Introduction

Humans can rapidly detect faces in their visual field, despite their variety. A cognitive template for face detection must therefore be able to accommodate face variability. In this study, we investigate whether a statistical average of previously encountered faces could form a generic detection template. The visual system uses statistical summaries, such as averages, to represent sets of objects, including faces (de Fockert & Wolfenstein, 2009). Averages capture commonalities across faces, both within a single identity (Burton et al., 2005) and across multiple identities (Sutherland et al., 2013). Yet an effective template would also need to be stable (see Jenkins & Burton, 2011). Here we manipulate the number and variability of faces contributing to an average to determine when stability is reached.

Methods

- **Participants:** Total N = 112 (75 females). Experiment 1 N = 32, Experiment 2, N = 40, Experiment 3 N = 40.
- **Face stimuli:** 420 ambient faces from an online generator. Twelve categories: gender, age and race.
- **Averages:** 288 average faces from 120 source faces per condition. 12 set sizes (2, 4, 6, 8, ..., 18, 20, 30, 40).
- **Category number:** Six conditions: two single categories, one 2-category, 4-category, 6-category and 12-category.
- **Perceptual stability:** Participants judged the similarity of average pairs via a ‘same’/’different’ 2AFC task. Face pairs were constructed from an equal number of randomly selected, non-overlapping face identities.
- **Image stability:** Principal Components Analysis (PCA) used to estimate the variability of face images at each set size, by calculating the standard deviation across eigenvectors for randomly generated sets of averages.

Results

**Perceptual stability**

- **Experiment 1:** A 6 (Category Number) x 12 (Set Size) ANOVA. Effects of Category Number, F(5,155) = 70.91, p < .001, Set Size, F(11,341) = 294.18, p < .001, and an interaction, F(55,1705) = 5.67, p < .001.
- Matches increased with set size number. But few differences after SS10, stable from SS20 (dashed lines = n.s.).
- Matches decreased with category number, and higher category numbers (e.g., 4C, 6C, 12C) stabilised later.
- Differences between category numbers persisted even at the highest set sizes.
- **Experiment 2:** Direct replication of Experiment 1 with online participants (due to Covid-19).
- **Experiment 3:** Category conditions were presented in separate counterbalanced blocks, rather than intermixed, as high category number trials may appear more dissimilar in the context of low category number trials.
- Experiment 3 found effects of Category Number, F(5,195) = 10.27, p < .001, Set Size, F(11,429) = 250.89, p < .001, and an interaction, F(55,2145) = 5.10, p < .001. But found a smaller effect size for category number (ŋ²p = .21) compared to experiment 1 (ŋ²p = .70), and convergence across conditions from SS20.

**Image stability**

- Effects of Category Number, F(5,155) = 99.24, p < .001, Set Size, F(11,341) = 3637.23, p < .001, and an interaction, F(55,1705) = 3.11, p < .001.
- Declines in variance with set size number, then levels off.
- Variance increased with category number, but differences only consistent for 1CM and 12C.

Conclusions

- Cross-identity face averages reach stability surprisingly quickly, after being constructed from just 20 to 30 distinct identities.
- This is evident from both subjects’ similarity judgements and the PCA.
- Stability is reached regardless of whether averages are constructed from single or multiple face categories.
- A cognitive face detection template based on averaging could form rapidly, after exposure to just a small number of faces.