

Using principal component analysis to characterize eye movement fixation patterns during face viewing

Kira Wegner-Clemens

Department of Neurosurgery and Core for Advanced MRI, Baylor College of Medicine, Houston, TX

Johannes Rennig

Department of Neurosurgery and Core for Advanced MRI, Baylor College of Medicine, Houston, TX

John F. Magnotti

Department of Neurosurgery and Core for Advanced MRI, Baylor College of Medicine, Houston, TX

Michael S. Beauchamp

Department of Neurosurgery and Core for Advanced MRI, Baylor College of Medicine, Houston, TX



Human faces contain dozens of visual features, but viewers preferentially fixate just two of them: the eyes and the mouth. Face-viewing behavior is usually studied by manually drawing regions of interest (ROIs) on the eyes, mouth, and other facial features. ROI analyses are problematic as they require arbitrary experimenter decisions about the location and number of ROIs, and they discard data because all fixations within each ROI are treated identically and fixations outside of any ROI are ignored. We introduce a data-driven method that uses principal component analysis (PCA) to characterize human face-viewing behavior. All fixations are entered into a PCA, and the resulting eigenimages provide a quantitative measure of variability in face-viewing behavior. In fixation data from 41 participants viewing four face exemplars under three stimulus and task conditions, the first principal component (PC1) separated the eye and mouth regions of the face. PC1 scores varied widely across participants, revealing large individual differences in preference for eye or mouth fixation, and PC1 scores varied by condition, revealing the importance of behavioral task in determining fixation location. Linear mixed effects modeling of the PC1 scores demonstrated that task condition accounted for 41% of the variance, individual differences accounted for 28% of the variance, and stimulus exemplar for less than 1% of the variance. Fixation eigenimages provide a useful tool for investigating the relative importance of the different factors that drive human face-viewing behavior.

Introduction

Human faces are perhaps the most important visual stimulus that we encounter, prompting extensive investigations of eye movement behavior during face viewing. Among the dozens of visual features in a face, observers spend most of their time fixating the eyes and the mouth (Yarbus, 1967). Recently, it has been shown that there are substantial individual differences in face-viewing behavior: Some participants exclusively fixate the eyes or the mouth of the viewed face while others balance eye and mouth fixations to varying degrees. These interindividual differences are consistent for intervals between tests as long as 18 months (Mehoudar, Arizpe, Baker, & Yovel, 2014); are present when viewing static, silent faces (Mehoudar et al., 2014; Perlman et al., 2009; Rayner, Li, Williams, Cave, & Well, 2007; Royer et al., 2018) or dynamic, talking faces (Gurler, Doyle, Walker, Magnotti, & Beauchamp, 2015; Klin, Jones, Schultz, Volkmar, & Cohen, 2002) either in the laboratory or the real world (Peterson, Lin, Zaun, & Kanwisher, 2016); and may reflect individual differences in optimal behavior (Peterson & Eckstein, 2012). Individual differences in face-viewing behavior are linked to other important psychological phenomena. For instance, individuals who prefer to fixate the mouth of a viewed face are better able to understand noisy audiovisual speech (Rennig, Wegner-Clemens, & Beauchamp, in press).

Studies examining face viewing typically quantify eye movement behavior by measuring the amount of

Citation: Wegner-Clemens, K., Rennig, J., Magnotti, J. F., & Beauchamp, M. S. (2019). Using principal component analysis to characterize eye movement fixation patterns during face viewing. *Journal of Vision*, 19(13):2, 1–15, <https://doi.org/10.1167/19.13.2>.

<https://doi.org/10.1167/19.13.2>

Received March 20, 2019; published November 5, 2019

ISSN 1534-7362 Copyright 2019 The Authors



time that participants fixate within different regions of interest (ROIs). For instance, in the study of Gurler et al. (2015), rectangular ROIs were hand-drawn around each eye and the mouth region of the viewed face (Figure 1A). This approach has several potential problems. First, ROIs require an arbitrary decision by the experimenter about what parts of the face constitute a particular ROI. For instance, one could argue that a mouth ROI should encompass not only the mouth proper, but also the peri-mouth region of the face because the peri-mouth region contains visual information about mouth movements due to the structure of the facial musculature (Irwin et al., 2018). Although a spatially extensive peri-mouth ROI might seem logical, it has the potential to produce biased estimates of gaze behavior. If the time spent fixating a large peri-mouth ROI is greater than the time spent fixating a smaller eye ROI, this could be due to the size imbalance between the ROIs (under the null hypothesis that all face locations are fixated with equal probability, a larger ROI contains more fixations) rather than an actual preference for mouth fixations.

A second drawback of ROI analyses is that they discard data, a statistically undesirable property. Regardless of where ROI borders are placed, fixations occurring just outside the border are ignored although it seems reasonable that a fixation just outside the mouth ROI carries some evidence for mouth-viewing preference or vice versa for a fixation just outside the eye ROI. One solution to this concern is to create two very large but equal-sized ROIs, one that covers the entire top half of the face and one that covers the entire bottom half of the face. This approach is illustrated in Figure 1B, adapted from Rennig and Beauchamp (2018).

A disadvantage to splitting the face into upper and lower ROIs is that it decreases specificity; for instance, fixations of the eye and ear region are equivalent in this analysis despite the distance between them. Another solution is to draw multiple ROIs that together tile the entire face (Armann & Bühlhoff, 2009; Nguyen, Isaacowitz, & Rubin, 2009; Sæther, Van Belle, Laeng, Brennen, & Øvervoll, 2009; Schurgin et al., 2014). Although ensuring that all fixations are included in the analysis, this method requires arbitrary choices about the total number of ROIs and their boundaries. For each ROI, the researcher must hand-draw a spatially consistent region on each individual stimulus presented to participants, highlighting an additional difficulty of ROI analyses: their labor-intensive nature.

To address these concerns, we developed a new method to analyze face viewing that relies on principal component analysis (PCA), a mathematical technique to transform a high-dimensional data set (every fixation

location in every viewed face for every observer) into a lower-dimensional data set that captures the greatest possible amount of variance. Because the input to the PCA are 2-D fixation locations, the output consists of a set of orthogonal eigenimages, and each eigenimage is a spatial map of frequently fixated locations on the viewed face.

The PCA approach avoids the major drawbacks of ROI analyses. Rather than arbitrarily selecting the location, extent, and total number of ROIs, PCA uses a data-driven approach to select the eye-fixation patterns that account for the most variance. All fixations are included in the PCA calculation (unlike ROI analyses, which discard fixations that occur outside of any ROI), and labor-intensive manual tracing of individual face regions is not required. It additionally potentially allows multiple maps of differences in face-viewing behavior (for example: mouth vs. eyes, left face vs. right face) to be analyzed without creating new ROIs for each possible distribution).

To demonstrate the utility of the PCA approach, we applied it to characterize the relationship between individual differences in face-viewing behavior and changes in task and stimulus.

Task instructions influence face-viewing behavior (Kanan, Bseiso, Ray, Hsiao, & Cottrell, 2015). For instance, the task of determining facial gender increases fixations to the upper face relative to an emotion-determination task because the eye region is diagnostic for gender (Armann & Bühlhoff, 2009; Sæther et al., 2009; Schyns, Bonnar, & Gosselin, 2002). This task modulation could interact with individual differences in several ways. For instance, participants who already prefer to fixate the eyes might fixate the eyes even more during a gender-discrimination task, or they might show little change because they have accumulated greater expertise in processing information from the eye region of viewed faces.

Stimulus differences are also expected to change face-viewing behavior. The choice of fixation location is driven by visual saliency, and a moving mouth is highly salient. Therefore, one might predict that, for participants performing a given task (such as gender discrimination), viewing a dynamic talking face might lead to more mouth fixations than viewing a static face. Similarly, differences between the features of viewed faces could also drive viewing behavior. For instance, some faces might have more interesting mouths or eyes, leading to greater fixation of these regions.

We examined face-viewing behavior within individual participants under different stimulus, task, and exemplar conditions in order to estimate the relative importance of individual differences relative to task and stimulus effects.

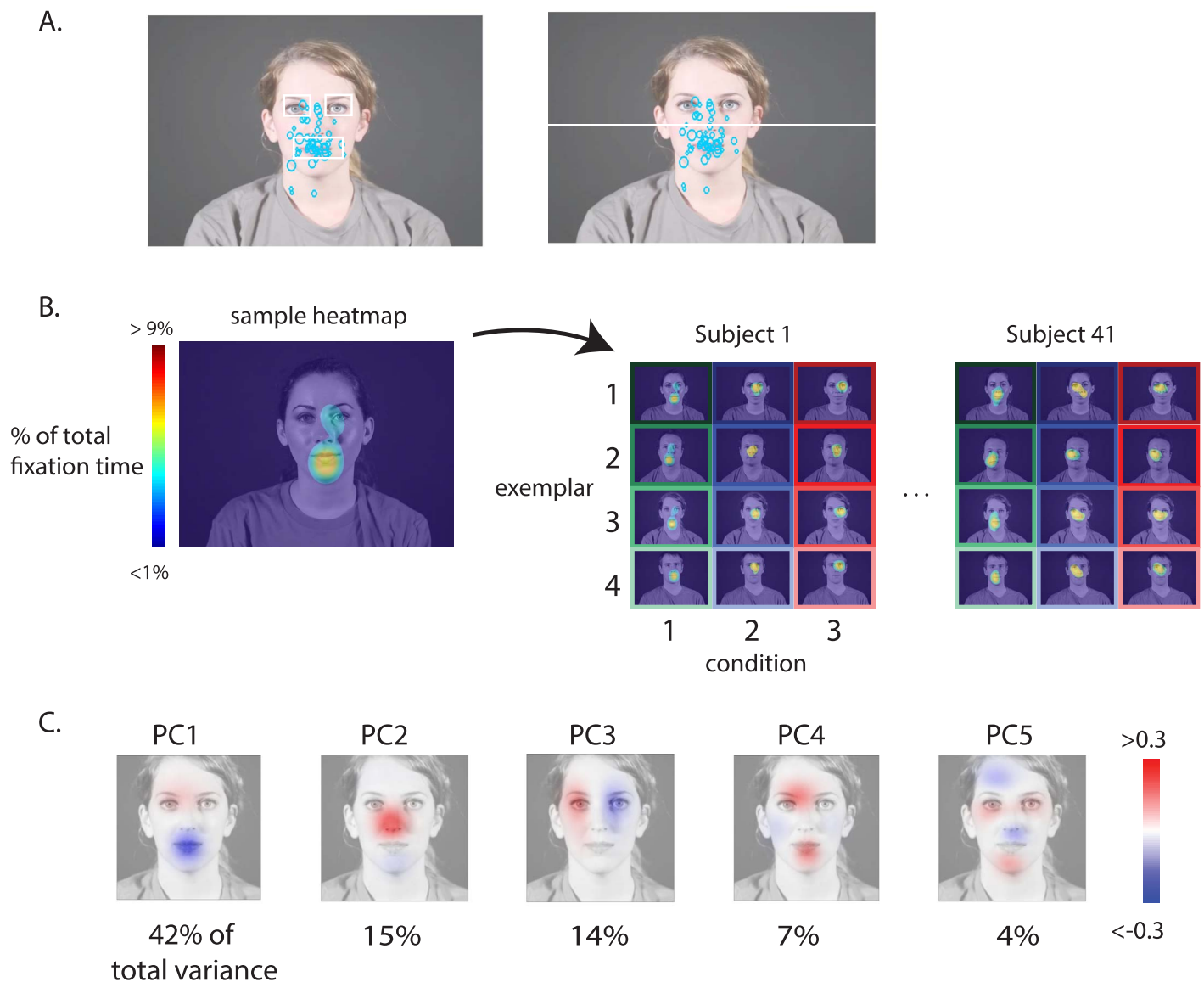


Figure 1. Comparison of ROI and PCA methods. (A) Two ROI methods for analyzing fixation data during face viewing. A single participant repeatedly viewed a face image. Cyan ellipses show the location of each individual fixation with the size of each ellipse proportional to the duration of each fixation. In the left panel, ROIs were manually drawn on the eye and mouth regions of the face (white boxes). Fixations within the eye ROI were classified as eye fixations, and fixations within the mouth ROI were classified as mouth fixations (Gurler et al., 2015). In the right panel, the face was bisected (white line). Fixations in the upper half of the face were classified as upper face fixations and fixations in the lower half of the face were classified as lower face fixations (Rennig & Beauchamp, 2018). (B) The PCA method for analyzing fixation data during face viewing. A heat map was constructed for each participant's fixations for each exemplar during each condition. The overlay color indicates the percentage of total fixation time spent at that location with warmer colors indicating more time. All heat maps from all participants, all stimulus exemplars, and all task conditions were entered in the PCA (shown schematically to the right as 12 heat maps for two participants). The color and shade of the border surrounding each heat map illustrates the exemplar and condition (same color scale used in other figures). (C) Results of PCA analysis, showing the first five PCs for the fixation data from panel B. The underlay shows the stimulus image converted to grayscale. The color overlay shows the fixation eigenimage for that component with the color corresponding to the parameter estimate for that location in the stimulus image (red indicates positive values, and blue indicates negative values; no color indicates values near zero). The percentage of variation accounted for by each component is displayed underneath each eigenimage.

Methods

Participants and materials

Forty-one participants (25 female, mean age 22, range 18–31) provided written, informed consent under an experimental protocol approved by the Committee for the Protection of Human Participants of the Baylor College of Medicine, Houston, Texas. The work was carried out in accordance with the Declaration of Helsinki.

Participants' eye movements were recorded using an infrared eye tracker (Eye Link 1000 Plus, SR Research Ltd., Ottawa, Ontario, Canada) as they viewed dynamic or static faces presented on a display (Display++ LCD Monitor, 32-in., $1,920 \times 1,080$, 120 Hz, Cambridge Research Systems, Rochester, UK) and listened to speech through speakers located on either side of the screen. To increase the stability of the eye tracking, participants rested on a chin rest (University of Houston School of Optometry, Houston, Texas) placed 90 cm from the display.

The stimulus set included eight 2-s audiovisual speech videos and four static frames from the same videos. Four speakers were filmed in the laboratory looking straight at the camera with a gray background and wearing a gray shirt to maintain as much similarity as possible across videos. To maintain alignment between speakers, the camera was not moved between recordings, and speakers were aligned with markers on the wall, which were later edited out in Final Cut Pro. Each speaker recorded two speech videos, one in which they said “ba” and the other in which they said “ga.”

Experimental design: Overview

Participants underwent two separate testing blocks within a single testing session. In the first testing block (duration of ~14 min), participants performed a speech identification task. The stimuli in the first block consisted of audiovisual speech movies of different speakers speaking the congruent audiovisual syllables “ba” or “ga.” Participants performed a two-alternative, forced choice, deciding which of the two syllables was presented on each trial.

In the second testing block (duration of ~16 min), participants performed a gender-discrimination task. The stimuli consisted of either the same audiovisual speech movies presented in the first block or silent still frames from the movies. Participants performed a two-alternative, forced choice, deciding whether the actor in each stimulus was male or female.

A goal of our study was to pit two sources of variability against each other: variability introduced by

task demands and by individual differences. Therefore, we chose two tasks (speech and gender identification) known to evoke very different patterns of eye movements. Speech tasks drive fixations to the mouth (Buchan, Paré, & Munhall, 2008; Vatikiotis-Bateson, Eigsti, Yano, & Munhall, 1998), and gender tasks drive fixations to the eyes (Armann & Bühlhoff, 2009; Sæther et al., 2009; Schyns et al., 2002). The speech and gender tasks were designed to be very simple and extremely easy (“ba” or “ga,” “male” or “female”) minimizing confounds of task difficulty or cognitive effort. Comparing the patterns of eye fixations made in response to the audiovisual speech movies viewed in the first (speech) and second (gender) testing blocks isolated task differences because the movies were identical. Comparing fixations between the movies and still frames presented within the second (gender) testing block isolated stimulus differences.

A secondary goal of the study was to examine the variability contributed by different stimuli relative to individual and task differences. In order to estimate this variability, we presented a limited number of stimulus exemplars recorded from four different actors many times. This allowed us to estimate the response to individual stimuli and determine whether the response to different stimuli differed. An alternative approach would be to present a different stimulus on each trial, but this would not allow for an estimation of variability contributed by different stimuli.

Experimental design: Details

Each trial was preceded by an interstimulus interval of 1 s, in which a fixation crosshair presented outside of the location where the face would appear in order to simulate natural viewing conditions in which faces rarely appear at the center of gaze (Gurler et al., 2015). Fixations appeared just outside of where the face image would appear, randomly selected from one of four different locations. As soon as the face image appeared in the center of the screen, the fixation crosshair disappeared, and participants were free to fixate anywhere. The face image remained on screen for 2 s, during which the participants reported their behavioral choice by pressing a key on a computer keyboard. The face stimuli subtended approximately 10 by 13 cm (6° wide by 8° high). Instead of requiring participants to fixate noncentrally followed by central face presentation, an alternative approach would be to present the face in different screen locations (Hsiao & Cottrell, 2008).

In the first testing block (task: speech), data from 160 trials was analyzed, consisting of 4 different speakers * 2 syllables (congruent audiovisual “ba” or “ga”) * 20 repetitions of each video. Two types of additional trials

were presented but not analyzed. First, trials in which incongruent audiovisual syllables (auditory “ba” and visual “ga”) were presented; these trials were not analyzed as the incongruence between auditory and visual speech might have influenced eye movements. Second were trials with an additional speaker that was not included in the gender task; these trials were not analyzed because a goal of the analysis was to measure the effect of speaker, independent of task.

In the second testing block (task: gender), 320 total trials comprising two different stimulus types were analyzed; 160 trials consisted of the same stimuli presented in the first block (4 speakers * 2 syllable movies * 20 repetitions) and 160 trials consisted of static versions of these stimuli (still frames with no sound): 4 speakers * 40 repetitions. The two stimulus types were randomly intermixed, and participants performed the same gender task on each stimulus.

Testing was paused midway through each testing block in order to recalibrate the eye tracker.

Behavioral performance was near ceiling in all three conditions. Participants accurately identified syllables when viewing dynamic faces (mean of 98% accuracy across participants, range 87.8%–100%) and accurately identified gender when viewing both static (mean 99%, range 94.3%–100%) and dynamic faces (mean 99%, range 95.6%–100%).

Eye-tracking analysis

Eye tracking was performed with a sampling rate of 500 Hz. The eye tracker was calibrated using a nine-target array four times within each testing session (before and halfway through each testing block). A fixation crosshair was presented at the center of the display six times within each testing block. The difference between the measured eye position during these epochs and the screen center was applied to correct the eye-tracking data in the preceding stimulus epoch.

Fixation heat maps were constructed from the eye-tracking data by calculating the percentage of total fixation time at each stimulus location across all trials for that participant, condition, and exemplar. The data was not filtered, thresholded, or smoothed in any other way. A separate fixation map was created for each of the 41 participants viewing each of the four speakers (stimulus exemplars) during each of the three conditions (audiovisual speech stimulus + speech task, audiovisual speech stimulus + gender task, static face stimulus + gender task). This resulted in a total of 492 fixation maps (41 participants * 4 exemplars * 3 conditions). The majority of the fixations were on the face, so heat maps were cropped closely to the face. Cropping also allowed us to align facial features on

each exemplar. The x and y coordinates for the tip of the nose for each exemplar image was identified, and a heat map was constructed for a square region (300 pixels on a side) centered on the nose.

PCA was used to reduce the dimensionality of the 2-D heat maps. This analysis has been termed “eigenfaces” in the computer vision literature (Turk & Pentland, 1991) and, by analogy, “eigengaze” when applied to eye movements (Fookes & Sridharan, 2010).

To reduce processing time, the heat maps were scaled down in size. Each heat map was reduced from the 300×300 stimulus size to a 150×150 matrix. These matrices were reshaped into a linear $22,500 \times 492$ vector, one column per heat map. The matrix was mean-centered by subtracting out the column mean, the covariance matrix ($22,500 \times 22,500$ elements) calculated, and the eigenvectors and eigenvalues determined. Each eigenvalue was divided by the sum of all eigenvalues to determine the percentage variance accounted for.

For visualization, each eigenvector was displayed as a heat map overlaid on the original stimuli image or a still frame from the video clip if the original stimulus was in this format.

Results

Figure 1 illustrates the steps involved in the PCA analysis of face-viewing behavior. First, a fixation heat map was generated for all presentations of a particular stimulus exemplar/task condition/participant combination (Figure 1B). Then, all 492 fixation heat maps generated from the 41 participants * 4 exemplars * 3 conditions were entered into a PCA. Each principal component (PC) consisted of a 2-D matrix with each cell in the matrix corresponding to a location on the face. These PCs were then visualized as fixation eigenimages, heat maps overlaid on a still frame from a single stimulus exemplar (Figure 1C).

The first PC corresponds to eye and mouth looking

The first principal component (PC1) accounted for 42% of the total variance. Strikingly, the PC1 eigenimage contained a positive peak around the eyes and a negative peak around the mouth of the face stimulus, demonstrating that differences in the tendency to fixate the eyes or mouth accounted for the largest proportion of variance in the data set.

PC2 (15% of total variance) was characterized by a positive peak on the center of the face and PC3 (14%) by a positive peak on the right half of the face and a

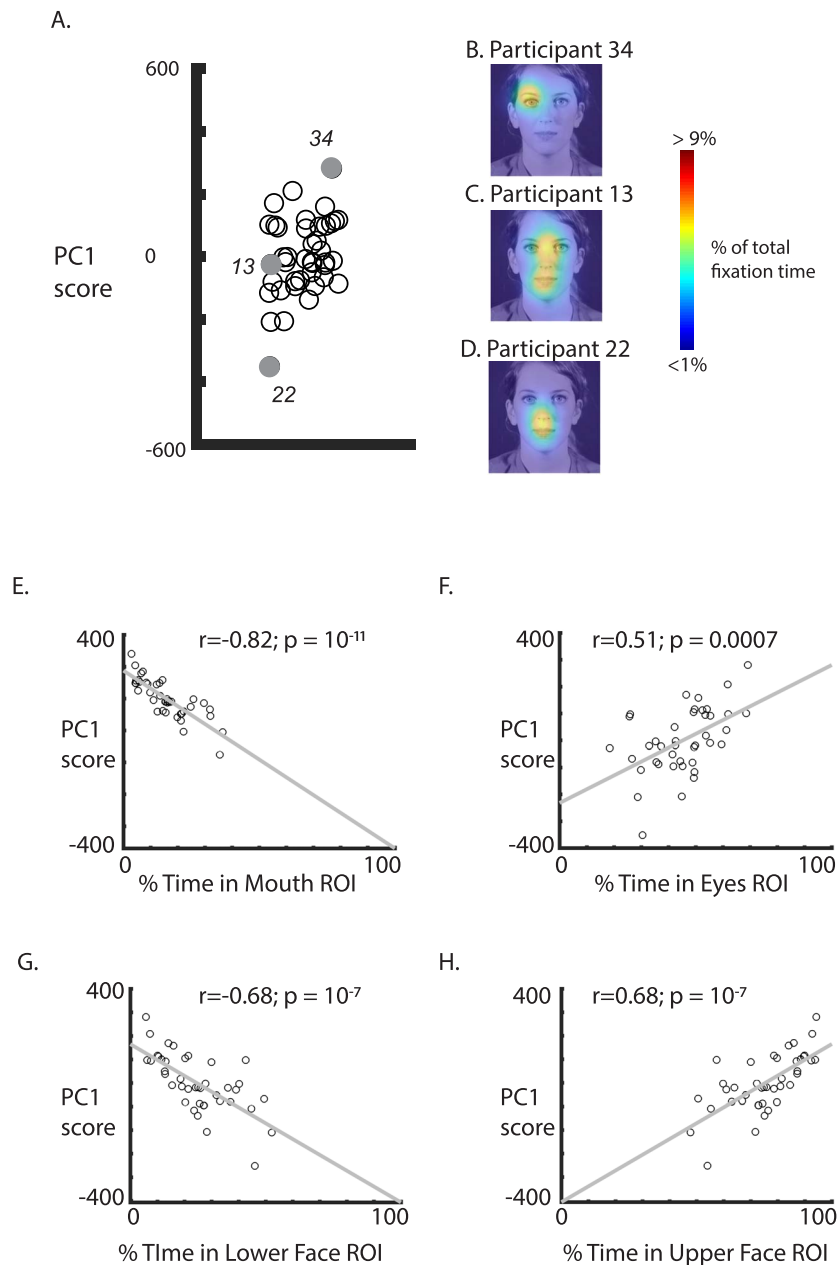


Figure 2. PC1 score distributions and comparison to ROI fixation percentages. (A) PC1 scores are shown for each participant, averaged across stimuli, exemplar, and task. Score represented on the y-axis with the x-axis values jittered for visibility. The symbols representing the subjects with the lowest score (subject 22, PC1 score: -352), median score (subject 13, PC1 score: -19), and highest score (subject 34, PC1 score: $+280$) are labeled and filled in with gray. (B) Fixation heat map for subject 34 based on all fixations. Color overlay indicates percentage of total fixation time spent on that stimulus location. (C) Fixation heat map for subject 13. (D) Fixation heat map for subject 22. (E) PC1 scores for each subject compared to time spent in the mouth region of interest as depicted in the left panel of Figure 1A. Values show correlation and significance values for the least square regression line. (F) PC1 scores compared to time spent in the eye region of interest as depicted in the left panel of Figure 1A. (G) PC1 scores compared to time spent in the lower face region of interest as depicted in the right panel of Figure 1A. (H) PC1 scores compared to time spent in the upper face region of interest as depicted in the right panel of Figure 1A.

negative peak on the left half. PC4 (7%) separated a peak on the forehead and chin regions of the face, and PC5 (4%) separated fixations on the eyes and chin from the forehead and nose.

PCs are a linear transform of the original data set, allowing a weight score for each PC to be calculated for each stimulus exemplar, condition, or participant that entered the analysis. We calculated a PC1 score for

each participant, averaged across stimulus exemplars and conditions (Figure 2A).

To verify that PC1 characterized individual differences in the propensity to fixate the eyes or mouth, we examined the raw fixation heat maps for individual participants across the range of PC1 scores. Participant 34, with the highest PC1 score of +280, exclusively fixated the right eye of the speaker (Figure 2B). Participant 13, with a PC1 score of −19, equivalent to the median PC1 score across participants, showed a balanced fixation pattern, fixating both the upper and lower halves of the face (Figure 2C). Participant 22, with the lowest PC1 score of −352, exclusively fixated the mouth of the speaker (Figure 2D).

We compared PC1 score as a measure of individual differences in face-viewing behavior with two ROI analyses. For the first ROI analysis, an ROI was created around the mouth of each viewed face and the percentage of total fixation time in the mouth ROI calculated (as in Figure 1A), and for the second ROI analysis, two large ROIs were created spanning the top and bottom halves of the face (as in Figure 1B). Because the PC1 eigenimage contained a positive peak at the eyes and a positive peak at the mouth, higher PC1 scores corresponded to a preference to fixate the eyes, leading us to predict a negative correlation between PC1 score and ROI measures of mouth looking and a positive correlation between PC1 and ROI measures of eye looking. As expected, there was a strong negative correlation between PC1 score and mouth looking for both ROI analyses: $r = -0.82$, $p = 10^{-11}$, between PC1 score and percentage of fixation time in mouth ROI (Figure 2E) and $r = -0.68$, $p = 10^{-7}$, between PC1 score and percentage of fixation time in lower face ROI (Figure 2F). Similarly, there was a positive correlation of $r = 0.51$, $p = 0.0007$, between PC1 and percentage of fixation time in the eyes ROI (Figure 2G) and of $r = -0.68$, $p = 10^{-7}$, between PC1 and percentage of fixation time in the upper face ROI.

Stimulus and task differences modulate PC1 face-looking behavior

Having established PC1 score as a measure of mouth looking, we examined how changing the stimulus and task modulated this behavior (Figure 3A). In the first condition, Dynamic_Speech, participants viewed 2-s movie clips of speakers pronouncing syllables; participants indicated the syllable that they heard. In the second condition, Dynamic_Gender, participants viewed the same movie clips but indicated the speaker's gender. In the third condition, Static_Gender, participants viewed still frames from the movie clips for 2 s and indicated the face's gender.

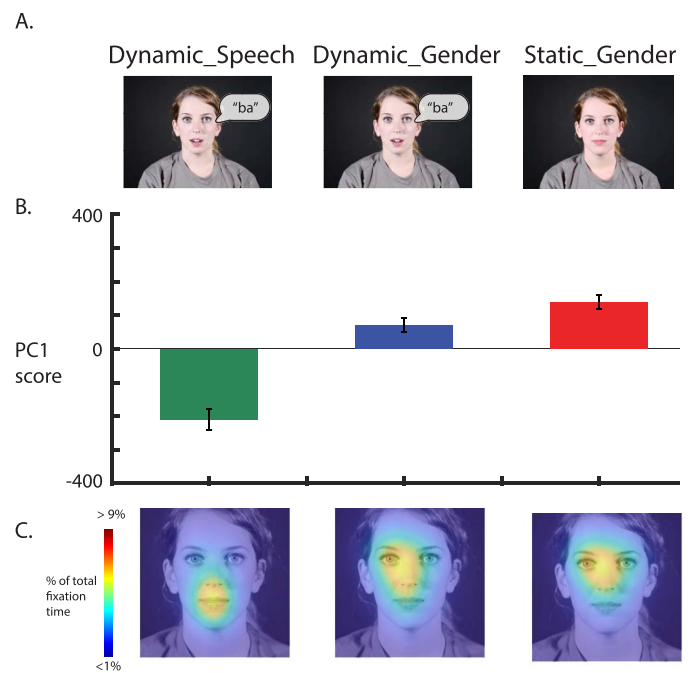


Figure 3. Face-viewing behavior by condition. (A) The experiment contained three stimulus/task conditions. In Dynamic_Speech, participants viewed 2-s audiovisual movies of spoken syllables and identified the syllable. In Dynamic_Gender, participants viewed the same movies and identified the gender of the speaker. In Static_Gender, participants viewed still frames from the movies presented for 2 s without sound and identified the gender. One hundred sixty trials of each condition were presented. (B) Mean PC1 scores for each condition. Error bars show standard error of the mean. (C) Fixation heat maps for each condition, averaged across all participants and exemplars.

We calculated the PC1 score for each participant for each condition (Figure 3B). Averaged across participants, the PC1 score for the Dynamic_Speech condition was strongly negative (−210), indicating more time fixating the mouth of the viewed face. The mean PC1 score for the Dynamic_Gender condition was +71, and the mean PC1 score for the Static_Gender condition was +139, indicating progressively greater time spent fixating the eyes. The Dynamic_Speech heat map showed many fixations to the mouth, and the Dynamic_Gender and Static_Gender heat maps included fixations to both the eye and mouth regions (Figure 3C).

Quantitatively, linear mixed effects (LME) with condition as a fixed effect and exemplar and participant as random effects demonstrated a large and significant effect of condition, parameter estimate = 175, $t(448) = 25$, $p < 10^{-16}$ (see Table 1 for full LME results). The main effect of condition was driven by low PC1 scores for Dynamic_Speech, intermediate PC1 scores for

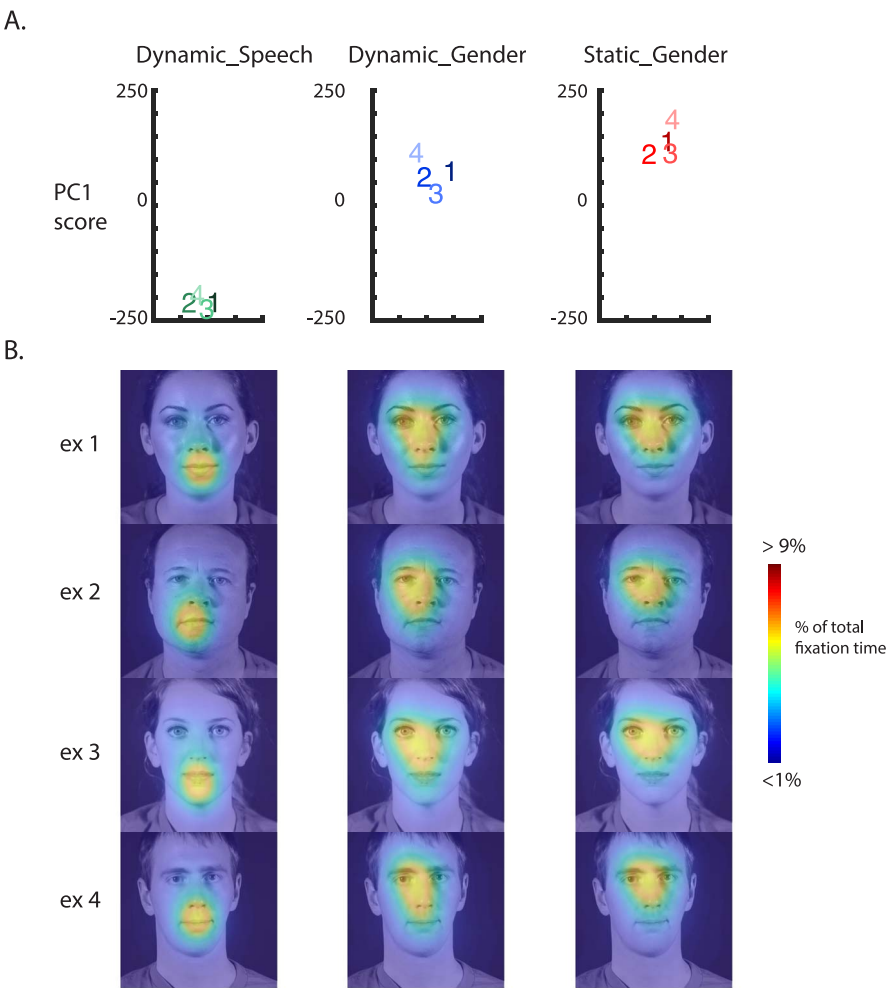


Figure 4. Face-viewing behavior by exemplar. (A) Each digit represents the average score for one exemplar. The digit indicates the exemplar. The y-axis reflects the score. The x-axis is jittered for visibility. (B) Fixation heat maps for each condition and exemplar, averaged across all participants. Heat map is based on all fixations for the given exemplar by condition pair.

Dynamic_Gender, and high PC1 scores for Static_Gender as confirmed by post hoc *t* tests: Dynamic_Speech −210 vs. Dynamic_Gender +71, $t(163) = -18$, $p < 10^{-16}$; Dynamic_Gender +71 vs. Static_Gender +139, $t(163) = -11$, $p < 10^{-16}$; Dynamic_Speech −210 vs. Static_Gender +139, $t(163) = -20$, $p < 10^{-16}$.

Exemplar differences have little effect on PC1

For each condition, there were four different face exemplars. To determine the effect of face exemplars, we calculated the PC1 score for each face exemplar in each condition (Figure 4A). Qualitatively, the raw fixation heat maps for each exemplar in each condition appeared similar (Figure 4B). In the LME model in which exemplar was included as a random effect, it accounted for less variance than participant (Table 1). For the different exemplars, the PC1 values were similar within conditions. For Dynamic_Speech, the

PC1 scores across exemplars ranged from −224 to −194 (mean of −210); for Dynamic_Gender, the range was +27 to +116 (mean of +71); and for Static_Gender, the range was +113 to +188 (mean of +139).

Effects	Estimate	SE	DF	t value	p value
Fixed effects					
Condition	175	7	448	25	<10 ^{−16}
Random effects					
	Variance	SD			
Exemplar	548	24			
Participant	13,233	115			

Table 1. LME model of PC1 score. Notes: Results from an LME model of PC1 score, created with the lme4 package in R (Bates et al., 2015). Condition was included as a fixed effect, and participant and exemplar were included as random effects. Statistical values were calculated according to the Satterthwaite approximation using the lmerTest package (Kuznetsova, Brockhoff, & Christensen, 2017).

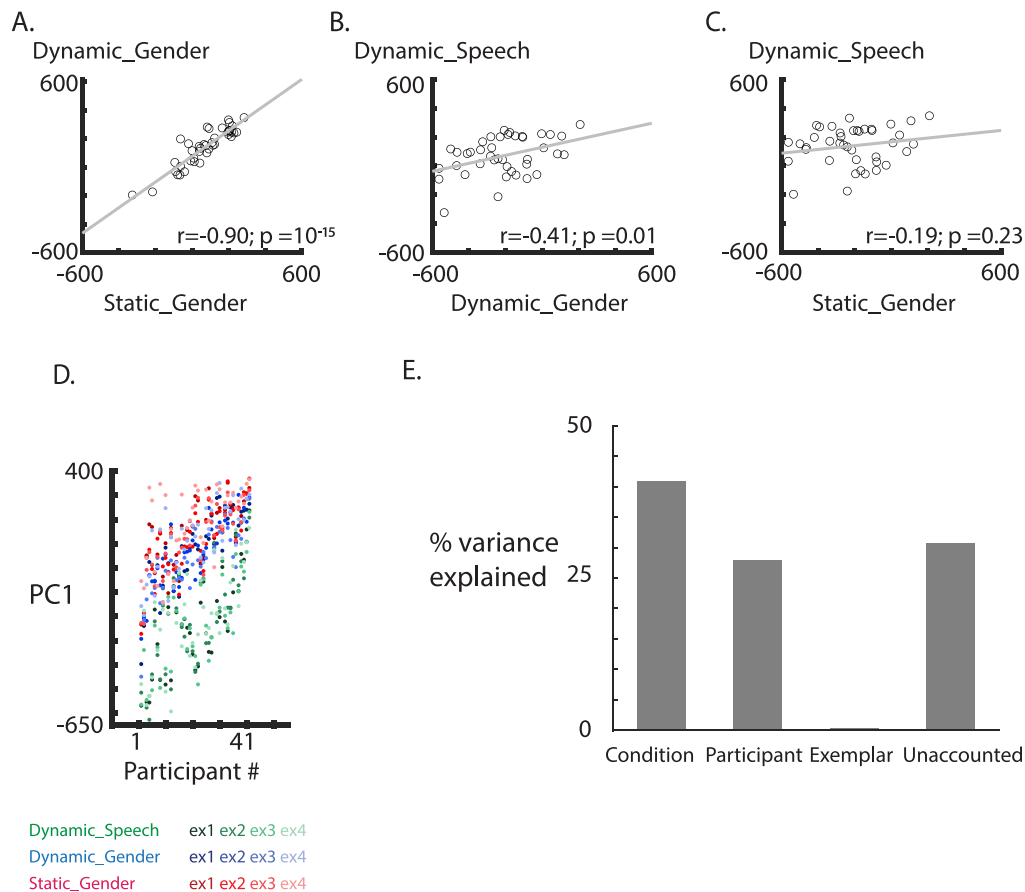


Figure 5. Comparison of PC1 score across task and stimulus. (A) Comparison of PC1 scores for conditions in which task is shared. PC1 scores for the Dynamic_Gender are shown on the y-axis, and scores for Static_Gender are shown on the x-axis. Each symbol corresponds to one participant. Values indicate correlation and significance of least square regression line. (B) Comparison of PC1 scores for conditions in which stimulus type is shared. (C) Comparison of PC1 scores for conditions in which neither task or stimulus type is shared. (D) PC1 scores for all participants, conditions, and exemplars. Each symbol is a subject's score for a given exemplar and condition. PC1 score is shown on the y-axis. Participant is shown on the x-axis, ordered from lowest to highest mean score (does not correspond to participant numbers in Figure 2). Condition is indicated by color of the point (Dynamic_Speech in green, Dynamic_Gender in blue, Static_Gender in red). Exemplar is indicated by the shade of the point with the darkest color indicating exemplar 1 and the lightest color indicating exemplar 4 as in Figure 1B. (E) Variance explained by each factor in the LME model of PC1 score (Table 1) and percentage unaccounted for. Factors are shown on the x-axis, and percentage of variance explained is shown on the y-axis.

Consistency of individual differences in PC1 across stimulus and task

Our PCA analyses showed that face looking varied dramatically across participants (Figure 2) and conditions (Figure 3) but not exemplars (Figure 4). To determine if individual differences were consistent across conditions, we calculated a PC1 score for each participant and condition and correlated them. As shown in Figure 5A, there was a high correlation ($r = 0.90$, $p = 10^{-15}$; Figure 5A) between the PC1 scores of each participant for Dynamic_Gender and Static_Gender. There was a weaker correlation between Dynamic_Gender and Dynamic_Speech ($r = 0.41$, $p = 0.01$; Figure 5B). The weakest correlation was between

Static_Gender and Dynamic_Speech ($r = 0.19$, $p = 0.23$; Figure 5C).

PC1 by condition, participant, and exemplar

To visualize the variance in the data, we plotted all 492 PC1 scores across participants, conditions, and exemplars (Figure 5A). The relative contribution of the different factors was determined by constructing LMEs with and without each factor and determining the difference in variance accounted for (Figure 5B). Condition accounted for 41% of the observed variance, participant for 28%, and exemplar for only 0.4%,

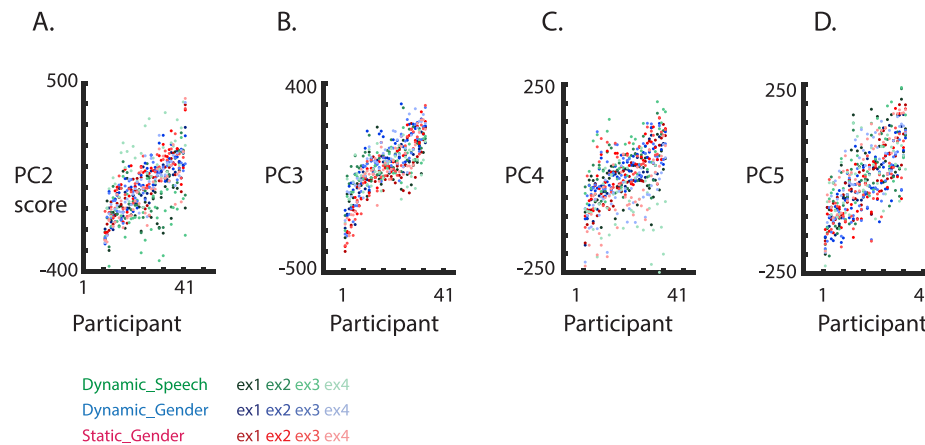


Figure 6. Analysis of PC2-5. (A) PC2 scores for all participants, conditions, and exemplars as in Figure 5D. There is high variance between participants (arranged across the x-axes) but little separation between scores for different conditions and exemplars as indicated by the colors and shades of the points. (B) PC3 scores for all participants, conditions, and exemplars. (C) PC4 scores for all participants, conditions, and exemplars. (D) PC5 scores for all participants, conditions, and exemplars.

leaving 31% of the variance unaccounted for by the model.

The variance in the other PCs is primarily between participants, not conditions or exemplars

We repeated the above analysis on the remaining PCs, plotting all 492 scores for each component across participants, conditions, and exemplars (Figure 6A–D). LMEs for each component were constructed with and without each factor in order to calculate the variance accounted for. Condition explained the largest amount of variance in PC1 (41%) but explained only a small amount of variance in the other components (<2%). In the other components, participant accounted for the largest amount of variance (41%, 63%, 50%, 24% for PC2, PC3, PC4, and PC5, respectively, compared with 28% for PC1). Stimulus exemplar accounted for little variance in any component (<4% for all PCs).

In contrast to PC1, which measured differences along the vertical (bottom to top) axis of the face, PC3 measured differences across the face with a positive peak on the left side of the face and a negative peak on the right side of the face (Figure 1C). One participant with a high PC3 score, participant 34, ($PC3 = +190$) primarily fixated the left side of the image (Figure 7B), and a participant with a low PC3 participant, participant 27 ($PC3 = -207$) primarily fixated the right side of the image (Figure 7C). Participants with PC3 scores near zero, such as participant 38 ($PC3 = +15$) primarily fixated the middle of the face (Figure 7B).

Discussion

During face viewing, humans tend to fixate the eyes and the mouth, but the amount of time spent on each feature varies from individual to individual. These individual differences have been observed for static and dynamic face stimuli in laboratory and real-world viewing conditions (Gurler et al., 2015; Mehoudar et al., 2014; Peterson et al., 2016) and are maintained across test sessions up to 18 months apart (Mehoudar et al., 2014). A separate line of research demonstrates that face-viewing behavior is sensitive to task demands. For instance, participants tend to fixate the mouth when asked to identify whether a face is happy (Pearson, Henderson, Schyns, & Gosselin, 2003; Schurgin et al., 2014) or when attempting to understand noisy speech (Buchan et al., 2008; Rennig et al., in press; Vatikiotis-Bateson et al., 1998), reflecting the additional positive valence and speech information available in the mouth region, respectively, and participants performing a gender task tended to fixate the eye regions of the viewed face (Armann & Bülhoff, 2009; Pearson et al., 2003; Sæther et al., 2009; Schyns et al., 2002), reflecting the additional information about gender available in the eye region.

The interaction between individual differences in face viewing and task demands is poorly understood. Face-viewing behaviors for particular tasks may reflect individual differences in optimal behavior (Peterson & Eckstein, 2012). For instance, an individual with a propensity to fixate the mouth might be confronted with a gender task best performed with fixations to the eye region of the viewed face. Under this circumstance, the individual's propensity might prevail, the task demands might prevail, or there could be some balance between the two.

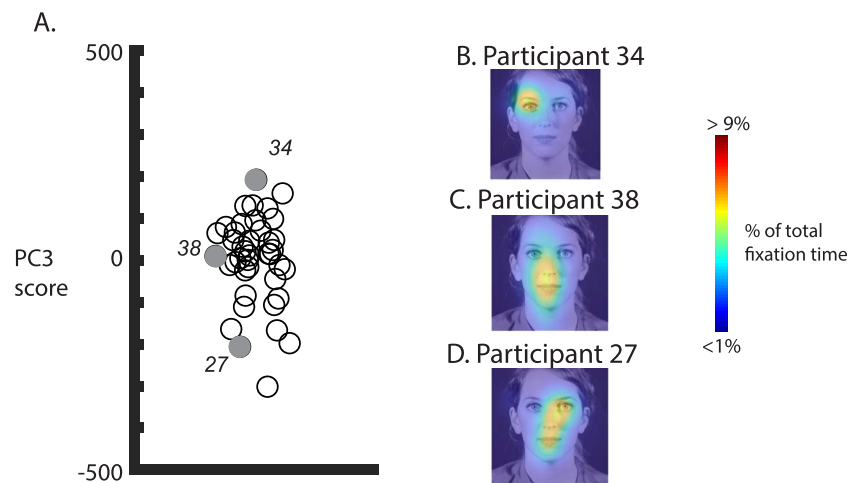


Figure 7. PC3 scores by individual. (A) PC3 scores are shown for each participant, averaged across stimuli, exemplar, and task. Score represented on the y-axis, and the x-axis is jittered for visibility. Representative subjects are shown for high (subject 34, PC3 score: +190), average (subject 38, PC3 score: +15), and low (subject 27, PC3 score: −207); scores are labeled and filled in with gray. (B) Fixation heat map for subject 34, a representative participant with a high PC3 score (+190) averaged across condition and exemplar. Heat map is based on all fixations by the subject. The overlap color indicates the percentage of total fixation time spent on that location with warmer colors indicating more time. (C) Fixation heat map for subject 38, a representative participant with a middle PC3 score (+15) averaged across condition and exemplar. Heat map is based on all fixations by the subject. Color on stimulus image indicates time spent, percentage of total fixation time spent on that location with warmer colors indicating more time. (D) Fixation heat map for subject 27, a representative participant with a low PC3 score (−207) averaged across condition and exemplar. Heat map is based on all fixations by the subject. Color on stimulus image indicates time spent, percentage of total fixation time spent on that location with warmer colors indicating more time.

PCA as a technique for investigating fixation behavior

To investigate this question, we developed a new method for examining individual differences in fixation behavior. Previous studies have drawn ROIs on the viewed face and calculated the fraction of time spent fixating each region. This technique has several limitations. Determining the location, size, and number of ROIs requires a number of arbitrary decisions by the experimenter, each with a large influence on the results of the analysis. ROI analyses also discard data: Fixations that lie outside of any ROI borders are not considered at all in the analysis, and all fixations within an ROI are treated equally. For instance, in the common approach of dividing the face into an upper- and a lower-half ROI, a fixation to the forehead is treated the same as a fixation to the eye. In contrast, our PCA method uses data from all fixations to determine the patterns that account for the most variance across individuals and task conditions. Viewing the fixation eigenimages produced by PCA provides a natural way to visualize data that can be interpreted the same way as the commonly used fixation heat maps.

Eye fixation eigenimages are related to the technique of “eigenfaces” in computer vision in that they allow for the identification of key face regions for classification and recognition (Sirovich & Kirby, 1987; Turk &

Pentland, 1991). Rather than apply PCA to the face images themselves, our technique analyzes participant fixations while viewing the face images.

The different eye-fixation eigenimages identified by PCA can be related to previous studies of face-viewing behavior. PC1, accounting for 42% of total variance, differentiated the eye and mouth regions of the face. There was a large range of PC1 scores across participants, reflecting large individual differences in the propensity to fixate the eyes or the mouth, consistent with previous studies (Gurler et al., 2015; Klin et al., 2002; Mehoudar et al., 2014; Perlman et al., 2009; Peterson & Eckstein, 2013; Peterson et al., 2016; Royer et al., 2018). PC3, accounting for 14% of the variance, differentiated the left and right sides of this face, consistent with previous literature describing left–right asymmetries in face viewing (Butler et al., 2005; Everdell, Marsh, Yurick, Munhall, & Paré, 2007; Mertens, Siegmund, & Grüsser, 1993; Schyns et al., 2002).

Differences between tasks

In the Dynamic_Speech condition, fixations were concentrated around the mouth of the viewed face, and in both Gender conditions, fixations were concentrated around the eye of the viewed face. These findings are

consistent with the literature. During speech perception, mouth fixations are more frequent (Buchan, Paré, & Munhall, 2007; Vo, Smith, Mital, & Henderson, 2012), and during gender tasks, eye fixations predominate (Armann & Bühlhoff, 2009; Sæther et al., 2009; Schyns et al., 2002). This similarity with previous results mitigates a drawback of our study, which is that task order was not counterbalanced: Participants always performed the speech task first and the gender task second. Therefore, if our results were not consistent with this literature, differences between the gender and speech tasks could instead be attributed to fatigue.

Other face-related tasks, such as emotion identification and age judgments, are likely to produce other fixation patterns (Nguyen et al., 2009; Pérez-Moreno, Romero-Ferreiro, & García-Gutiérrez, 2016; Schurgin et al., 2014; Smith, Gosselin, Cottrell, & Schyns, 2010), consistent with a large body of works showing that both task (Armann & Bühlhoff, 2009; Sæther et al., 2009; Schurgin et al., 2014; Schyns et al., 2002) and stimulus (Buchan et al., 2008; Everdell et al., 2007; Masciocchi, Mihalas, Parkhurst, & Niebur, 2009; D. Parkhurst, Law, & Niebur, 2002; D. J. Parkhurst & Niebur, 2004; Pérez-Moreno et al., 2016; Vatikiotis-Bateson et al., 1998) influence face-viewing behavior.

In our experiment, the effect of exemplar was small, accounting for only 0.4% of the total variance in PC1 scores compared with 41% for condition. Across conditions, the Dynamic_Speech and Dynamic_Gender conditions evoked very different patterns of eye fixations even though the stimuli were identical; conversely, the Dynamic_Gender and Static_Gender evoked similar viewing behavior even though the physical stimuli were very different (dynamic audiovisual faces vs. static faces). Taken together, these results suggest that behavioral task is a key driver of fixation behavior with stimulus differences (dynamic vs. static faces or different face exemplars) having less influence.

Our study design replicates previous reports of differences in face-viewing behavior between individuals and between task conditions and allows us to directly compare the importance of these two effects. In our LME model, condition accounted for 41% of the observed variance and participant for 28%, suggesting that both variables make a significant contribution to face-viewing behavior. Of course, this ratio is only valid for the individuals and conditions that we tested. Comparisons between very similar task conditions would result in interindividual differences accounting for a greater fraction of total variance, and comparisons between dissimilar tasks, such as adding auditory noise to the stimulus during the speech task in order to drive fixations toward the mouth (Buchan et al., 2008; Rennig et al., in press; Vatikiotis-Bateson et al., 1998),

would result in a greater contribution of task condition to total variance.

Individual differences in other PCs

In addition to its utility in characterizing individual and task differences in eye vs. mouth viewing, PCA also allows for an examination of other face regions. For instance, PC3 separated fixations made to the left side of the face from fixations made to the right side of the face. Unlike PC1, for which condition explained a large fraction (41%) of the variance, for PC3, condition explained little variance (<2%), showing that neither gender nor speech tasks drove fixations to a particular half of the face. Like PC1, there were large individual differences in PC3 with participant accounting for 63% of the variance (compared with only 28% for PC1). Examining the fixation heat maps for different participants, we found that some participants strongly preferred to fixate the left half of the face, and others strongly preferred to fixate the right half. These results are consistent with previous studies that used ROI analyses to examine individual preferences or fixating the left or right side of dynamic or still talking faces (Buchan et al., 2008; Everdell et al., 2007) in the presence or absence of auditory noise (Buchan et al., 2008).

Generalizability of the method

In the present study, PCA was a useful tool for characterizing individual differences and stimulus/task differences in the eye fixations made while viewing faces. This is useful because face-viewing behavior is linked to other cognitive abilities, such as the ability to understand noisy audiovisual speech (Rennig et al., in press). It seems likely that PCA eye fixation eigenimages would also be useful for other experimental designs, such as characterizing the different fixation patterns made when viewing different facial emotions with joy evoking mouth fixations and anger evoking eye fixations (Schurgin et al., 2014). We would expect significantly different PC1 scores between these two emotional face-viewing conditions. PCA could also be useful for characterizing between-group differences in face-viewing behavior. For instance, people with autism spectrum disorder (ASD) have different face-viewing patterns than typically developing control (Chita-Tegmark, 2016; Frazier et al., 2017; Papagianopoulou, Chitty, Hermens, Hickie, & Lagopoulos, 2014). If people with ASD are less likely to fixate the eyes and more likely to fixate the mouth of the speaker, we would expect them to have significantly lower PC1 scores. PCA could also be useful to characterize

differences in face-viewing fixations across the life span (Franchak, Heeger, Hasson, & Adolph, 2016) or across cultures (Chua, Boland, & Nisbett, 2005; Rayner et al., 2007). PCA analysis of eye fixations could also be useful for nonface objects. There are large individual differences in fixation patterns between participants viewing real-world scenes (de Haas, Iakovidis, Schwarzkopf, & Gegenfurtner, 2019). If, for instance, some participants scan scenes from top to bottom and other participants scan scenes from bottom to top, PCA would be expected to capture this individual difference.

Comparison with other approaches and limitations of the PCA method

A number of data-driven methods to analyze eye movement data have been previously proposed. In particular, clustering is an important alternative to prespecified ROIs (Drusch, Bastien, & Paris, 2014; Göbel & Martin, 2018; Latimer, 1988; Naqshbandi, Gedeon, & Abdulla, 2017; Santella & DeCarlo, 2004). In general terms, these approaches use clustering algorithms, such as *k*-means to group spatially or temporally proximate fixations. Clustering techniques have been used to identify where individuals look when viewing web pages (Drusch et al., 2014) or scenes (Santella & DeCarlo, 2004) as well as to decode the task performed by participants (Naqshbandi et al., 2017). Clustering approaches require additional steps to characterize individual differences. In contrast, for the PCA approach described here, PC scores can be immediately applied as a measure of looking behavior for a particular individual or task.

PCA itself has previously been applied to eye movements in a variety of contexts. In one study, PCA was used to characterize eye movements during scene viewing in order to create individual-specific biometrics; specifics of the eye movement patterns themselves were not considered (Fookes & Sridharan, 2010). In another study, 120 eye movement variables were calculated for each participant, and PCA was used to extract the combinations of variables most able to classify participants as typically developing or with ASD (Ben Mlouka, Martineau, Voicu, Hernandez, & Girault, 2009).

Summary

Fixation eigenimages provide a useful tool for summarizing the large volume of data collected during eye-tracking studies without relying on predefined ROIs. In our data set, the top five PCs were spatially distinct and provided useful information about inter-

individual differences, intertask differences, and the interaction between individual and task differences.

Keywords: *face perception, principal component analysis, eye movements, speech perception, gender classification*

Acknowledgments

This work was supported by the National Institutes of Health (R01NS065395 to MSB) and the Deutsche Forschungsgemeinschaft (RE 3693/1-1 to JR).

Commercial relationships: none.

Corresponding author: Michael S. Beauchamp.

Email: michael.beauchamp@bcm.edu.

Address: Department of Neurosurgery and Core for Advanced MRI, Baylor College of Medicine, Houston, TX.

References

- Armann, R., & Bühlhoff, I. (2009). Gaze behavior in face comparison: The roles of sex, task, and symmetry. *Attention, Perception, & Psychophysics*, 71(5), 1107–1126.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48, <https://doi.org/10.18637/jss.v067.i01>.
- Buchan, J. N., Paré, M., & Munhall, K. G. (2007). Spatial statistics of gaze fixations during dynamic face processing. *Social Neuroscience*, 2(1), 1–13.
- Buchan, J. N., Paré, M., & Munhall, K. G. (2008). The effect of varying talker identity and listening conditions on gaze behavior during audiovisual speech perception. *Brain Research*, 1242, 162–171, <https://doi.org/10.1016/j.brainres.2008.06.083>.
- Butler, S., Gilchrist, I. D., Burt, D. M., Perrett, D. I., Jones, E., & Harvey, M. (2005). Are the perceptual biases found in chimeric face processing reflected in eye-movement patterns? *Neuropsychologia*, 43(1), 52–59.
- Chita-Tegmark, M. (2016). Attention allocation in ASD: A review and meta-analysis of eye-tracking studies. *Review Journal of Autism and Developmental Disorders*, 3(3), 209–223.
- Chua, H. F., Boland, J. E., & Nisbett, R. E. (2005). Cultural variation in eye movements during scene perception. *Proceedings of the National Academy of Sciences, USA*, 102(35), 12629–12633.

- de Haas, B., Iakovidis, A. L., Schwarzkopf, D. S., & Gegenfurtner, K. R. (2019). Individual differences in visual salience vary along semantic dimensions. *Proceedings of the National Academy of Sciences, USA*, 116(24), 11687–11692.
- Drusch, G., Bastien, J. M. C., & Paris, S. (2014). Analysing eye-tracking data: From scanpaths and heatmaps to the dynamic visualisation of areas of interest. *International Conference on Applied Human Factors and Ergonomics*. Krakow, Poland.
- Everdell, I. T., Marsh, H., Yurick, M. D., Munhall, K. G., & Paré, M. (2007). Gaze behaviour in audiovisual speech perception: Asymmetrical distribution of face-directed fixations. *Perception*, 36(10), 1535–1545.
- Fookes, C., & Sridharan, S. (2010). Eigengaze—Covert behavioral biometric exploiting visual attention characteristics. *10th International Conference on Information Science, Signal Processing, and Their Applications*, 320–323.
- Franchak, J. M., Heeger, D. J., Hasson, U., & Adolph, K. E. (2016). Free viewing gaze behavior in infants and adults. *Infancy*, 21(3), 262–287.
- Frazier, T. W., Strauss, M., Klingemier, E. W., Zetzer, E. E., Hardan, A. Y., Eng, C., & Youngstrom, E. A. (2017). A meta-analysis of gaze differences to social and nonsocial information between individuals with and without autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, 56(7), 546–555.
- Göbel, F., & Martin, H. (2018). Unsupervised clustering of eye tracking data. *Spatial Big Data and Machine Learning in GIScience (Workshop at GIScience 2018)*, 25–28.
- Gurler, D., Doyle, N., Walker, E., Magnotti, J., & Beauchamp, M. (2015). A link between individual differences in multisensory speech perception and eye movements. *Attention, Perception, & Psychophysics*, 77(4), 1333–1341.
- Hsiao, J. H. W., & Cottrell, G. (2008). Two fixations suffice in face recognition. *Psychological Science*, 19(10), 998–1006.
- Irwin, J., Avery, T., Brancazio, L., Turcios, J., Ryherd, K., & Landi, N. (2018). Electrophysiological indices of audiovisual speech perception: Beyond the McGurk effect and speech in noise. *Multisensory Research*, 31(1–2), 39–56.
- Kanan, C., Bseiso, D. N. F., Ray, N. A., Hsiao, J. H., & Cottrell, G. W. (2015). Humans have idiosyncratic and task-specific scanpaths for judging faces. *Vision Research*, 108, 67–76.
- Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry*, 59(9), 809–816.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26.
- Latimer, C. R. (1988). Eye-movement data: Cumulative fixation time and cluster analysis. *Behavior Research Methods, Instruments, & Computers*, 20(5), 437–470.
- Masciocchi, C. M., Mihalas, S., Parkhurst, D., & Niebur, E. (2009). Everyone knows what is interesting: Salient locations which should be fixated. *Journal of Vision*, 9(11):25, 1–22, <https://doi.org/10.1167/9.11.25>. [PubMed] [Article]
- Mehoudar, E., Arizpe, J., Baker, C. I., & Yovel, G. (2014). Faces in the eye of the beholder: Unique and stable eye scanning patterns of individual observers. *Journal of Vision*, 14(7):6, 1–11, <https://doi.org/10.1167/14.7.6>. [PubMed] [Article]
- Mertens, I., Siegmund, H., & Grüsser, O. J. (1993). Gaze motor asymmetries in the perception of faces during a memory task. *Neuropsychologia*, 31(9), 989–998.
- Mlouka M. B., Martineau J., Voicu I., Hernandez N., & Girault J. M. (2009) Principal component analysis of eye-tracking data during visual perception of human faces in adults and children with autism. In: S. Vlad, R. V. Ciupa, A. I. Nicu (Eds.), *International Conference on Advancements of Medicine and Health Care through Technology. IFMBE Proceedings*, vol 26. Berlin, Heidelberg: Springer.
- Naqshbandi, K., Gedeon, T., & Abdulla, U. A. (2017). Automatic clustering of eye gaze data for machine learning. *2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016—Conference Proceedings*, 1239–1244.
- Nguyen, H. T., Isaacowitz, D. M., & Rubin, P. A. D. (2009). Age- and fatigue-related markers of human faces: An eye-tracking study. *Ophthalmology*, 116(2), 355–360.
- Papagiannopoulou, E. A., Chitty, K. M., Hermens, D. F., Hickie, I. B., & Lagopoulos, J. (2014). A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. *Social Neuroscience*, 9(6), 610–632.
- Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research*, 42(1), 107–123.
- Parkhurst, D. J., & Niebur, E. (2004). Texture contrast

- attracts overt visual attention in natural scenes. *European Journal of Neuroscience*, 19, 783–789.
- Pearson, A. M., Henderson, J. M., Schyns, P. G., & Gosselin, F. (2003). Task-dependent eye movements during face perception. *Abstracts of the Psychonomic Society*, 8: 84.
- Pérez-Moreno, E., Romero-Ferreiro, V., & García-Gutiérrez, A. (2016). Where to look when looking at faces: Visual scanning is determined by gender, expression and tasks demands. *Psicológica*, 37, 127–150.
- Perlman, S. B., Morris, J. P., Vander Wyk, B. C., Green, S. R., Doyle, J. L., & Pelphrey, K. A. (2009). Individual differences in personality predict how people look at faces. *PLoS One*, 4(6): e5952.
- Peterson, M. F., & Eckstein, M. P. (2012). Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences, USA*, 109(48), E3314–E3323.
- Peterson, M. F., & Eckstein, M. P. (2013). Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. *Psychological Science*, 24(7), 1216–1225.
- Peterson, M. F., Lin, J., Zaun, I., & Kanwisher, N. (2016). Individual differences in face looking behavior generalize from the lab to the world. *Journal of Vision*, 16(7):12, 1–18, <https://doi.org/10.1167/16.7.12>. [PubMed] [Article]
- Rayner, K., Li, X., Williams, C. C., Cave, K. R., & Well, A. D. (2007). Eye movements during information processing tasks: Individual differences and cultural effects. *Vision Research*, 47(21), 2714–2726.
- Rennig, J., & Beauchamp, M. S. (2018). Free viewing of talking faces reveals mouth and eye preferring regions of the human superior temporal sulcus. *NeuroImage*, 183, 25–36.
- Rennig, J., Wegner-Clemens, K., & Beauchamp, M. S. (in press). Face viewing behavior predicts multi-sensory gain during speech perception. *Psychonomic Bulletin and Review*.
- Royer, J., Blais, C., Charbonneau, I., Déry, K., Tardif, J., Duchaine, B., ... Fiset, D. (2018). Greater reliance on the eye region predicts better face recognition ability. *Cognition*, 181, 12–20.
- Sæther, L., Van Belle, W., Laeng, B., Brennen, T., & Øvervoll, M. (2009). Anchoring gaze when categorizing faces' sex: Evidence from eye-tracking data. *Vision Research*, 49(23), 2870–2880.
- Santella, A., & DeCarlo, D. (2004). Robust clustering of eye movement recordings for quantification of visual interest. *Eye Tracking Research and Applications (ETRA) Symposium*.
- Schurigin, M. W., Nelson, J., Iida, S., Ohira, H., Chiao, J. Y., & Franconeri, S. L. (2014). Eye movements during emotion recognition in faces. *Journal of Vision*, 14(13):14, 1–16, [https://doi.org/10.1167/14\(13\):14](https://doi.org/10.1167/14(13):14). [PubMed] [Article]
- Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, 13(5), 402–409.
- Sirovich, L., & Kirby, M. (1987). Low-dimensional procedure for the characterization of human faces. *Journal of the Optical Society of America A*, 4(3), 519–524.
- Smith, M. L., Gosselin, F., Cottrell, G. W., & Schyns, P. G. (2010). Transmitting and decoding facial expressions of emotion. *Journal of Vision*, 4(8):909, <https://doi.org/10.1167/4.8.909>. [Abstract]
- Turk, M. A., & Pentland, A. P. (1991). Face recognition using eigenfaces. *Proceedings 1991 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 586–591.
- Vatikiotis-Bateson, E., Eigsti, I.-M., Yano, S., & Munhall, K. G. (1998). Eye movement of perceivers during audiovisual speech perception. *Perception & Psychophysics*, 60(6), 926–940.
- Vo, M. L.-H., Smith, T. J., Mital, P. K., & Henderson, J. M. (2012). Do the eyes really have it? Dynamic allocation of attention when viewing moving faces. *Journal of Vision*, 12(13):3, 1–14, <https://doi.org/10.1167/12.13.3>. [PubMed] [Article]
- Yarbus, A. L. (1967). *Eye Movements and Vision*. New York: Springer.