Supplementary Methods

***3D Face Database*.** The face database comprised 197 females, 158 males, 233 Western Caucasian, 122 East Asian, age between 16 and 86, SD = 15.06, scanned in-house with a Di4D face capture system, at a high resolution in shape (4,735 3D vertex coordinates) and texture (800*600 RGB pixels, see Supplementary Figure 1A). All 3D models were in full color with hair removed, posing with a neutral facial expression.

***Fine-tuning Beta_2 Coefficients*.** In a self-adaptive procedure, we initialized Beta_2 amplification with equally spaced values between 0 and 50, with 10 unit increments. We then narrowed the amplification range to participant's responses until convergence, keeping the same total number of stimuli (i.e. 6 faces) per trial. Supplementary Figure 2B2 illustrates the adaptive procedure.

The experiment comprised one session per familiar face, with familiar face order randomized across participants. Each session started with the screen presentation of the front view of one familiar face target to instruct participants as to the target of the session. On each trial, 6 faces initially amplified between 0 and 50 appeared on the screen, randomly positioned in a 2 by 3 array against a black background. We instructed participants to choose the face that best resembled the familiar identity by pressing one of six response buttons. The 6 faces remained on the screen until response, immediately followed by the next trial. We repeated the trial 5 times, with the same 6 faces in different random array positions, to determine the next amplification range. We narrowed the amplification range every 5 trials by finding the minimum and maximum values that bound the participant's 5 choices. With this new range, we produced 6 new faces by evenly sampling the amplification values and again tested the participant over 5 new trials. We iteratively repeated sequences of 5 testing trials, updates of the amplification range, until it stabilized—i.e. remained constant over three blocks of 5 trials. We used the median of the final amplification range as value to generate the fine-tuned Beta_2 coefficients that we call *mental representation* in our analyses (see Supplementary Figure 2B2).

Non-negative Matrix Factorization (NNMF). We applied NNMF to the full 4735 by 56 (i.e. vertex-by-model) binary faithful representation matrix to identify the main face shape features that represent faithful memory representations of identities across all participants. NNMF factorizes the multi-dimensional (and positive) data into non-negative additive components by minimizing the reconstruction error. We performed NNMF with an alternating-least square algorithm and repeated the factorization using 1 to 20 components. We determined N, the optimal number of components, by repeating 1,000 times NNMF for each number of components, randomizing initial conditions and recalculating the mean squared residual reconstruction error. The boxplots in Supplementary Figure 4A illustrate the curve of residual errors. Its 2nd derivative shows that residual errors flatten when N equals 8

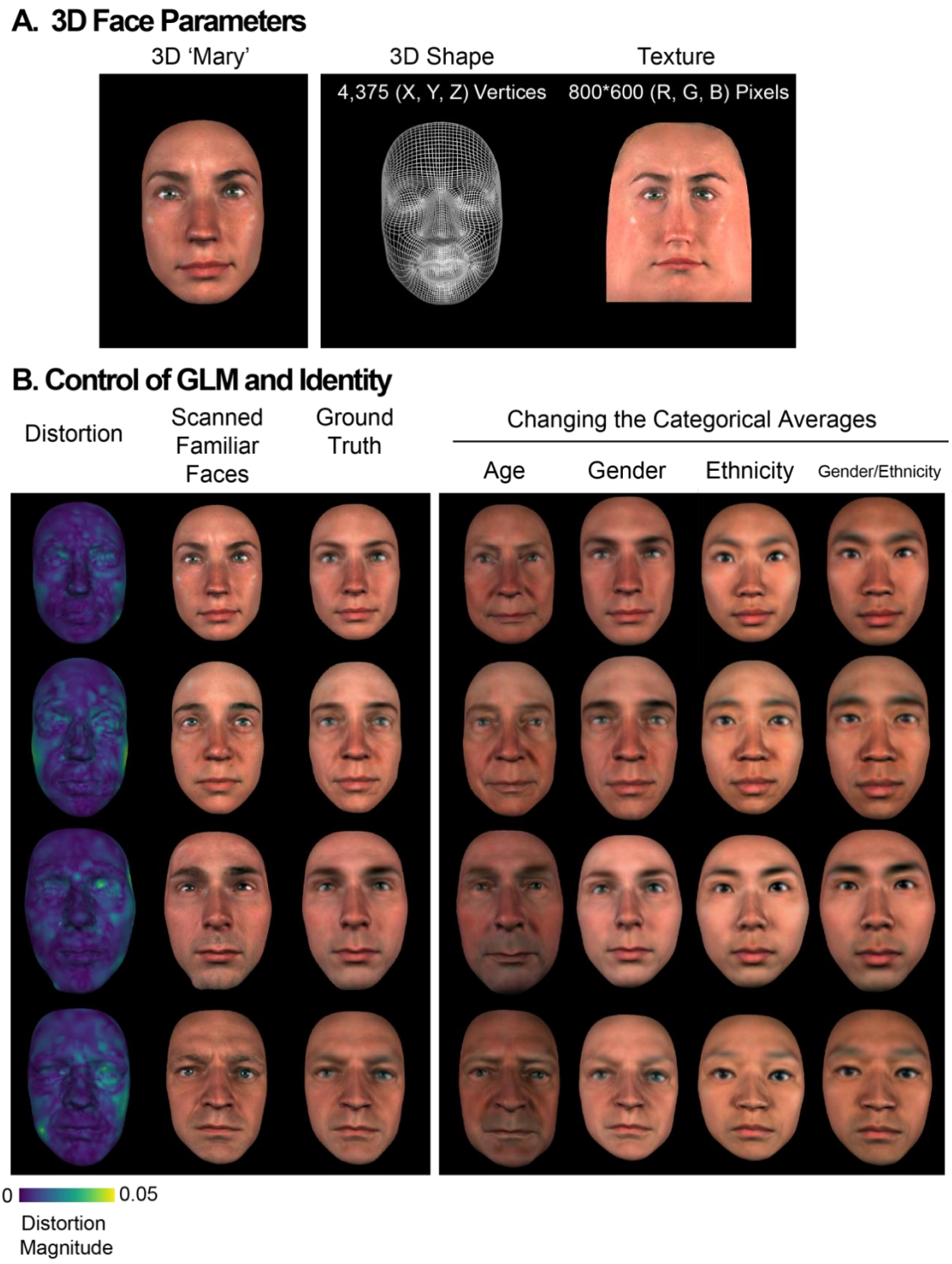
components. Supplementary Figure 4B shows the face reconstructions based on these reduced set of 8 components against the faithful representations of the original faces (cf. Figure 2B), demonstrating the reliability of our NNMF additive reconstructions.

Linear Mixed Effect Model of Face Type by Amplification Interaction. We fitted a linear mixed effects model (i.e. fitlme, Matlab 2016b) to the data by collapsing across 4 identities, using Wilkinson's formulae:

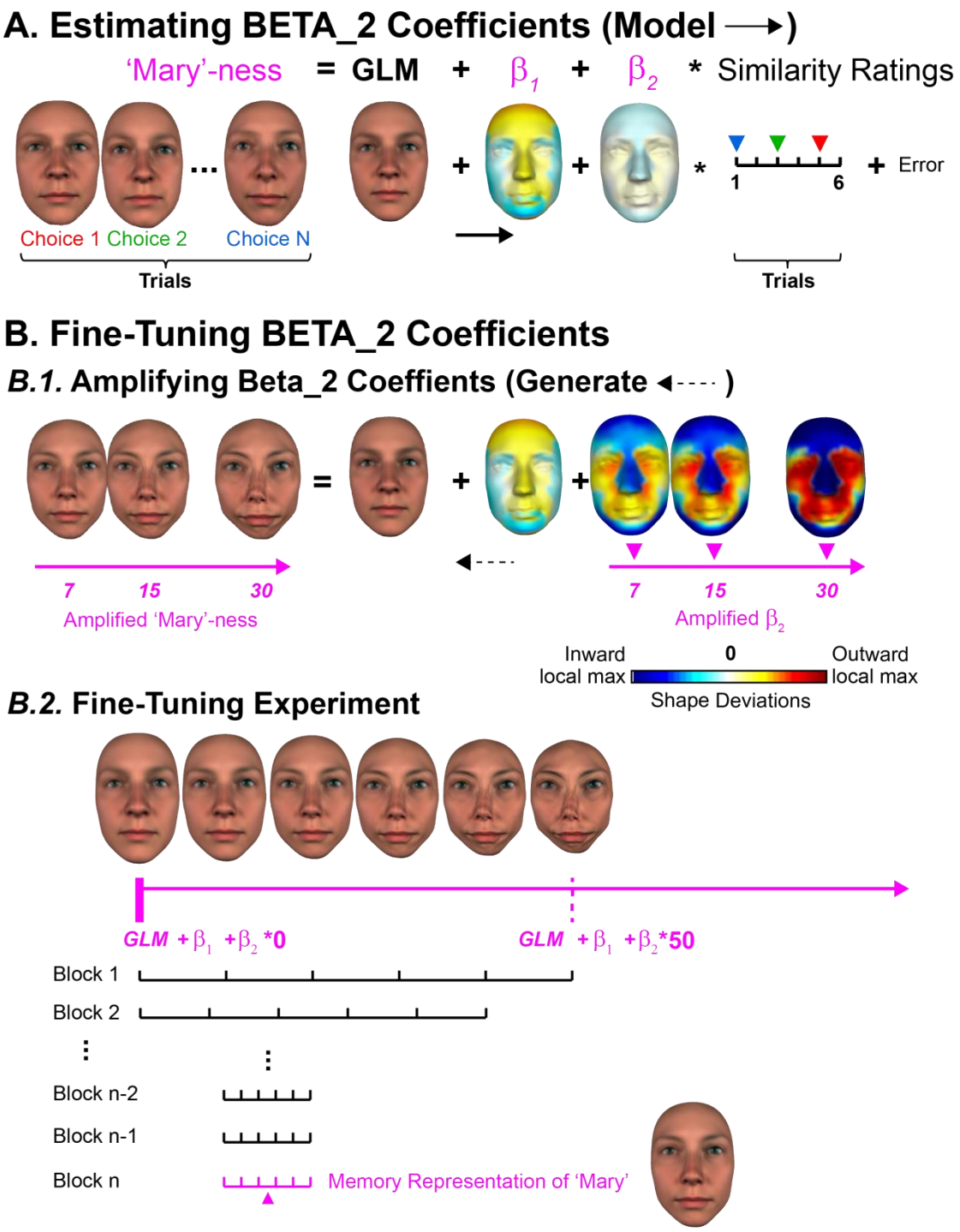
$$Performance \sim 1 + Face\ Type + Amplification + Face\ Type * Amplification + (Face\ Type + Amplification + Face\ Type * Amplification - 1|Subject)$$

The model had fixed factors of Face Type (i.e. diagnostic vs. nondiagnostic), Feature Amplification (i.e. 0.33, 0.67, 1, 1.33, 1.67) and the interaction between Face Type and Amplification as explanatory variables, and participants' response variability as random factor. We tested the specified fixed effect factors using ANOVA in Matlab 2016b, and reported the full statistics in Supplementary Table 7.

Supplementary Figures and Tables

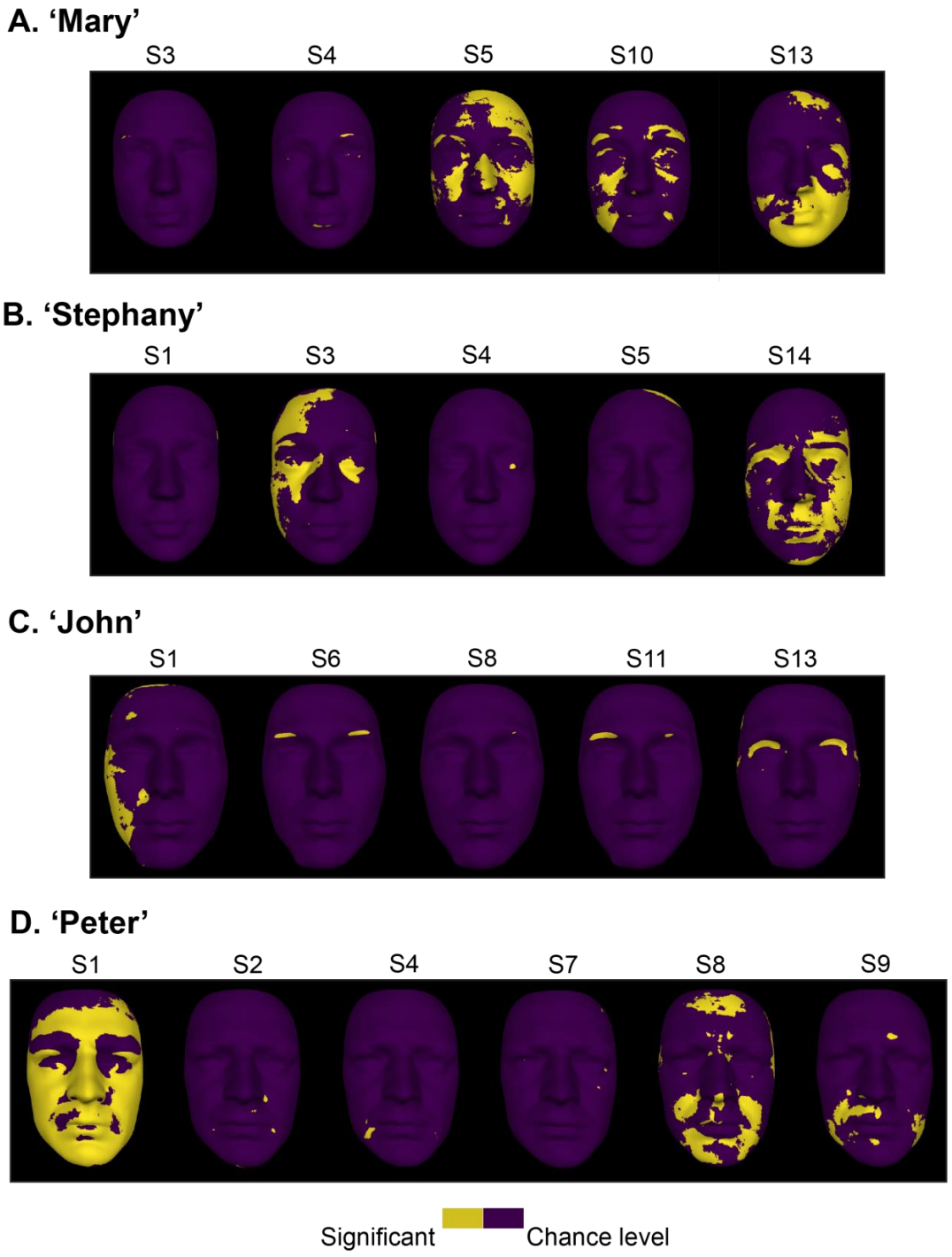


Supplementary Figure 1. (A) 3D Face parameters. We parametrized the shape of a face with the 3D coordinates of 4,735 vertices and its texture with 800*600 RGB 2D pixels. (B) GLM control of the categorical averages and identity components. *Distortion.* Distortion quantifies, vertex per vertex the quality of the 3D GLM fit of the scanned familiar faces. *Changing the categorical averages.* In each column, the GLM controls the factors of sex, ethnicity and age using local averages, while the identity residuals are kept constant.



62

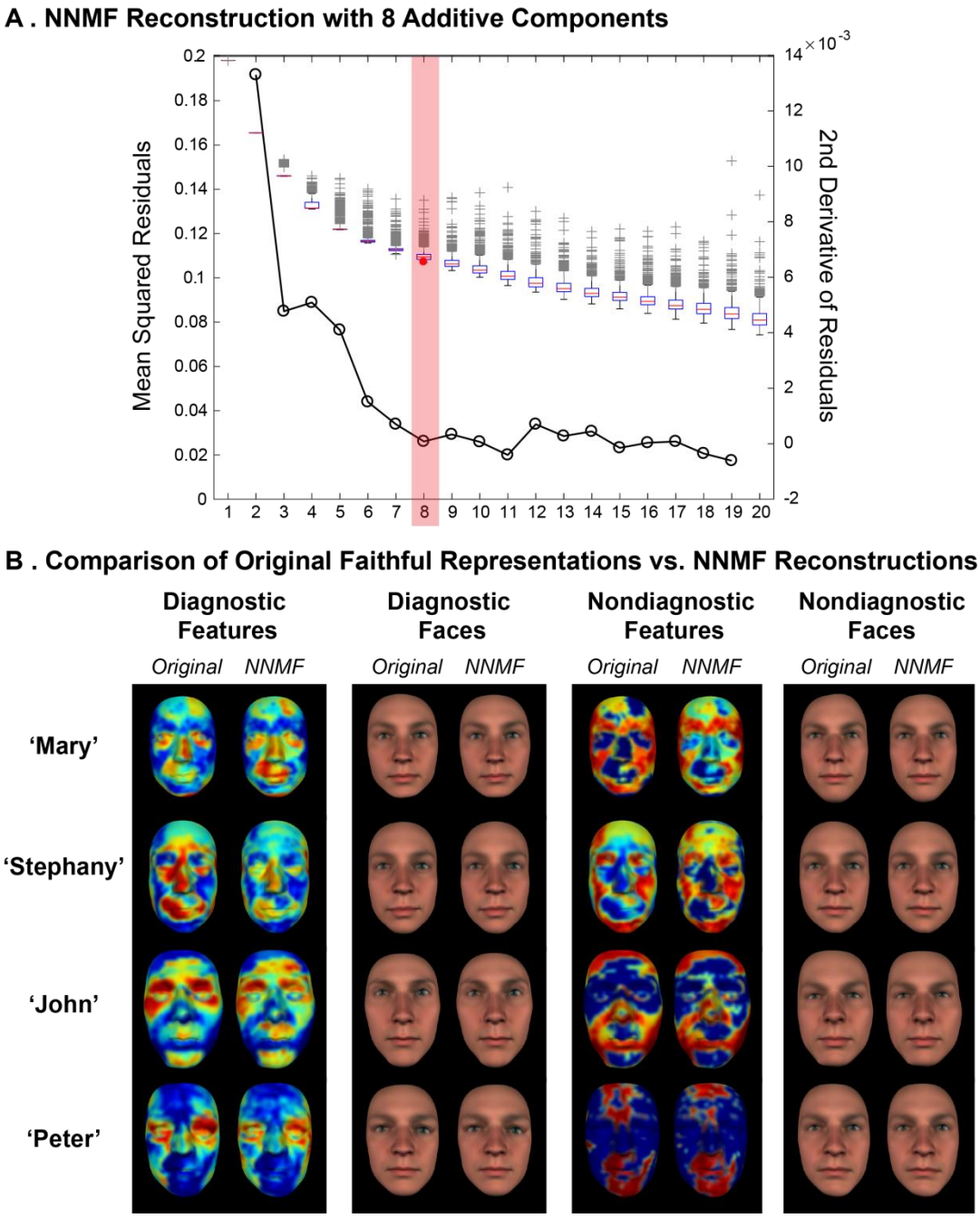
63 **Supplementary Figure 2. Reverse-correlating the visual information contents of familiar**
64 **face representations.** (A) Estimating Beta_2 Coefficients. We linearly regressed the 3D
65 vertices of shape (separately for the X, Y and Z coordinates, texture not illustrated) with
66 similarity judgments of the selected random identities (illustrated here for 'Mary'). For each
67 vertex, Beta_2 coefficients are color-coded according to their magnitude. Yellow-to-red
68 indicates an outward change from the categorical average; turquoise-to-blue indicates an
69 inward change from the categorical average. (B) Fine-tuning Beta_2 Coefficients. *(B.1)*
70 *Amplifying Beta_2 coefficients.* Illustration of the amplification of Beta_2 coefficients. *(B.2)*
71 *Illustration of the fine-tuning experiment.*



72

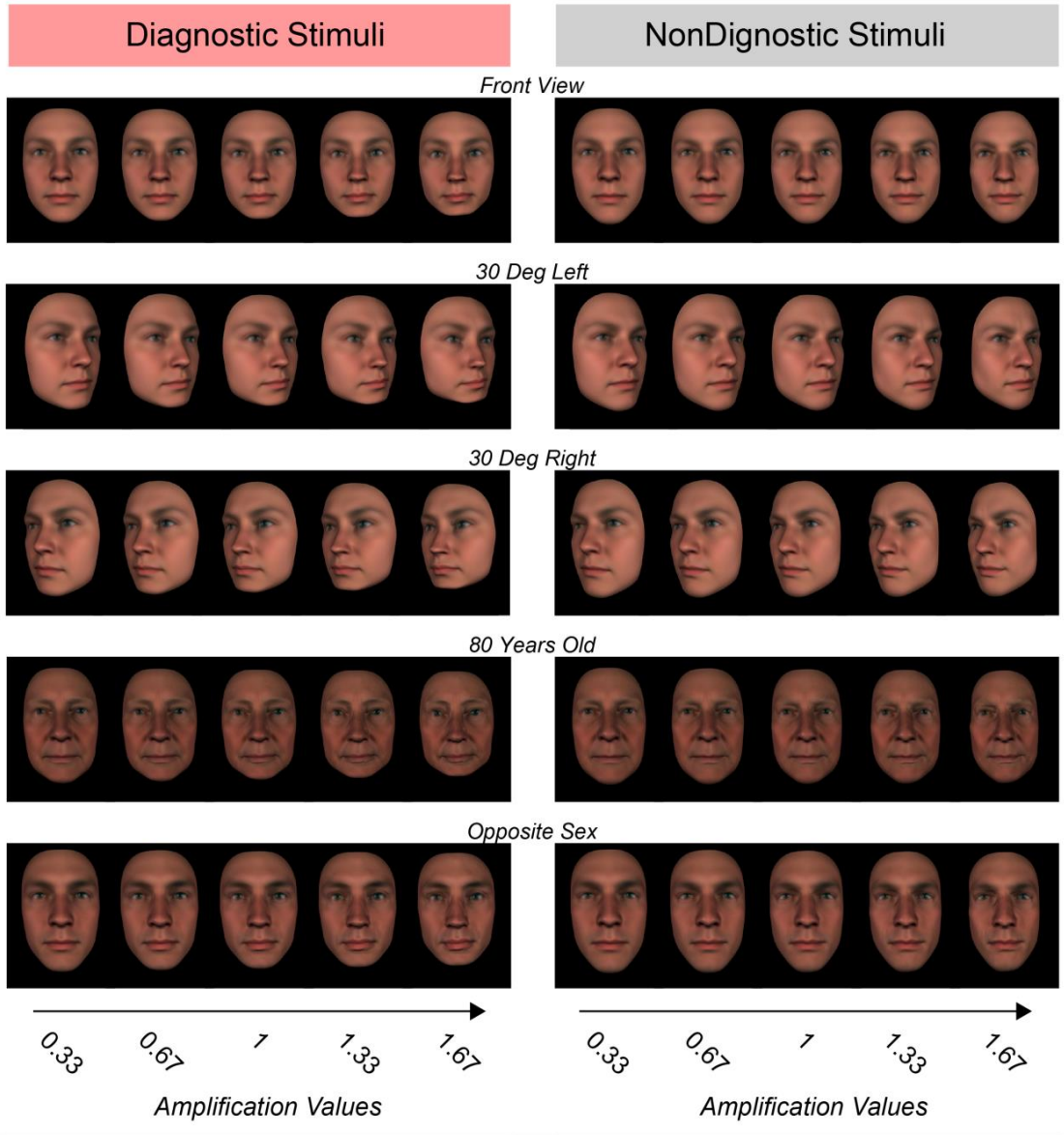
73 **Supplementary Figure 3. Beta_2 coefficients of texture.** Yellow colored overlays on each
74 familiar face illustrate the significant Beta_2 coefficients for RGB texture pixels in each
75 participant (labelled S1-S14). Dark purple pixels represent non-significant RGB coefficients.

76

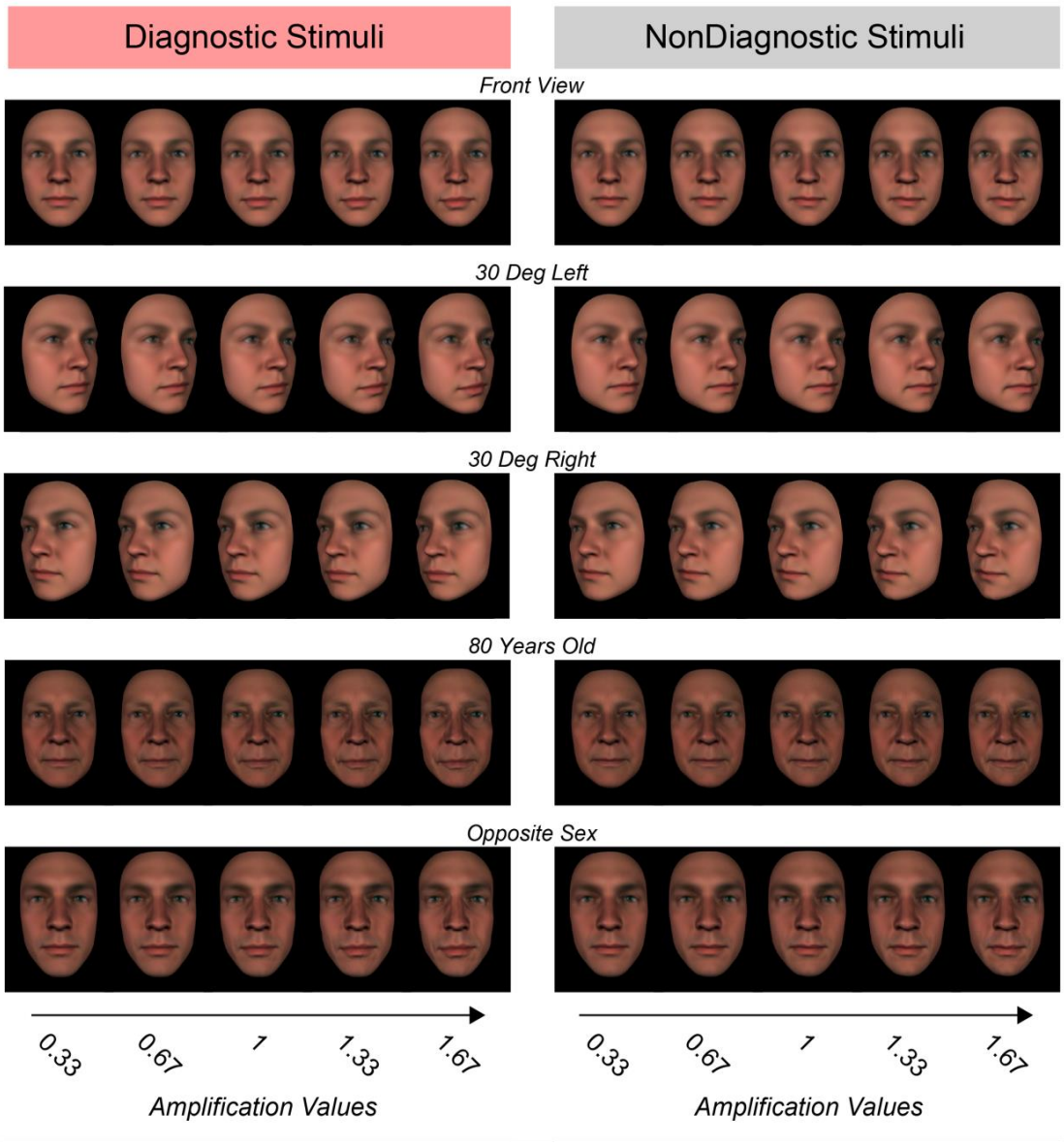


Supplementary Figure 4. Accuracy of reduced (i.e. 8 NNMF components) multivariate representations of faithful face representations. A. NNMF representation with 8 additive components. We performed NNMF across all identity models and participants to derive a reliable low-dimensional multivariate representation of their main shape features. We found that reconstruction error plateaued with 8 NNMF components (highlighted in red). B. Comparison of original faithful representations vs. NNMF reconstructions. We compared the reduced NNMF representations of each identity with their original faithful representations and, as can

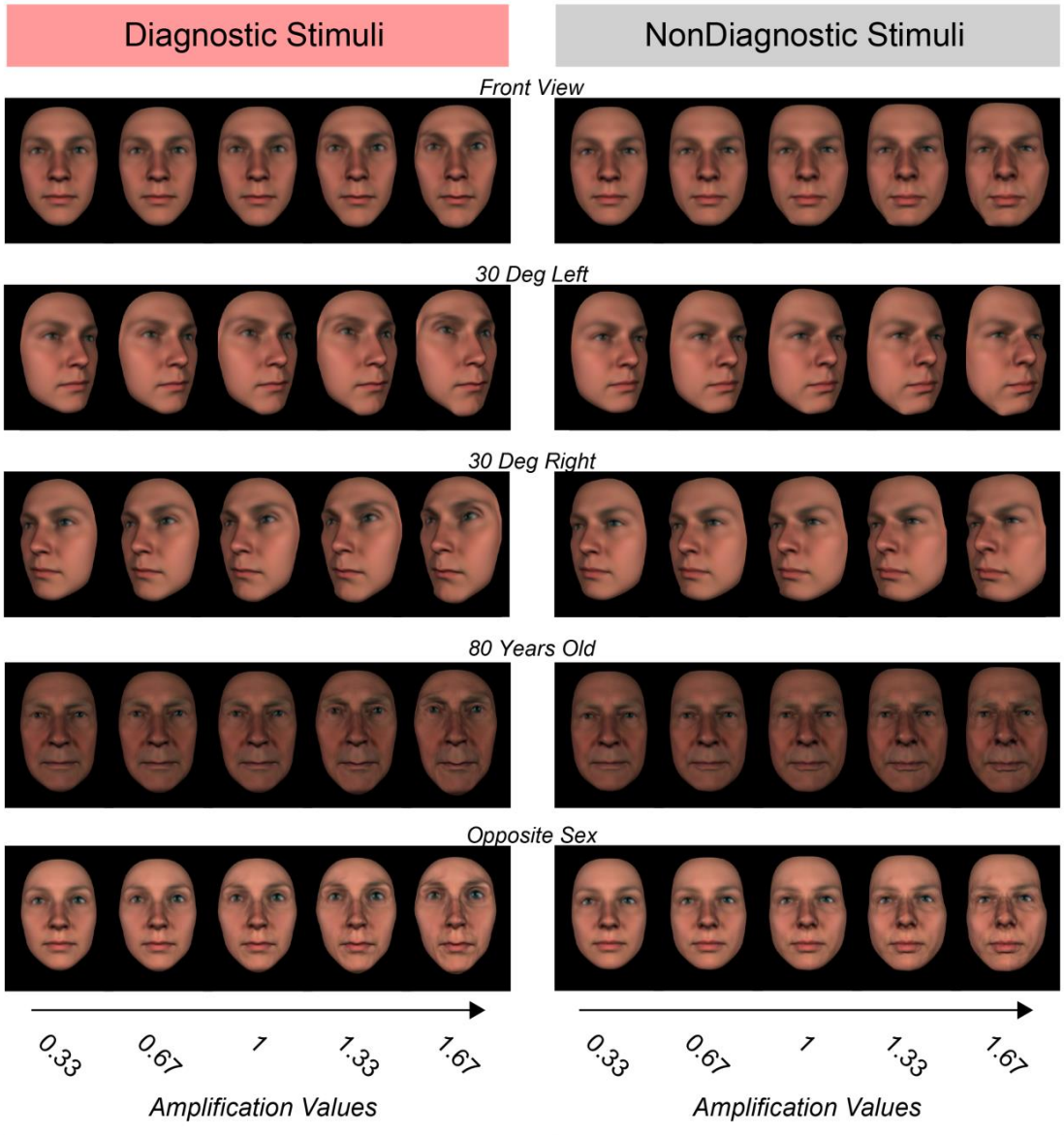
be seen both from the reconstructed diagnostic features and faces, the reconstructions with 8 components were reliable.



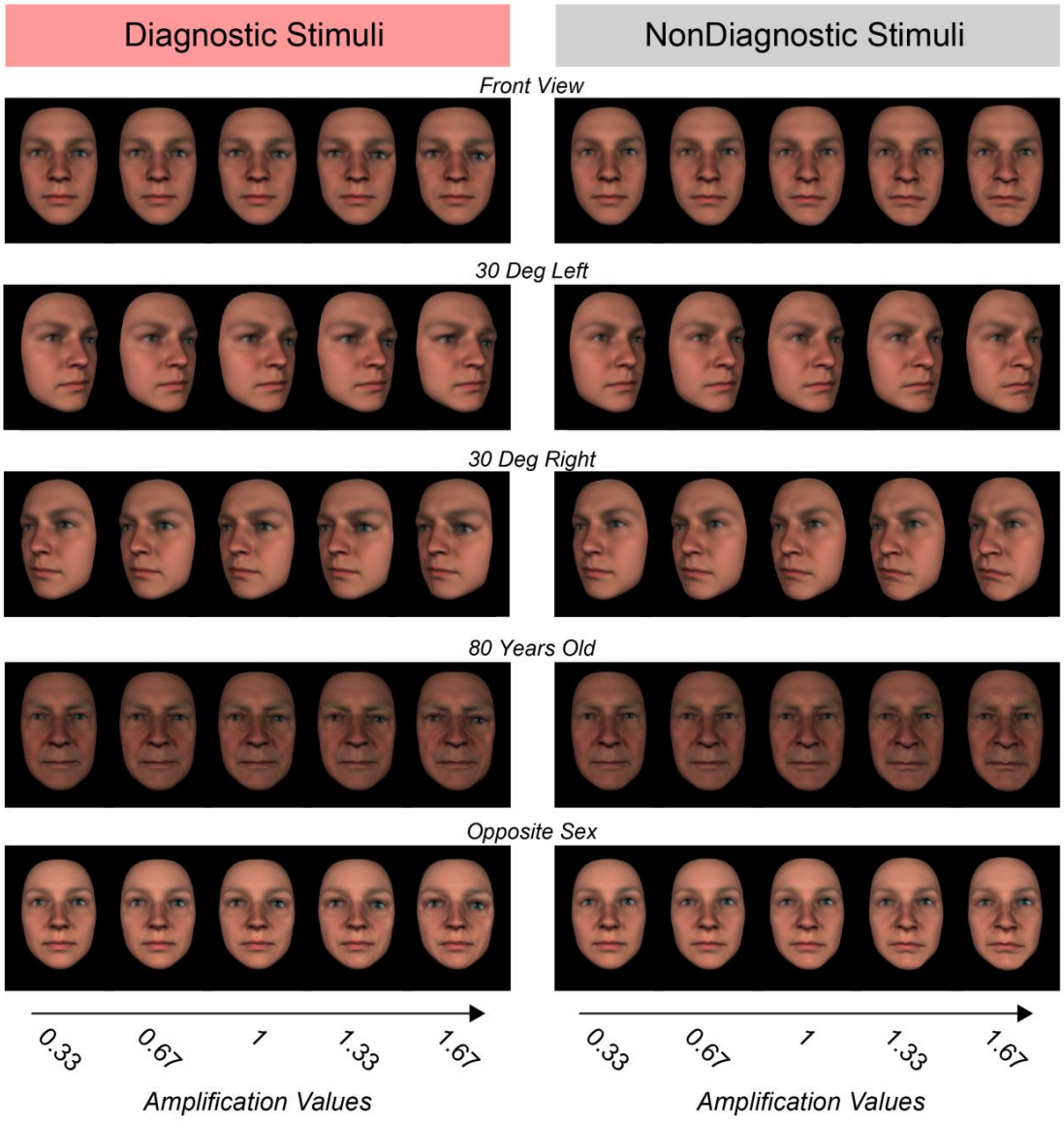
Supplementary Figure 5. Diagnostic (left) and nondiagnostic (right) faces of 'Mary'. Each row presents the main conditions of stimulus synthesis (i.e. 3 viewpoints, age and sex). Each column presents a level of diagnostic (vs. nondiagnostic) component amplification in the face.



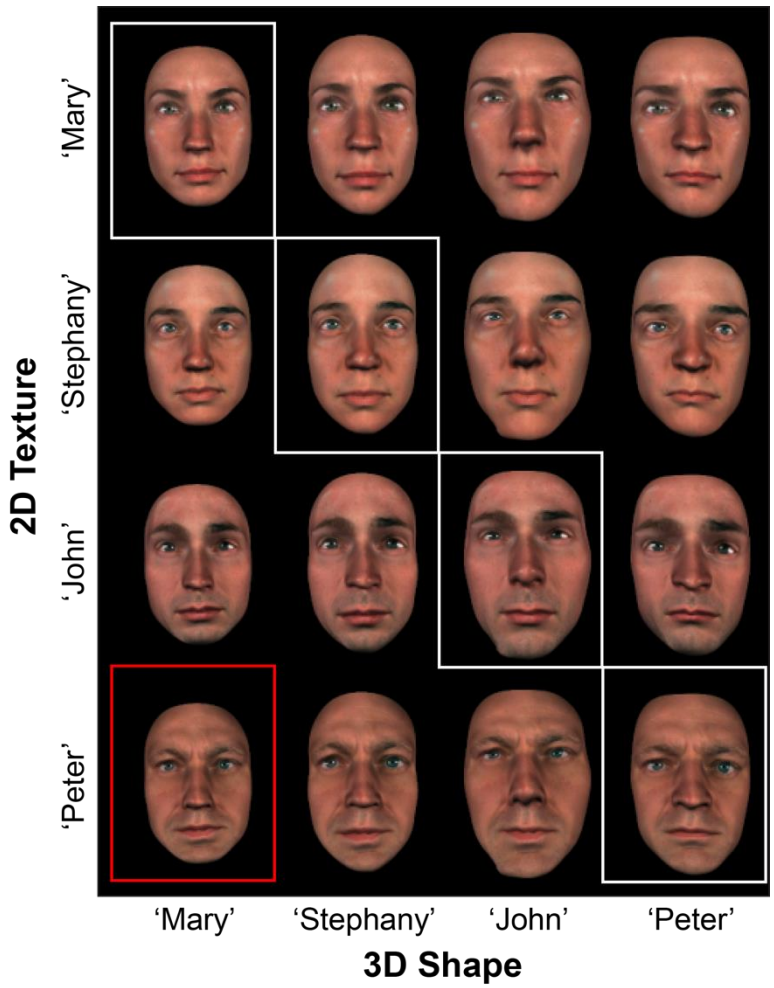
Supplementary Figure 6. Diagnostic (left) and nondiagnostic (right) faces of 'Stephany'. Same caption as in Supplementary Figure 5.



Supplementary Figure 7. Diagnostic (left) and nondiagnostic (right) faces of 'John'. Same caption as in Supplementary Figure 5.



Supplementary Figure 8. Diagnostic (left) and nondiagnostic (right) faces of 'Peter'. Same caption as in Supplementary Figure 5.



Supplementary Figure 9. Different 3D shapes (X axis) and different 2D textures (Y axis) are combined to synthesize the 4 original target faces on the diagonal (white framed). Off diagonal faces combine the X-axis shape of identity A with the Y-axis texture of identity B. For example, the red framed identity combines the shape of 'Mary' with the texture of 'Peter.'

Supplementary Table 1. Identity familiarity ratings of 14 participants in the reverse correlation experiment.

Participants	Mary	Stephany	John	Peter
1	6	4	3	7
2	7	8	9	6
3	7	9	7	3
4	8	8	7	8
5	3	4	4	4
6	3	3	4	5
7	5	6	9	4
8	5	6	9	5
9	8	7	7	4
10	8	6	8	10
11	7	7	8	7
12	7	6	6	7
13	5	3	9	7
14	9	9	9	9
Mean	6.29	6.14	7.07	6.14
SD	1.86	2.03	2.09	2.07

Note: Ratings are from 1 (not familiar at all) to 9 (highly familiar).

Supplementary Table 2. Identity familiarity ratings of 12 validators in the generalization experiment.

Validators	Mary	Stephany	John	Peter
1	9	7	6	6
2	7	5	6	4
3	9	9	9	9
4	7	8	9	6
5	7	6	5	4
6	6	4	5	4
7	8	6	9	8
8	9	6	9	9
9	8	5	7	6
10	6	9	3	4
11	9	9	7	9
12	9	5	9	7
Mean	7.83	6.58	7	6.33
SD	1.19	1.78	2.04	2.06

Note: Ratings are from 1 (not familiar at all) to 9 (highly familiar).

Supplementary Table 3. Linear mixed-effects model for ‘Mary’.

Estimation of Linear-mixed Effect Models						Statistics of Fixed Factors			
Fixed Factors		Estimated Slope	SE	95% Confidence Intervals		*F value*	*DF1*	*DF2*	*p value (two-sided)*
				Lower	Upper				
Intercept		0.146	0.021	0.104	0.188	46.23	1	564	< 0.001
Face Type	(Diag vs. Nondiag)	0.297	0.017	0.264	0.33	315.49	1	12.763	< 0.001
Amplification		0.321	0.025	0.272	0.37	165.99	1	46.502	< 0.001
Task Type A	(30 Deg Left vs. Front)	-0.038	0.021	-0.079	0.002	7.8653	4	20,386	< 0.001
Task Type B	(30 Deg Right vs. Front)	0.105	0.025	0.055	0.155				
Task Type C	(80 Years Old vs. Front)	-0.098	0.022	-0.142	-0.055				
Task Type D	(Opposite Sex vs. Front)	-0.107	0.037	-0.18	-0.034				

Note: Diag = diagnostic; Nondiag = nondiagnostic; Deg = degree.

Supplementary Table 4. Linear mixed-effects model for ‘Stephany’.

Estimation of Linear-mixed Effect Models						Statistics of Fixed Factors			
Fixed Factors		Estimated Slope	SE	95% Confidence Intervals		*F value*	*DF1*	*DF2*	*p value (two-sided)*
				Lower	Upper				
Intercept		0.343	0.023	0.299	0.388	231.72	1	552	< 0.001
Face Type	(Diag vs. Nondiag)	0.058	0.012	0.035	0.081	25.068	1	20.624	< 0.001
Amplification		0.156	0.052	0.054	0.258	9.047	1	15.714	0.008
Task Type A	(30 Deg Left vs. Front)	0.139	0.051	0.039	0.239	9.815	4	14.734	< 0.001
Task Type B	(30 Deg Right vs. Front)	-0.069	0.042	-0.152	0.014				
Task Type C	(80 Years Old vs. Front)	-0.083	0.029	-0.139	-0.027				
Task Type D	(Opposite Sex vs. Front)	0.067	0.046	-0.024	0.159				

Note: Diag = diagnostic; Nondiag = nondiagnostic; Deg = degree.

Supplementary Table 5. Linear mixed-effects model for ‘John’.

Estimation of Linear-mixed Effect Models						Statistics of Fixed Factors			
Fixed Factors		Estimated Slope	SE	95% Confidence Intervals		*F value*	*DF1*	*DF2*	*p value (two-sided)*
				Lower	Upper				
Intercept		0.398	0.025	0.35	0.447	261.9	1	540	< 0.001
Face Type	(Diag vs. Nondiag)	0.143	0.031	0.083	0.204	21.639	1	12.004	< 0.001
Amplification		0.162	0.059	0.045	0.278	7.385	1	15.297	0.016
Task Type A	(30 Deg Left vs. Front)	0.032	0.027	-0.022	0.085	1.591	4	14.468	0.23
Task Type B	(30 Deg Right vs. Front)	-0.01	0.022	-0.053	0.033				
Task Type C	(80 Years Old vs. Front)	-0.062	0.028	-0.117	-0.006				
Task Type D	(Opposite Sex vs. Front)	0.007	0.033	-0.058	0.071				

Note: Diag = diagnostic; Nondiag = nondiagnostic; Deg = degree.

Supplementary Table 6. Linear mixed-effects model for ‘Peter’.

Estimation of Linear-mixed Effect Models						Statistics of Fixed Factors			
Fixed Factors		Estimated Slope	SE	95% Confidence Intervals		*F value*	*DF1*	*DF2*	*p value (two-sided)*
				Lower	Upper				
Intercept		0.251	0.023	0.206	0.295	122.73	1	564	< 0.001
Face Type	(Diag vs. Nondiag)	0.095	0.04	0.017	0.173	5.76	1	12.007	0.034
Amplification		0.394	0.038	0.32	0.469	107.72	1	20.592	0.008
Task Type A	(30 Deg Left vs. Front)	-0.028	0.027	-0.082	0.025				
Task Type B	(30 Deg Right vs. Front)	-0.012	0.024	-0.058	0.035	2.696	4	20.831	0.059
Task Type C	(80 Years Old vs. Front)	0.047	0.029	-0.011	0.104				
Task Type D	(Opposite Sex vs. Front)	-0.047	0.037	-0.119	0.026				

Note: Diag = diagnostic; Nondiag = nondiagnostic; Deg = degree.

Supplementary Table 7. Linear mixed-effects model with Face Type by Amplification Interaction.

Estimation of Linear-mixed Effect Models					Statistics of Fixed Factors			
Fixed Factors	Estimated Slope	SE	95% Confidence Intervals		F value	DF1	DF2	p value (two-sided)
			Lower	Upper				
Intercept	0.285	0.015	0.256	0.314	317.31	1	2376	< 0.001
Face Type (Diag vs. Nondiag)	0.056	0.018	0.021	0.091	9.721	1	20.154	0.005
Amplification	0.258	0.036	0.188	0.328	52.341	1	15.301	< 0.001
Face Type * Amplification	0.092	0.024	0.046	0.138	15.318	1	13	0.002

Note: Diag = diagnostic; Nondiag = nondiagnostic; Deg = degree.