Testing the effect of depth on the perception of faces in an online study

Simon M. Hofmann[1,2], Abhay Koushik[1,3], Anthony Ciston[1], Felix Klotzsche[1,4], Vadim Nikulin[1], Arno Villringer[1,4], Michael Gaebler[1,4]

[1] *Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany* | [2] *Fraunhofer Institute Heinrich Hertz, Berlin, Germany*

[3] *Center de Research Interdisciplinaire (CRI), Université de Paris, France* | [4] *Humboldt-Universität zu Berlin, Berlin School of Mind and Brain, Germany*

simon.hofmann@cbs.mpg.de HofmannZimon

MAX PLANCK INSTITUTE FOR HUMAN COGNITIVE AND BRAIN SCIENCES

Introduction

Faces are socially relevant stimuli, distinguished by spatial arrangements of their features [1]. The perceptual system orders these features in a cognitive **"face space", where distance represents face similarity** [2].

Previously, this "face space" has been mostly investigated with 2D faces. We plan an online study to **investigate the effect of 2D vs 3D representations on face perception** using a **similarity judgment task**.

We present here the results from an advanced pilot experiment.

Hypotheses

To investigate the effect of depth on face perception and bridge the gap towards naturalistic stimuli, we hypothesize that:

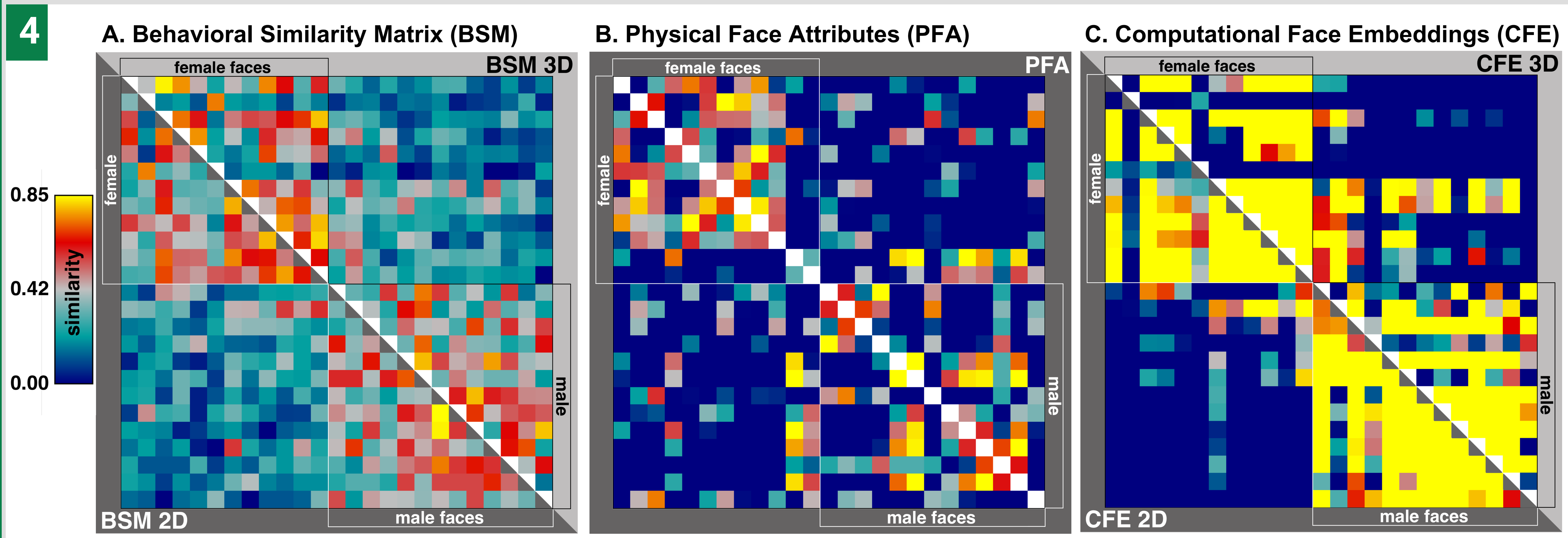
(1) Facial dimensions that span the face space differ between representations (2D, 3D).

(2) Volumetric features are more relevant in 3D

(3) Computational models fitted to behavioral data reflect these volumetric properties.

Results

4



Human similarity judgements in both viewing conditions (*lower left*: 2D; *upper right triangle*: 3D). Cells represent aggregated pairwise similarity judgements. First 12 columns/rows represent ♀ faces; last 13, represent ♂ faces.

Cosine similarities of physical attributes in face pairs. Initial 44 attributes (e.g., face width) were subject to PCA. Here, the first 5 PCs were used explaining 89.5% variance.

Similarities of VICE embeddings
In **2D** 4 & in **3D** 6 (out of initial 20) dimensions remained relevant.
Gender is a driving component; for further **interpretation** of single VICE dimensions see **Fig 5**.

RSA (BSM, 2D & 3D)
R=0.75, p<0.001, 43.54% of variance ($1-R^2$) in one viewing condition remains unexplained by the other.

RSA (BSM-PFA)

- **2D:** R = 0.21, p < 0.001
- **3D:** R = 0.14, p < 0.02

VICE prediction accuracy

- **2D**: 60.19%
- **3D**: 64.10%

Discussion

Differences in representations (2D, 3D)

- PFA & CFE explain parts of human judgements
- 2D-based PFA explains 2D BSM better than 3D
- BSMs are more heterogeneous, gender effect partially stronger in CFE & PFA

Outlook: main experiment & further analysis

- Using **100 faces** ($n_♀=n_♂=50$)
- Online study: **1000 participants** to sample **161,700 combinations** of face triplets
- Using non-linear models with explainable A.I. methods
- Extracting volumetric features, i.e., "3D PFA"

Implications

- 3D effect neglected in previous 2D research of faces
- Encourage more naturalistic 3D designs
- Pipeline usable for further research (e.g., ethnical groups, clinical populations, psychophysiological studies)

References

[1] Eng et al. (*Vis. Res.*, 2017)
[2] Jozwik et al. (*PNAS*, 2022)
[3] Ma et al. (*Behav. Res. Methods*, 2015)
[4] Feng et al. (*ACM Trans. Graph.*, 2021)
[5] Hebart et al (*Nat. Hum. Behav.*, 2020)
[6] Brookes al. (*Behav. Res. Methods*, 2020)
[7] Kriegeskorte & Kievit (*TiCS*, 2013)
[8] Muttenthaler et al. (*arXiv*, 2022)

Acknowledgement

We would like to thank Martin Hebart & Laura Stoinski for their valuable inputs on the similarity judgement task.

Methods

Stimulus Preparation

- Images from standardized 2D Chicago-Face-Database (**CFD**) [3]
- Random sample of neutral faces ($n_♀=12$, $n_♂=13$)
- **Deep learning-based pipeline (DECA)** for **3D-face-reconstruction** [4] (**Fig 1**)

Sampling the cognitive face space

Similarity Judgement Task

- **Triplet odd-one-out task** [5] (**Fig 3**)
- Implemented in Unity-based *UXF 2.0* [6]
- Computing pair-wise **behavioral similarity matrices (BSM)** (**Fig 4A**)
- Between-subject-design:
- **2D**: 3D-reconstructed but static (n=14, 8 ♀)
- **3D**: with rotating faces (n=16, 8 ♀; **Fig 2**)
- 180 trials in 3 blocks per participant (**Fig 3**)

1 3D reconstruction using DECA

2 Dynamic 3D representation

Representational similarity analysis (RSA) [7]
Similarity matrices are analyzed using Spearman's *rho* (R) to quantify their differences.

Explaining the cognitive face space

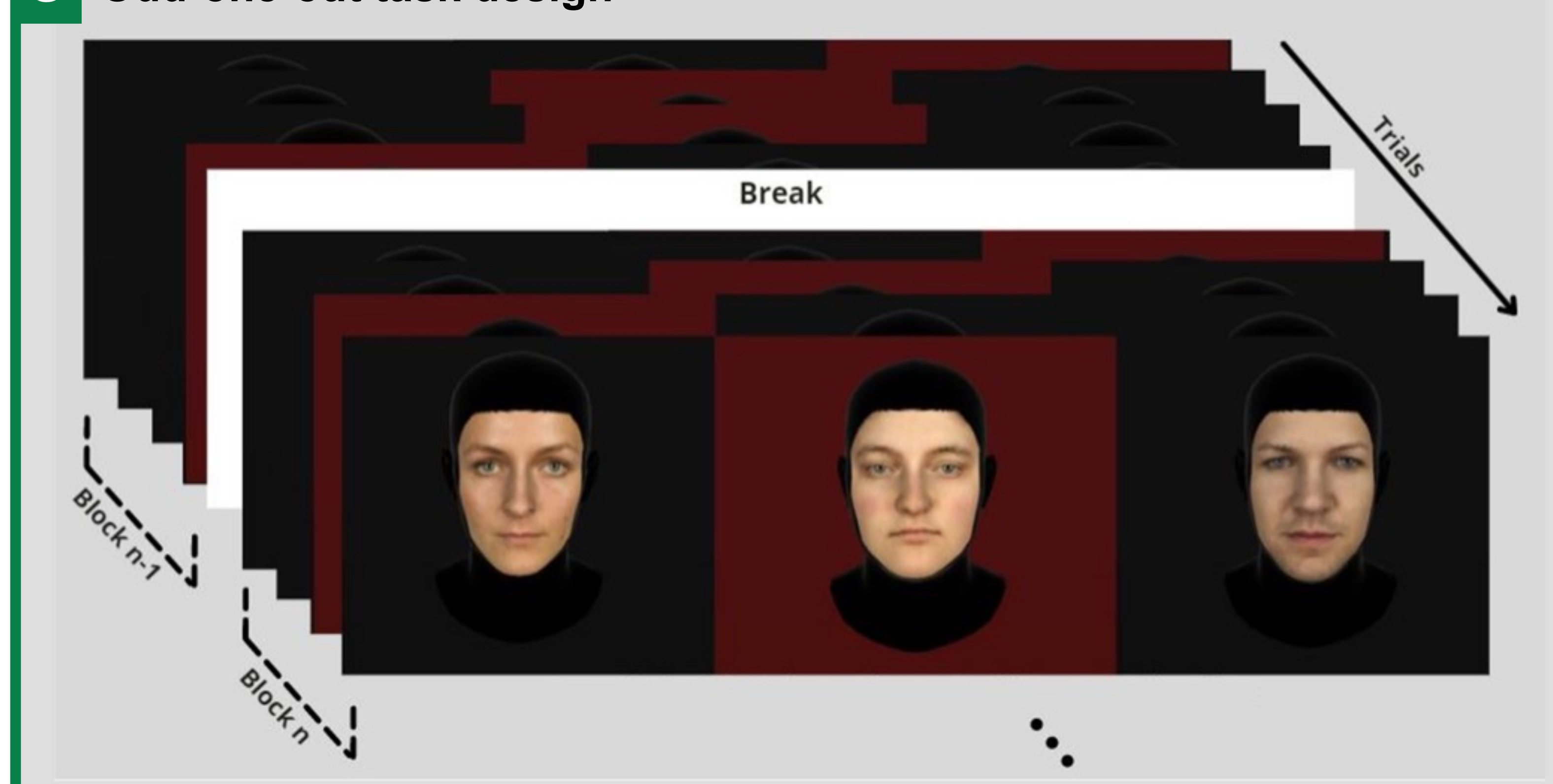
Similarity of physical face attributes (PFA)

- Initial 44 PFA in CFD [3] (e.g., face width)
- Subject to a principal component analysis
- Extraction of 5 most informative PC's
- Compute cosine similarity of reduced PFA for each face pair (**Fig 4B**)

Computational face embeddings (CFE)

- **Variational Interpretable Concept Embeddings** (VICE) models [8]
- VICE extracts the most relevant & interpretable dimensions (**Fig 4C**)
- Trained on 90% of trials (n_{train} = 2070)
- Predict human judgements in 2D & 3D, respectively on 10% of trials (n_{test} = 230)

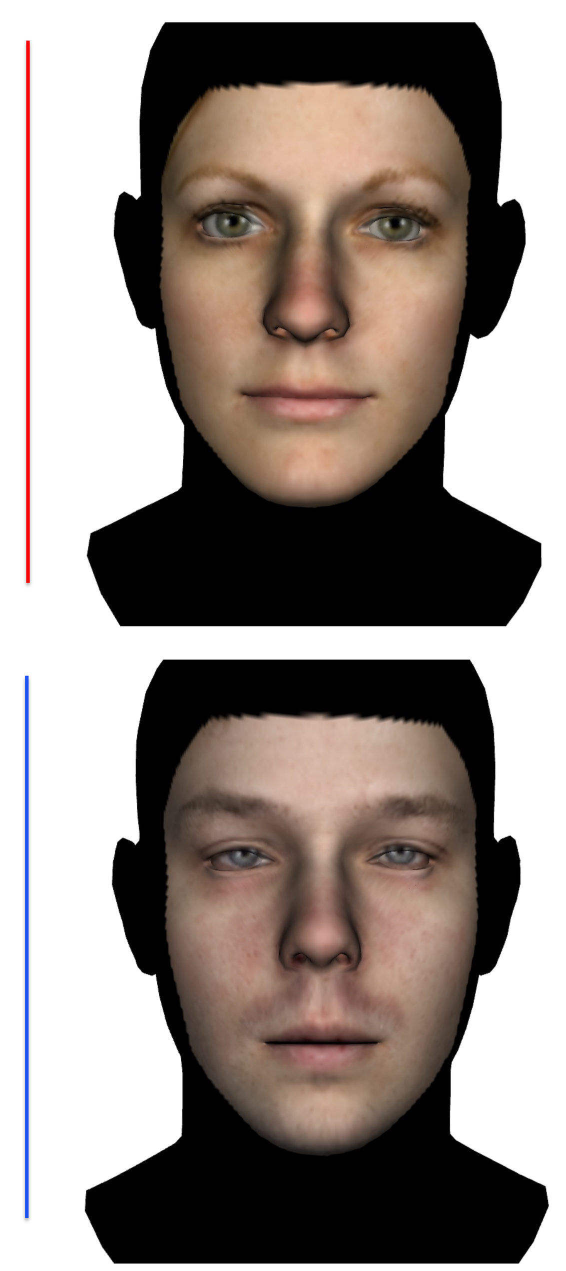
3 Odd-one-out task design



5 Interpretation of VICE embeddings

Dimension 1 of VICE model 3D correlated with 8 eye-related PFAs (e.g., eye size; all 0.76 > R > 0.64). *Upper* face (index 3 in **Fig 4**) has strongest weight in this dimension, *bottom* face the weakest (index 16).

Other dimensions correlated with the shape of chins, cheeks & noses, and the luminance of faces, for both VICE models (2D, 3D).



CCN Conference, San Francisco, 2022