsequently influence

users' multidimensional impression judgments. It thereby visualizes the full pipeline of impression formation, from physical facial features to psychological evaluations.

DISCUSSION

Structure of Avatar-Based Impression Formation

This study supports H1: impressions of AI-generated avatars are structured along three interrelated dimensions—*Relaxing Impression (RI)*, *Trustworthiness (Tw)*, and *Emergency Suitability (ES)*. Latent visual components predicted impressions more effectively than individual features.

PC1 (“*Large and wide eyes*”, *high EFR & low ESI*) was the strongest predictor across all dimensions and both groups, reinforcing the role of eye-related cues as universal trust signals.

In contrast, *PC2* (“*Narrow smiling mouth*”) significantly influenced Chinese participants but had weaker or no effects among Japanese participants, especially for *RI*. *PC3* (“*Wide and short face*”) had minimal impact. The PCA model captured holistic impression patterns through a hierarchical path: facial features → components → impressions (Fig. 6).

Cultural Differences in Trust Evaluation Pathways

The results support the Cultural Moderation Hypothesis (H2), showing cultural differences in interpreting identical avatar features. Japanese participants gave higher ratings overall, particularly for *Trustworthiness* and *Emergency Suitability*, and responded more consistently to avatars with prominent eye-area configurations (*PC1*), a universal predictor across both groups. This may reflect a cultural emphasis on institutional trust and preparedness.

In contrast, Chinese participants showed a more differentiated evaluation process. Structural equation modeling revealed a sequential path from *Relaxing Impression* to *Emergency Suitability* to *Trustworthiness*, suggesting that emotional comfort plays a foundational role in building trust. Regression results further showed that *PC2* (“*Narrow smiling mouth*”) significantly predicted all impression dimensions for Chinese participants, but not for Japanese participants, indicating greater sensitivity to smile-related cues in the Chinese context.

Contrary to expectations, *PC3* (“*Wide and short face*”) and individual features like *ESI* and *Mouth Width* had limited predictive power in either group, reinforcing that meaningful cultural effects emerge more clearly through latent components (e.g., *PC1*, *PC2*) rather than isolated metrics.

In summary, these findings partially support H2: while both groups rely on eye-area cues (*PC1*), Chinese participants place more weight on subtle emotional signals in smiles, whereas Japanese participants evaluate avatars more uniformly across impression types.

Design Implications and Limitations

These findings highlight the importance of culturally adaptive avatar design. While features like “*large and horizontally wide eyes*” (PC1) can foster baseline trust across cultures, smile-related traits (e.g., *Mouth Width, Smile-Arc-Angle*) may require cultural tuning to avoid misinterpretation.

Static facial features alone, however, may be insufficient for building trust. Future avatar systems should incorporate dynamic cues such as gaze, blinking, and subtle expressions, along with personalization options that consider users’ cultural and demographic contexts.

Several limitations should be noted. The avatars used in this study were static and silent, reducing ecological validity. The dataset, based on commercially generated imagery, may also carry biases. Furthermore, the emergency communication scenario may have influenced participants’ impressions. Future studies should examine how impressions vary across more diverse formats, including casual, instructional, and commercial scenarios, to advance the design of trustworthy and adaptive avatar systems.

CONCLUSION

This study examined how geometric facial features of AI-generated avatars shape cross-cultural impressions across three dimensions: *Relaxing Impression, Trustworthiness, and Emergency Suitability*. Results supported a structured, three-dimensional model of trust perception, with eye-area configurations emerging as consistent predictors.

Cross-cultural analysis further validated our hypothesis on effect of cultural background, revealing that Chinese and Japanese participants interpreted identical facial cues through distinct cognitive pathways. While eye-area configurations functioned as a universal trust cue, smile-related features showed culturally specific effects.

By integrating facial feature mappings and principal component-based regression analysis, we proposed a layered impression model. This framework provides a foundation for designing avatars that are both culturally resonant and perceptually effective.

Future research should expand on dynamic, interactive avatars and explore real-world application contexts to enhance trust-sensitive digital communication systems.

ACKNOWLEDGMENT

This research was partially supported by the JST-Mirai Program (the Japan Science and Technology Agency); PI: Prof. Mihoko Niitsuma (2022–2024), and by the Joint Research Fund of the Institute of Science and Engineering, Chuo University; PI: Prof. Toshikazu Kato (2022–2024).

REFERENCES

- Bhattacharya, A., Qiu, L., Kumar, N.: Cross-cultural perceptions of AI-generated synthetic faces. *ACM Trans. Comput.-Hum. Interact.* 27(4), Article 25 (2020). <https://doi.org/10.1145/3380957>.

- Hofstede Insights: Compare countries. <https://www.hofstede-insights.com/product/compare-countries/> (Accessed 26 June 2025).
- Hu, R., Zhang, J., Sun, Y., Wang, C.: The role of baby schema in perceived trustworthiness of virtual faces in older adults. *Comput. Hum. Behav.* 123, 106869 (2021). <https://doi.org/10.1016/j.chb.2021.106869>.
- Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural Confusions Show that Facial Expressions Are Not Universal. *Current Biology*, 19(18), 1543–1548. <https://doi.org/10.1016/j.cub.2009.07.051>.
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98(2), 224–253.
- Mayer, R. C., Davis, J. H., Schoorman, F. D.: An integrative model of organizational trust. *Acad. Manage. Rev.* 20(3), 709–734 (1995). <https://doi.org/10.5465/amr.1995.9508080335>.
- Oosterhof, N. N., Todorov, A.: The functional basis of face evaluation. *Proc. Natl. Acad. Sci. USA* 105(32), 11087–11092 (2008). <https://doi.org/10.1073/pnas.0805664105>.
- Sobieszek, A., Siemiątkowski, M., Imbir, K. K.: Generative neural networks for experimental manipulation: Examining dominance-trustworthiness face impressions with data-efficient models. *Br. J. Psychol.* (2024). <https://doi.org/10.1111/bjop.12732>.
- Song, Y., Luximon, A., Luximon, Y.: The effect of facial features on facial anthropomorphic trustworthiness in social robots. *Appl. Ergon.* 94, 103043 (2021). <https://doi.org/10.1016/j.apergo.2021.103043>.
- Wienrich, C., Latoschik, M. E.: eXtended artificial intelligence: New prospects of human–AI interaction. *i-com* 20(3), 225–239 (2021). <https://doi.org/10.1515/icom-2021-0018>.
- Yuki, M., Maddux, W. W., & Masuda, T. (2007). Are the windows to the soul the same in the East and West? Cultural differences in using the eyes and mouth as cues to recognize emotions in Japan and the United States. *Journal of Experimental Social Psychology*, 43(2), 303–311.
- Zebrowitz, L. A., Montepare, J. M.: Social psychological face perception: Why appearance matters. *Soc. Pers. Psychol. Compass* 2(3), 1497–1517 (2008). <https://doi.org/10.1111/j.1751-9004.2008.00109.x>.