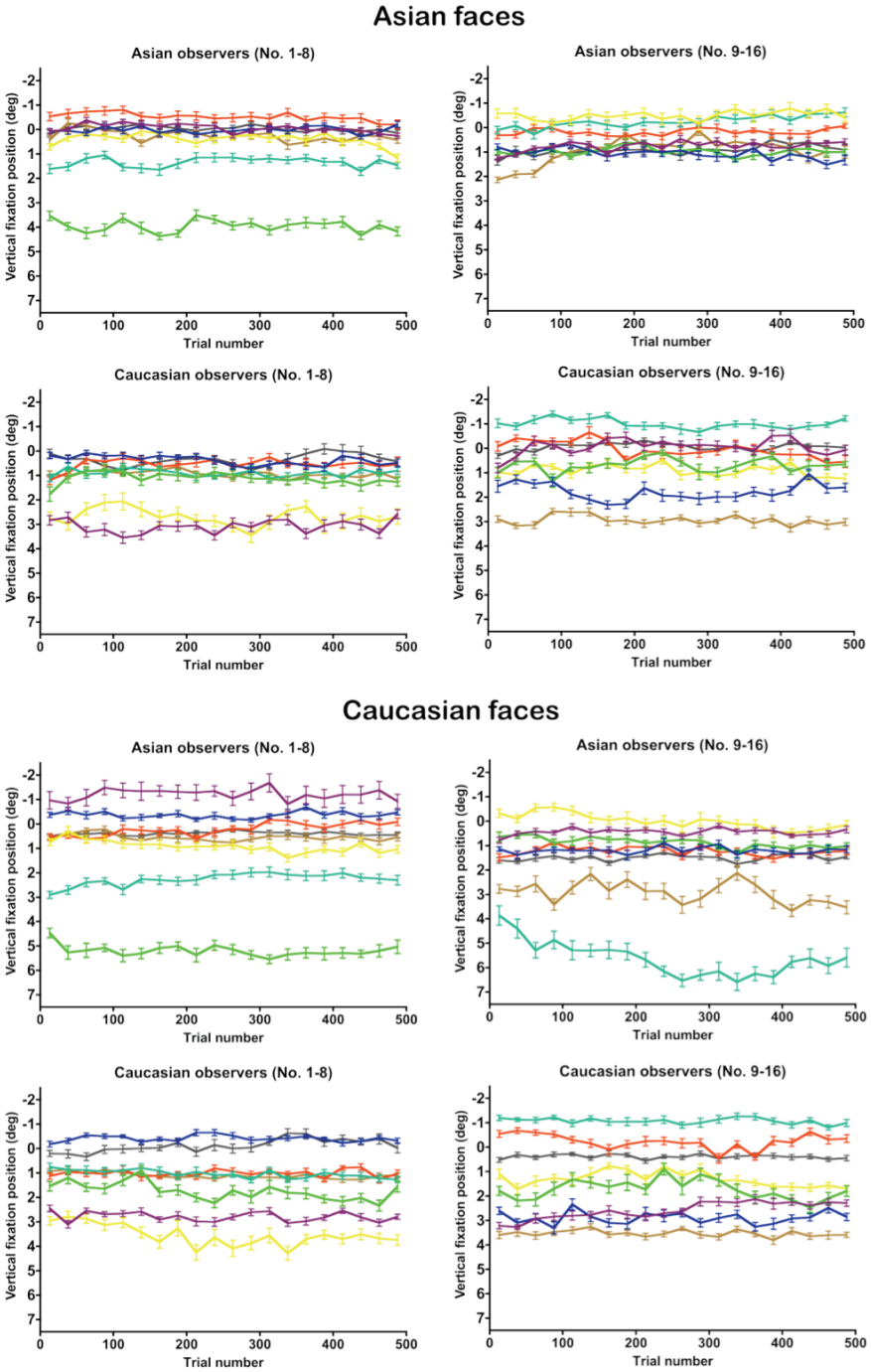
Or, Peterson, and Eckstein: Initial eye movements during face identification are optimal and similar across cultures

Supplementary materials

Experiment 1: Free eye movements (350 ms presentation)

(a)



(b)

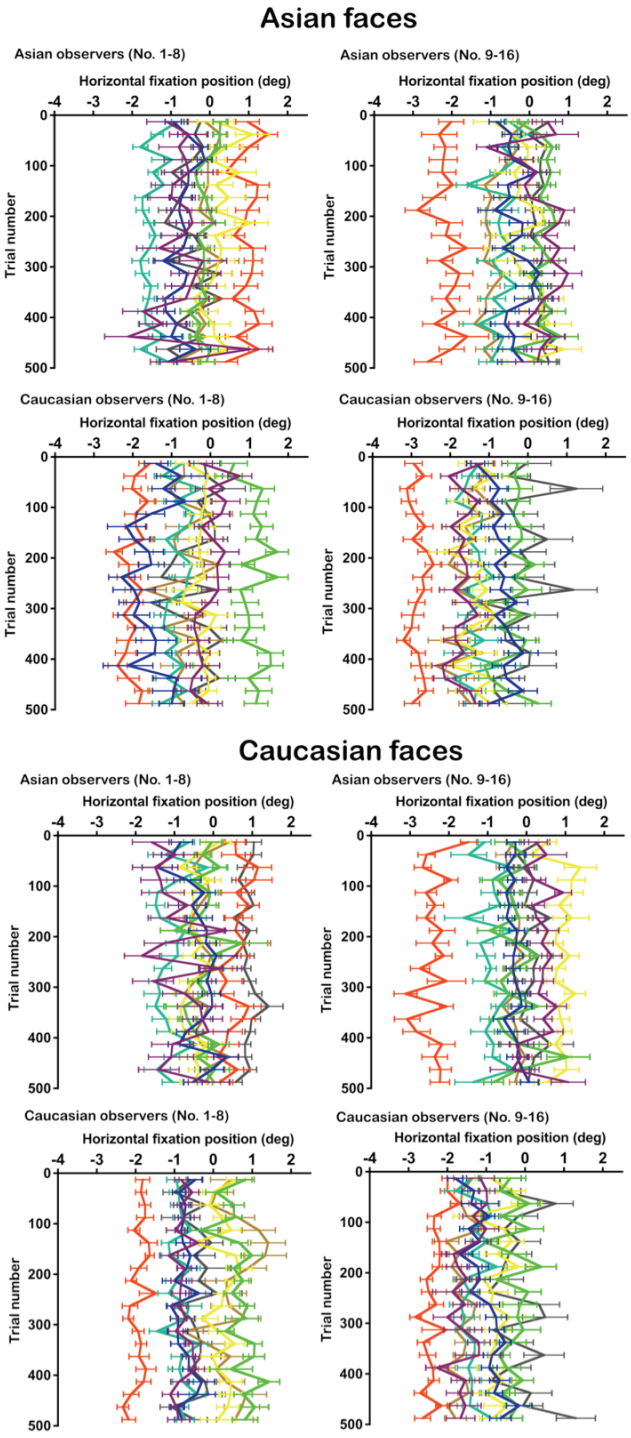


Figure S1 Individual observers' first fixation positions, separately plotted along (a) the vertical, and (b) the horizontal, dimensions, throughout Experiment 1 (Free eye movements: 350 ms presentations; 500 trials for each condition). Each point is an average fixation position across 25 consecutive trials (i.e., trials 1–25, 26–50, 51–75, and so on). The error bar denotes ±1 *SEM*.

Vertical dimension:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Trial group (T)	2.07	1	.16
Face race (F)	7.30	1	.01 *
Main effects (between-subjects)			
Observer culture (O)	0.17	1	.69
Two-way interactions			
T × O	0.46	1	.50
T × F	1.70	1	.20
O × F	0.67	1	.42
Three-way interaction			
T × O × F	1.32	1	.26
Residual		30	

Horizontal dimension:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Trial group (T)	0.81	1	.38
Face race (F)	1.29	1	.27
Main effects (between-subjects)			
Observer culture (O)	3.22	1	.08
Two-way interactions			
T × O	0.88	1	.36
T × F	1.77	1	.19
O × F	0.13	1	.72
Three-way interaction			
T × O × F	0.65	1	.43
Residual		30	

Table S1 Mixed-model repeated measures ANOVA results on potential variations of the first fixation positions across trials in Experiment 1 (Free eye movements: 350 ms presentations), analyzed separately for the vertical and the horizontal dimensions. The significant *p*-value (i.e., $p < .05$) is marked with an asterisk (*).

Experiment 2: Free eye movements (1500 ms presentation)

Vertical dimension:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Trial group (T)	3.41	1	.08
Face race (F)	6.77	1	.01 *
Main effects (between-subjects)			
Observer culture (O)	1.58	1	.22
Two-way interactions			
T × O	0.97	1	.33
T × F	0.82	1	.37
O × F	0.57	1	.46
Three-way interaction			
T × O × F	1.33	1	.26
Residual		30	

Horizontal dimension:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Trial group (T)	0.58	1	.45
Face race (F)	0.20	1	.66
Main effects (between-subjects)			
Observer culture (O)	4.45	1	.04 *
Two-way interactions			
T × O	1.17	1	.29
T × F	0.96	1	.34
O × F	0.20	1	.66
Three-way interaction			
T × O × F	0.58	1	.45
Residual		30	

Table S2 Mixed-model repeated measures ANOVA results on potential variations of the first fixation positions across trials in Experiment 2 (Free eye movements: 1500 ms presentations), analyzed separately for the vertical and the horizontal dimensions. Significant *p*-values (i.e., $p < .05$) are marked with asterisks (*).

The first fixation: Bootstrapping results

As in Experiment 1, the means and standard deviations of the fixation distributions were compared across conditions using the block bootstrapping method (Table S3). For fixation means in the vertical dimension, no significant differences were found across cultures. There were significant differences across face races for Asian observers, where fixations on Asian faces were slightly higher than on Caucasian faces. The same trend was observed for Caucasian observers but the differences failed to reach significance. In the horizontal dimension, the cultural differences just reached significance when Caucasian faces were displayed, where Caucasian observers directed their first eye movements slightly more leftward than Asian observers. On the contrary, the cultural differences with Asian faces barely failed to reach significance. There were no significant differences across face races. Nevertheless, the bootstrapping results were largely consistent with those from the ANOVA. In addition, the fixation standard deviations did not vary significantly across conditions.

Difference of fixation means

Vertical dimension	95% CI (deg)
Cultural difference: Asian – Caucasian observers	
Stimuli: Asian faces	(–1.43, 0.28)
Stimuli: Caucasian faces	(–1.24, 0.51)
Stimulus difference: Asian – Caucasian faces	
Observers: Asian	(–0.80, –0.21) *
Observers: Caucasian	(–0.71, 0.17)

Horizontal dimension	95% CI (deg)
Cultural difference: Asian – Caucasian observers	
Stimuli: Asian faces	(–0.03, 0.96)
Stimuli: Caucasian faces	(0.01, 1.06) *
Stimulus difference: Asian – Caucasian faces	
Observers: Asian	(–0.36, 0.15)
Observers: Caucasian	(–0.21, 0.22)

Difference of fixation standard deviations

Vertical dimension	95% CI (deg)
Cultural difference: Asian – Caucasian observers	
Stimuli: Asian faces	(–0.60, 0.58)
Stimuli: Caucasian faces	(–0.42, 0.57)
Stimulus difference: Asian – Caucasian faces	
Observers: Asian	(–0.28, 0.09)
Observers: Caucasian	(–0.32, 0.32)

Horizontal dimension	95% CI (deg)
Cultural difference: Asian – Caucasian observers	
Stimuli: Asian faces	(–0.67, 0.18)
Stimuli: Caucasian faces	(–0.45, 0.16)
Stimulus difference: Asian – Caucasian faces	
Observers: Asian	(–0.17, 0.25)
Observers: Caucasian	(–0.21, 0.44)

Table S3 The 95% confidence intervals (CI) of the differences of bootstrapped means and standard deviations of the distributions of the first fixation positions from Experiment 2 (Free eye movements, 1500 ms presentations). Significant differences are marked with asterisks (*).

Comparison between Experiment 1 and 2 results (the first fixation)

Vertical dimension:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Presentation time (T)	4.56	1	.04 *
Face race (F)	8.31	1	.007 *
Main effects (between-subjects)			
Observer culture (O)	0.71	1	.41
Two-way interactions			
T × O	3.39	1	.08
T × F	1.99	1	.17
O × F	0.74	1	.40
Three-way interaction			
T × O × F	0.21	1	.65
Residual		30	

Horizontal dimension:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Presentation time (T)	0.58	1	.45
Face race (F)	0.81	1	.38
Main effects (between-subjects)			
Observer culture (O)	4.02	1	.054
Two-way interactions			
T × O	0.004	1	.95
T × F	0.78	1	.39
O × F	0.001	1	.97
Three-way interaction			
T × O × F	0.84	1	.37
Residual		30	

Table S4 Mixed-model repeated measures ANOVA results on potential variations of the first fixation positions across Experiments 1 and 2 (350 ms vs. 1500 ms presentation times), analyzed separately for the vertical and the horizontal dimensions. Significant *p*-values (i.e., $p < .05$) are marked with asterisks (*).

The second fixation: Comparing fixation frequencies between left and right eyes

Observer culture	Face race	*t*	*df*	*P*-value
Asian	Asian	1.50	15	.16
	Caucasian	1.03	15	.32
Caucasian	Asian	0.70	15	.49
	Caucasian	0.85	15	.41

Table S5 Paired-samples t-test (two-tailed) results comparing fixation frequencies of the left-eye and the right-eye regions for each condition (Experiment 2: the second fixation). No significant difference was found for each comparison.

Subsequent fixations

Observers made on average 3.65 ± 0.57 (*SD*) fixations during the 1500 ms face presentations. We grouped the third and subsequent fixations and examined their distributions (Fig. S2a and Table S6) using similar methods as for earlier fixations. Similar to the second fixations, the third and subsequent fixations were scattered mainly on the eyes and nose regions, with a preference to the eyes when viewing Asian faces but a preference to the nose when viewing Caucasian faces, regardless of observer cultures. There were no significant differences between Asian and Caucasian observers regarding the fixation distributions. This indicates little change in fixation patterns after the second fixations.

All fixations

Finally, the distributions of all fixations were analyzed (Fig. S2b and Table S7). Again, observers from both cultures looked chiefly at the eyes and nose regions, but with different preferences to the eyes and nose for Asian and Caucasian faces respectively. There were also no significant differences in fixation distributions across observer cultures. Thus, Experiment 2 showed that Asian and Caucasian observers did not differ in saccades to faces within the 1500 ms face presentation, though they both employed different fixation strategies to faces of different races.

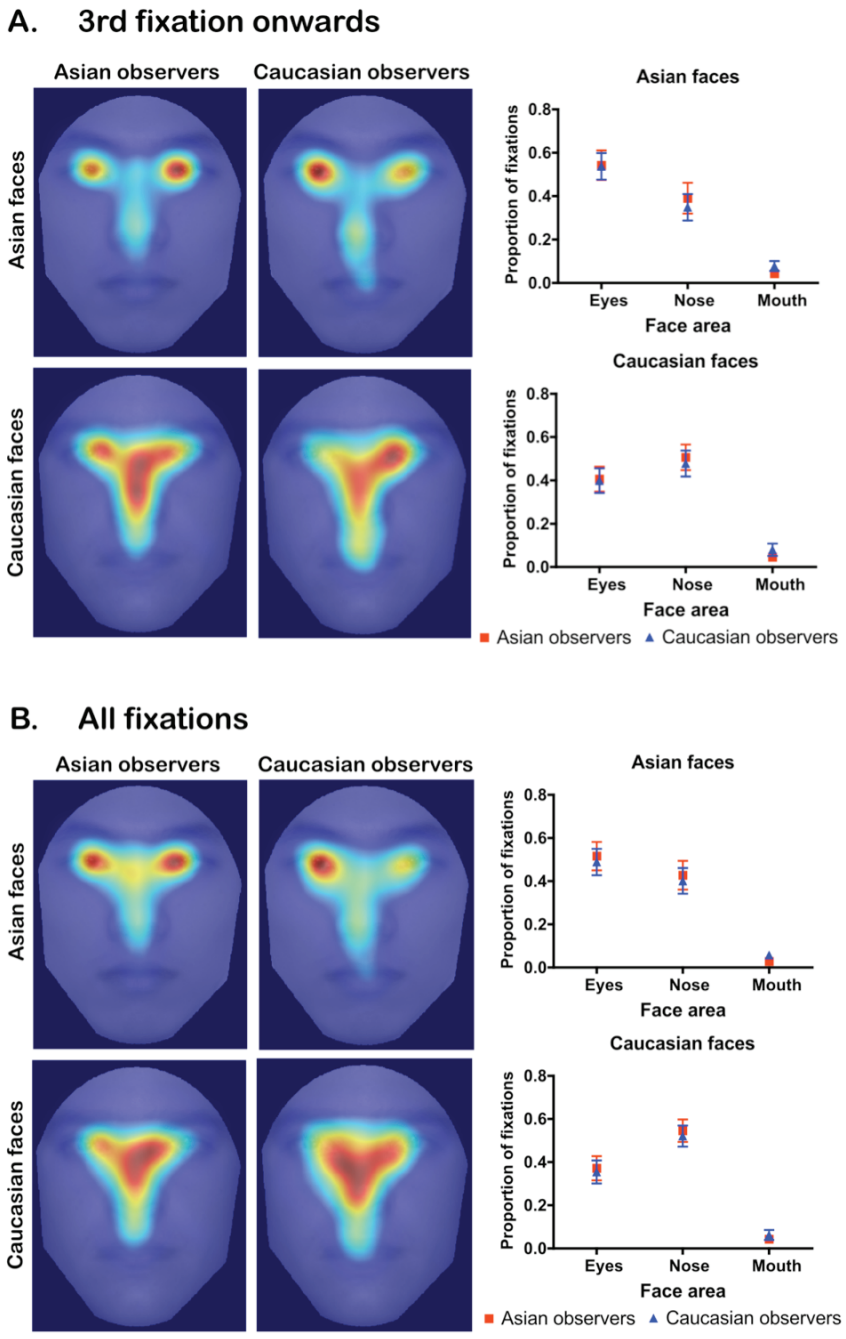


Figure S2 Distributions of (a) the third and subsequent fixation positions and (b) all fixations from Experiment 2. The heat maps (left) were generated using similar methods as in Fig. 3a, each of which considers all 500 trials from the 16 observers in a condition. Warmer colour indicates a higher fixation frequency of the pixel, and vice versa. Proportions of fixations (right) on distinct face areas covering the eyes, nose, and mouth (see definition of face areas in Fig. 5b), averaged by observer culture and face race, were also plotted. The error bar denotes ±1 *SEM*.

3rd fixation onwards

(a)

Observer culture	Face race	*t*	*df*	*P*-value
Asian	Asian	1.10	15	.29
	Caucasian	0.88	15	.39
Caucasian	Asian	1.57	15	.14
	Caucasian	0.70	15	.50

(b)

Eyes:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Face race (F)	34.70	1	<.001 *
Main effects (between-subjects)			
Observer culture (O)	0.005	1	.94
Two-way interactions			
O × F	0.002	1	.96
Residual		30	

Nose:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Face race (F)	27.49	1	<.001 *
Main effects (between-subjects)			
Observer culture (O)	0.17	1	.68
Two-way interactions			
O × F	0.08	1	.78
Residual		30	

Mouth:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Face race (F)	0.04	1	.85
Main effects (between-subjects)			
Observer culture (O)	1.42	1	.24
Two-way interactions			
O × F	0.001	1	.97
Residual		30	

Table S6 Statistical results on the third and subsequent fixations (Experiment 2). (a) Paired-samples t-test (two-tailed) results comparing fixation frequencies of the left-eye and the right-eye regions for each condition. No significant difference was found for each comparison. (b) Mixed-model repeated measures ANOVA results on fixation proportions of the eyes, nose, and mouth. Significant *p*-values (i.e., $p < .05$) are marked with asterisks (*).

All fixations

(a)

Observer culture	Face race	*t*	*df*	*P*-value
Asian	Asian	0.67	15	.51
	Caucasian	1.63	15	.12
Caucasian	Asian	0.73	15	.48
	Caucasian	1.69	15	.11

(b)

Eyes:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Face race (F)	43.97	1	<.001 *
Main effects (between-subjects)			
Observer culture (O)	0.08	1	.79
Two-way interactions			
O × F	0.06	1	.81
Residual		30	

Nose:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Face race (F)	34.80	1	<.001 *
Main effects (between-subjects)			
Observer culture (O)	0.11	1	.75
Two-way interactions			
O × F	0.002	1	.97
Residual		30	

Mouth:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Face race (F)	1.50	1	.23
Main effects (between-subjects)			
Observer culture (O)	1.10	1	.30
Two-way interactions			
O × F	0.39	1	.54
Residual		30	

Table S7 Statistical results on all fixations (Experiment 2). (a) Paired-samples t-test (two-tailed) results comparing fixation frequencies of the left-eye and the right-eye regions for each condition. No significant difference was found for each comparison. (b) Mixed-model repeated measures ANOVA results on fixation proportions of the eyes, nose, and mouth. Significant *p*-values (i.e., $p < .05$) are marked with asterisks (*).

Experiment 3: Famous face familiarity

Difference of fixation means

Vertical dimension	95% CI (deg)
Cultural difference: Asian – Caucasian observers	
Stimuli: Asian faces	(–2.33, 0.79)
Stimuli: Caucasian faces	(–1.77, 0.97)
Stimulus difference: Asian – Caucasian faces	
Observers: Asian	(–1.09, 0.41)
Observers: Caucasian	(–0.32, 0.59)

Horizontal dimension	95% CI (deg)
Cultural difference: Asian – Caucasian observers	
Stimuli: Asian faces	(–0.12, 1.07)
Stimuli: Caucasian faces	(–0.44, 0.92)
Stimulus difference: Asian – Caucasian faces	
Observers: Asian	(0.18, 0.62) *
Observers: Caucasian	(–0.19, 0.50)

Difference of fixation standard deviations

Vertical dimension	95% CI (deg)
Cultural difference: Asian – Caucasian observers	
Stimuli: Asian faces	(–1.37, 0.64)
Stimuli: Caucasian faces	(–1.10, 0.48)
Stimulus difference: Asian – Caucasian faces	
Observers: Asian	(–0.15, 0.45)
Observers: Caucasian	(–0.15, 0.56)

Horizontal dimension	95% CI (deg)
Cultural difference: Asian – Caucasian observers	
Stimuli: Asian faces	(–1.08, 0.22)
Stimuli: Caucasian faces	(–1.16, 0.25)
Stimulus difference: Asian – Caucasian faces	
Observers: Asian	(–0.45, 0.08)
Observers: Caucasian	(–0.36, 0.06)

Table S8 The 95% confidence intervals (CI) of the differences of bootstrapped means and standard deviations of the distributions of the first fixation positions from Experiment 3 (Free eye movements, Famous face familiarity). The significant difference is marked with an asterisk (*).

Comparisons between Experiment 1 and 3 results

Vertical dimension:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Experiment (T)	0.41	1	.54
Face race (F)	2.71	1	.13
Main effects (between-subjects)			
Observer culture (O)	0.09	1	.78
Two-way interactions			
T × O	3.00	1	.11
T × F	5.83	1	.03 *
O × F	1.35	1	.27
Three-way interaction			
T × O × F	1.15	1	.30
Residual		12	

Post hoc tests for T × F:

Face race (F)	Experiments (T)	Mean difference (deg)	*SEM* (deg)	*P*-value
Asian (A)	Exp 1 – Exp 3	–0.54	0.25	.051
Caucasian (C)	Exp 1 – Exp 3	0.27	0.29	.36

Experiment (T)	Face races (F)	Mean difference (deg)	*SEM* (deg)	*P*-value
Exp 1	A – C	–0.88	0.41	.053
Exp 3	A – C	–0.07	0.23	.78

Horizontal dimension:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Experiment (T)	4.56	1	.054
Face race (F)	2.24	1	.16
Main effects (between-subjects)			
Observer culture (O)	0.82	1	.38
Two-way interactions			
T × O	0.34	1	.57
T × F	3.84	1	.07
O × F	0.000	1	.998
Three-way interaction			
T × O × F	2.22	1	.16
Residual		12	

Table S9 Mixed-model repeated measures ANOVA results on potential variations of the first fixation positions between Experiments 1 and 3, analyzed separately for the vertical and the horizontal dimensions. The significant *p*-value (i.e., $p < .05$) is marked with an asterisk (*).

Experiment 4: Restricted number of fixations

The first fixation

Vertical dimension:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Trial group (T)	1.32	1	.27
Face race (F)	9.58	1	< .01 *
Main effects (between-subjects)			
Observer culture (O)	0.33	1	.58
Two-way interactions			
T × O	0.88	1	.36
T × F	0.87	1	.37
O × F	1.06	1	.32
Three-way interaction			
T × O × F	1.17	1	.30
Residual		14	

Horizontal dimension:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Trial group (T)	1.17	1	.30
Face race (F)	1.36	1	.26
Main effects (between-subjects)			
Observer culture (O)	0.94	1	.35
Two-way interactions			
T × O	0.84	1	.37
T × F	0.97	1	.34
O × F	0.12	1	.74
Three-way interaction			
T × O × F	1.07	1	.32
Residual		14	

Table S10 Mixed-model repeated measures ANOVA results on potential variations of the first fixation positions across trials in Experiment 4: Restricted number of fixations, analyzed separately for the vertical and the horizontal dimensions. Here, only the 16 observers who performed all four conditions are included in the analysis. The significant *p*-value (i.e., $p < .05$) is marked with an asterisk (*).

Vertical dimension:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Trial group (T)	0.56	1	.47
Face race (F)	7.43	1	.02 *
Main effects (between-subjects)			
Observer culture (O)	0.06	1	.81
Two-way interactions			
T × O	0.56	1	.47
T × F	0.42	1	.53
O × F	1.25	1	.29
Three-way interaction			
T × O × F	1.10	1	.32
Residual		10	

Horizontal dimension:

Parameter	*F*	*df*	*P*-value
Main effect (within-subjects)			
Trial group (T)	0.79	1	.39
Face race (F)	0.007	1	.94
Main effects (between-subjects)			
Observer culture (O)	0.61	1	.45
Two-way interactions			
T × O	0.81	1	.39
T × F	0.87	1	.37
O × F	0.06	1	.81
Three-way interaction			
T × O × F	0.34	1	.57
Residual		10	

Table S11 Mixed-model repeated measures ANOVA results on potential variations of the first fixation positions across trials in Experiment 4: Restricted number of fixations, analyzed separately for the vertical and the horizontal dimensions. Here, the 12 observers who performed only the 5000 ms condition are included in the analysis. The significant *p*-value (i.e., $p < .05$) is marked with an asterisk (*).

Difference of fixation means

Vertical dimension	95% CI (deg)	
Cultural difference: Asian – Caucasian observers		
Stimuli: Asian faces	(–0.63, 0.45)	
Stimuli: Caucasian faces	(–0.94, 0.28)	
Stimulus difference: Asian – Caucasian faces		
Observers: Asian	(–0.34, –0.09)	*
Observers: Caucasian	(–0.85, –0.15)	*

Horizontal dimension	95% CI (deg)
Cultural difference: Asian – Caucasian observers	
Stimuli: Asian faces	(–0.51, 1.35)
Stimuli: Caucasian faces	(–0.44, 1.26)
Stimulus difference: Asian – Caucasian faces	
Observers: Asian	(–0.05, 0.20)
Observers: Caucasian	(–0.05, 0.16)

Difference of fixation standard deviations

Vertical dimension	95% CI (deg)
Cultural difference: Asian – Caucasian observers	
Stimuli: Asian faces	(–0.18, 0.35)
Stimuli: Caucasian faces	(–0.47, 0.28)
Stimulus difference: Asian – Caucasian faces	
Observers: Asian	(–0.10, 0.07)
Observers: Caucasian	(–0.37, 0.01)

Horizontal dimension	95% CI (deg)
Cultural difference: Asian – Caucasian observers	
Stimuli: Asian faces	(–0.33, 0.39)
Stimuli: Caucasian faces	(–0.20, 0.46)
Stimulus difference: Asian – Caucasian faces	
Observers: Asian	(–0.08, 0.10)
Observers: Caucasian	(–0.05, 0.22)

Table S12 The 95% confidence intervals (CI) of the differences of bootstrapped means and standard deviations of the distributions of the first fixation positions from Experiment 4 (Restricted number of fixations). Here, only the 16 observers who performed all four conditions are included in the analysis. Significant differences are marked with asterisks (*).

The second fixation

The distributions of the second fixations are shown in Fig. S3a in form of heat maps considering fixations of all individual trials of the three stimulus presentation conditions (two fixations, three fixations, 5000 ms; the one-fixation condition was not included as it could not produce the second fixations). The second fixations reported here were generally consistent with those from Experiment 2 (with 1500 ms stimulus presentations), where an effect of observer culture was absent (Fig. S3b). As shown previously, the second fixations were more scattered than the first fixations but still concentrated on the eyes and nose regions. Observers, regardless of their cultures, directed their second fixations slightly more frequently to the eyes than to the nose of Asian faces, but the trend was reversed when shown Caucasian faces.

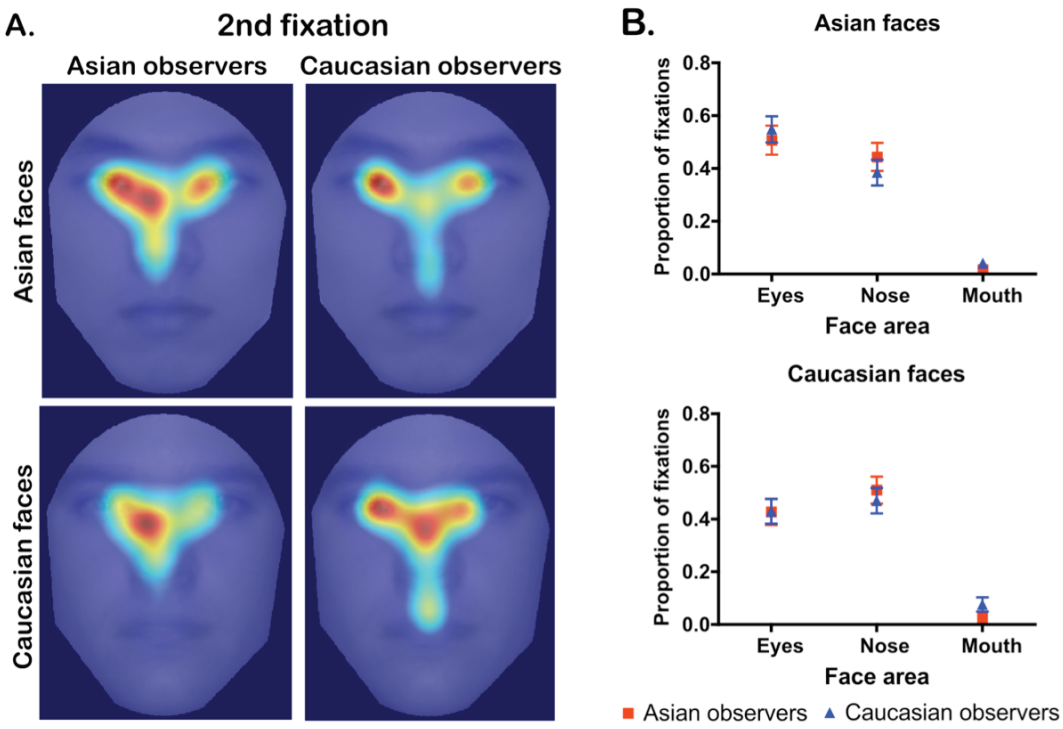


Figure S3 (a) Distributions of the second fixation positions from Experiment 4. Each heat map considers all individual trials obtained from the three stimulus presentation conditions (two fixations, three fixations, 5000 ms), generated using similar methods as in Fig. 3a. Warmer colour indicates a higher fixation frequency of the pixel, and vice versa. (b) Proportions of the second fixations on distinct face areas covering the eyes, nose, and mouth, averaged across the three stimulus presentation conditions for each observer culture and face race. The error bar denotes ±1 *SEM*.

All fixations

Figure S4 shows the distributions of all fixations for the four stimulus presentation conditions (i.e., the first fixations only for the one-fixation condition, the first and second fixations together for the two-fixation condition, the first, second, and third fixations together for the three-fixation condition, and all 10.89 ± 1.85 (*SD*) fixations for the 5000 ms condition). Importantly, we found that fixation patterns are highly consistent across observer culture and face race for prolonged stimulus presentations of at least 5000 ms. Fixations were initially concentrated just below the eyes close to the vertical midline of the face, and subsequently spread out to predominantly the eyes and nose regions, and remained there. During an average 5000 ms presentation, these regions together accounted for over 87% of total fixations (Fig. S5). Only 5.3% of saccades were directed to the mouth, on average. The rest of the face got less than 7% of saccades. Thus, despite prolonged stimulus exposure, our saccades remain focused at internal features of the face.

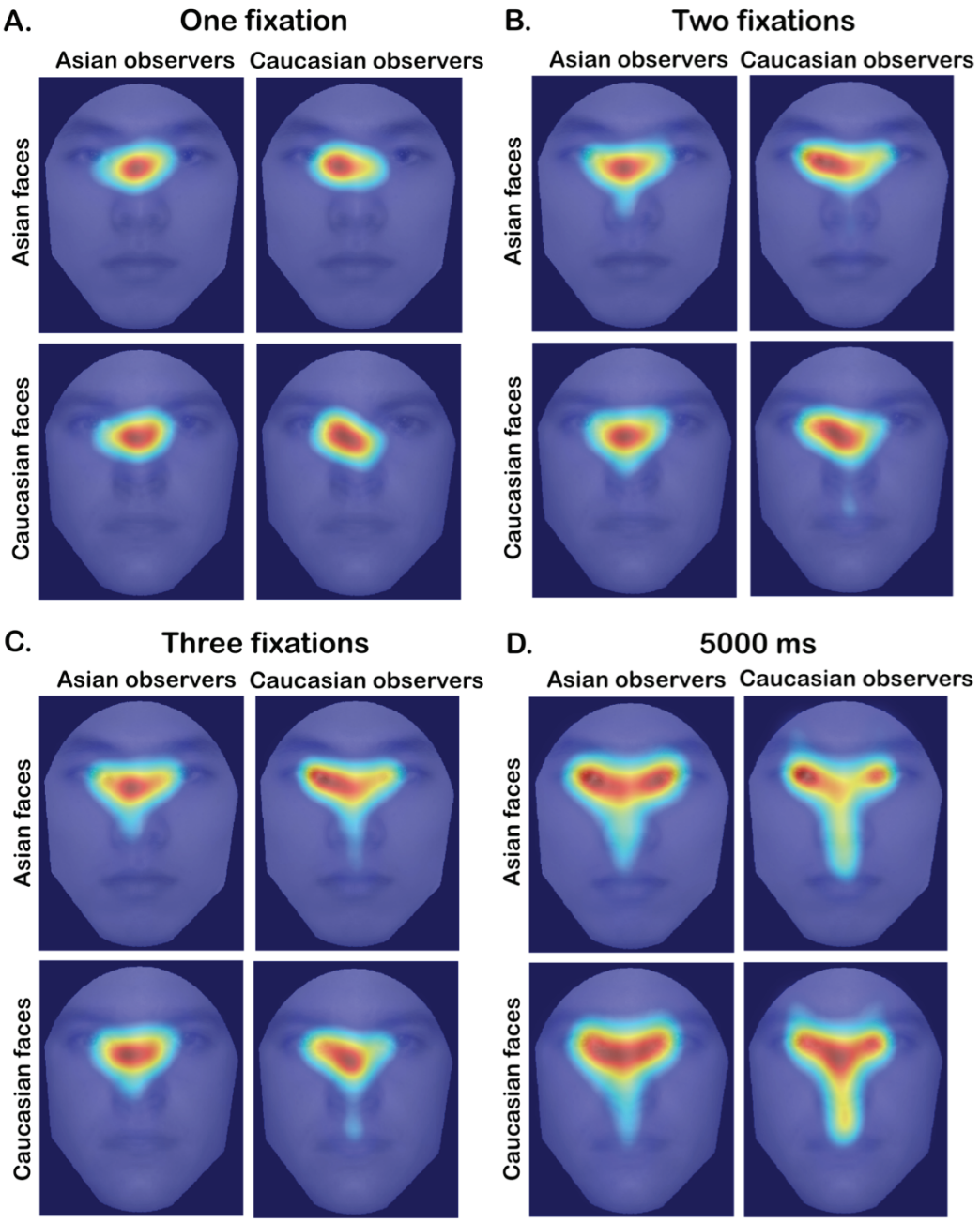


Figure S4 Fixation distributions from Experiment 4 where (a) one fixation, (b) two fixations, (c) three fixations, or (d) an unrestricted duration of 5000 ms were allowed on viewing the face stimulus. Each heat map considers all fixations throughout every trial, otherwise generated using similar methods as in Fig. 3a. Warmer colour indicates a higher fixation frequency of the pixel, and vice versa.

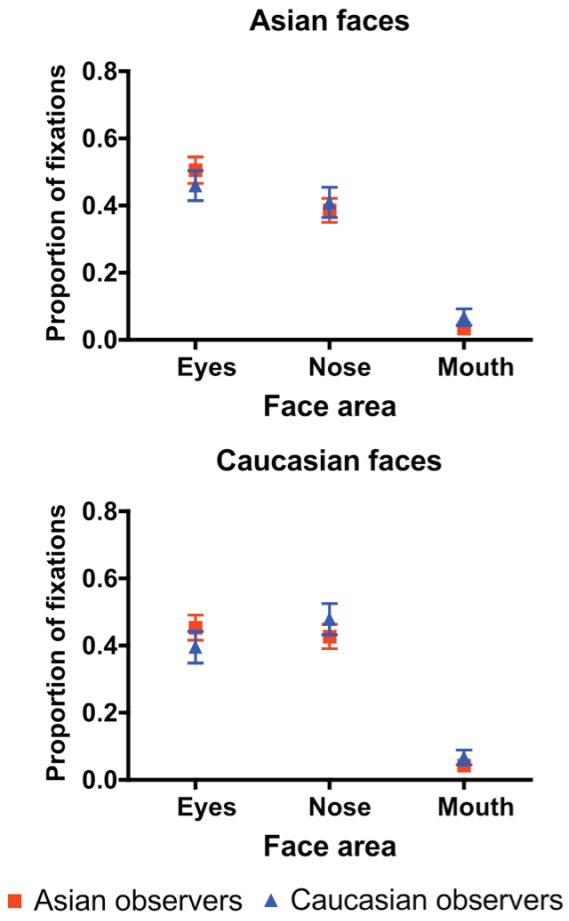


Figure S5 Proportions of all fixations on distinct face areas covering the eyes, nose, and mouth, averaged by observer culture and face race, for the 5000 ms condition in Experiment 4. The error bar denotes ±1 *SEM*.

Justifications of using 100 faces in the natural systems analysis

It could be possible that 100 faces might not reflect accurately the true value of the optimal fixation position. Thus, we expanded our sample to 150 Caucasian faces, all taken from a single database that were photographed in-house as the face stimuli used in the experiments, such that lighting, viewpoint, distance, and expression were tightly controlled. Then, these 150 faces were clustered into 15 groups of 10 faces. The FIO was conducted on each group, resulting in 15 maximum-performance fixation positions. We then computed the standard error across groups as a function of the number of total faces in the sample (Fig. S6). While error continues to decrease with increasing number of groups, it reaches a very small value (0.12°) within 10 groups (100 faces in total). Thus, extending our analysis to 100 faces for each race group gives a good estimate of the optimal point of fixation for the population of faces.

We also assessed how the precision of the FIO's maximum-performance fixation position is affected by the variation introduced through the use of multiple face-image databases. We collected 1,000 Caucasian faces (100 groups of 10) from diverse sources, leading to large variations in lighting and moderate variations in viewpoint, expression, and resolution. Not surprisingly, these added sources of variance increased the error across groups (Fig. S7). However, with 100 faces (10 groups) the average error was only 0.5°, lending confidence to our use of 100 faces for the NSA.

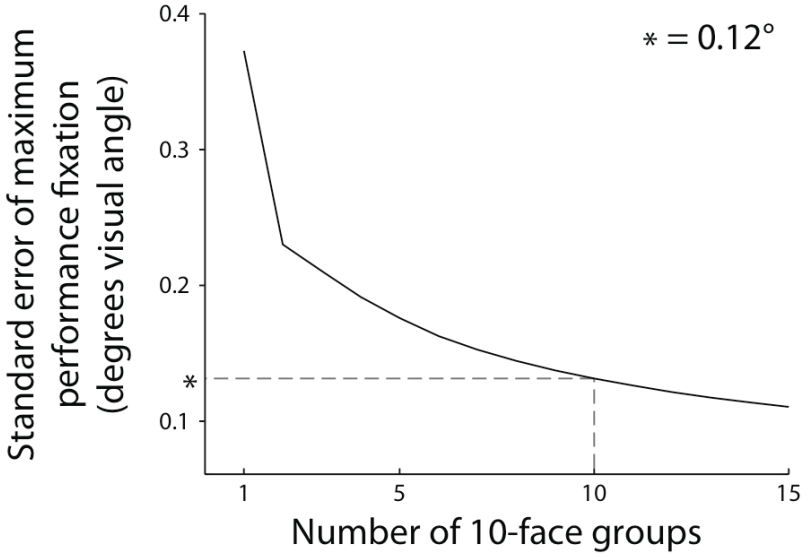


Figure S6 Standard error of the fixation position yielding maximum performance, averaged across face groups, as a function of the number of total faces in the sample. The 150 Caucasian faces (15 groups of 10 faces) used here were all taken from the same database.

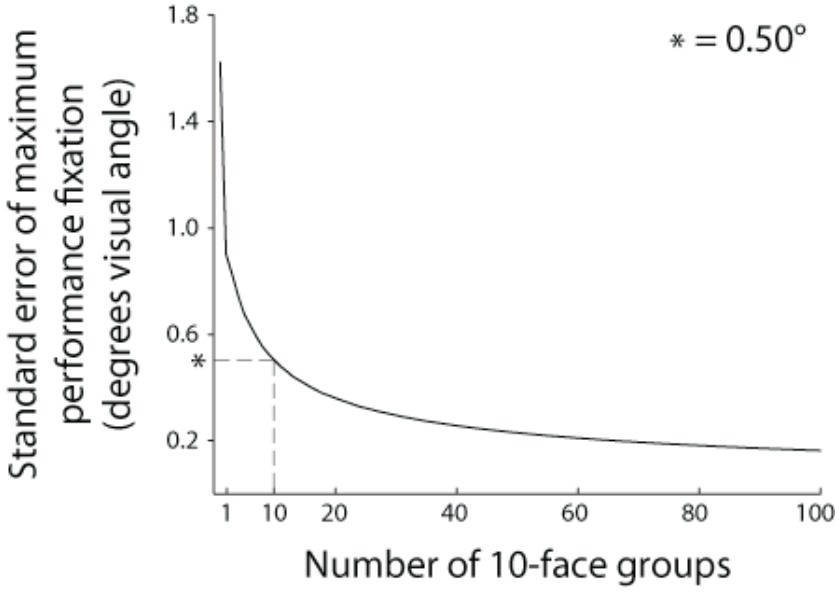


Figure S7 Standard error of the fixation position yielding maximum performance, averaged across face groups, as a function of the number of total faces in the sample. Here, we used 1,000 Caucasian faces (100 groups of 10 faces) collected from diverse sources.