More efficient scanning for familiar faces

Jennifer J. Heisz

Department of Psychology, Neuroscience, & Behaviour, McMaster University, Hamilton, Ontario, Canada

David I. Shore

Department of Psychology, Neuroscience, & Behaviour, McMaster University, Hamilton, Ontario, Canada

The present study reveals changes in eye movement patterns as newly learned faces become more familiar. Observers received multiple exposures to newly learned faces over four consecutive days. Recall tasks were performed on all 4 days, and a recognition task was performed on the fourth day. Eye movement behavior was compared across facial exposure and task type. Overall, the eyes were viewed for longer and more often than any other facial region, regardless of face familiarity. As a face became more familiar, observers made fewer fixations during recall and recognition. With increased exposure, observers sampled more from the eyes and sampled less from the nose, mouth, forehead, chin, and cheek regions. Interestingly, this change in scanning behavior was only observed for recall tasks, but not for recognition.

Keywords: eye movements, face perception, learning, familiarity


Introduction

Accurate face perception is critical for establishing and maintaining social relationships since identification of familiar faces provides effective retrieval cues for person-specific information. The importance of face identification has made this a highly specialized skill in typically developing adults providing the ability to establish familiarity for newly learned faces within a single exposure. Familiar and novel faces can produce distinctly different patterns of eye scanning (Althoff & Cohen, 1999). However, little research has investigated how and when these changes in processing occur as a new face becomes familiar. In the present study, we measured eye movements across multiple exposures to newly learned faces.

Our representation of facial identity is highly robust for familiar faces but fallible for newly learned faces. Familiar faces are recognized faster and more accurately than unfamiliar faces (Ellis, Shepherd, & Davies, 1979; Klitzky & Forrest, 1984; Stacey, Walker, & Underwood, 2005). Poor image quality disrupts recognition of unfamiliar faces, but not familiar faces (Burton, Wilson, Cowen, & Bruce, 1999). Moreover, unfamiliar face recognition is disrupted by changes in image context, such as facial expression (Bruce, 1982; Bruce et al., 1999; Patterson & Baddeley, 1977), viewpoint (Bruce, 1982; Bruce et al., 1999; Hill & Bruce, 1996; O’Toole, Edelman, & Bülthoff, 1998; Roberts & Bruce, 1989), and lighting (Hill & Bruce, 1996). In contrast, familiar face recognition prevails despite variable image context (e.g., Bruce, 1982).

These face recognition differences have lead many researchers to theorize that independent neurological processes subserve unfamiliar and familiar face recognition processes (e.g., Benton, 1980). Lesions in different brain regions contributed to deficits in unfamiliar and familiar face recognition (Warrington & James, 1967). In addition, prosopagnosics (e.g. individuals with face recognition deficits) are often able to recognize either unfamiliar or familiar faces, but not both (Malone, Morris, Kay, & Levin, 1982).

Underlying neurological processes can also be inferred through the analysis of eye movements patterns across familiar and unfamiliar faces. Eye movements are thought to provide us with direct insight into brain processes (Just & Capenter, 1980). In highly meaningful scenes, eye gaze is driven by semantic factors (Buswell, 1935; Henderson, Weeks, & Hollingworth, 1999; Yarbus, 1967). Eye movements are highly goal directed and vary depending on task constraints (Henderson, 2003; Loftus & Mackworth, 1978). Moreover, scanning behavior is modified by scene properties, such as global contextual cues (e.g., Torralba, Oliva, Castelhano, & Henderson, 2006).

Indeed, eye movements differ for familiar and unfamiliar faces. O’Donnell and Bruce (2001) demonstrated that observers are sensitive to internal feature changes of familiar face images, but not of unfamiliar face images. Familiar faces were learned via 20-second video clips viewed 18 times, observers were considered familiar only if they were able to correctly identify the name of the face; unfamiliar faces were novel at time of test. During the test phase, observers performed a “same/different” identity-matching task in which two facial images of the same individuals were presented; one of the two images was manipulated, such that eyes or hair were altered. Observers were very proficient at detecting changes in
hair of both familiar and unfamiliar faces, whereas observers were only able to detect changes in the eyes of familiar faces. Similar results were reported in a recent eye-tracking study (Stacey et al., 2005). In a face-matching task, observers looked longer at the internal features of famous faces than at the external features. In contrast, observers looked longer at external features than internal features of unfamiliar faces. Interestingly, this pattern of results was only observed for the matching task; in a familiarity judgment task, observers looked longer at internal features relative to external features of both famous and unfamiliar faces (Stacey et al., 2005). Together these findings suggest that internal features convey more important information about face identity than external features and are particularly useful when the identity of the face is well known.

Differences have also been reported in the way we sample information from famous and unfamiliar faces. In an extensive eye-tracking study, Althoff and Cohen (1999) contrasted various aspects of eye movement behavior while observers viewed famous and unfamiliar faces. Observers performed a fame judgment task in which they had to consider the identity of each facial image. Overall, observers looked longer at the eyes than any other face region and internal features were sampled more than external features regardless of face identity. When viewing famous faces, observers looked more at the eyes than the mouth. However, when viewing unfamiliar faces, observers made more fixations and sampled more regions of the face. These differences emerged early in viewing (as early as the first five fixations) and seem to be available to the face-processing system prior to the recognition decision. Further, the scan pattern for unfamiliar faces appeared to be more idiosyncratic than for famous faces, such that there were more constraint between successive fixations when viewing unfamiliar faces (Althoff & Cohen, 1999). In other words, the probability of making a fixation to current facial region was highly contingent on the location of the immediately preceding fixation for unfamiliar but not famous faces. Moreover, under constrained viewing conditions in which observers were given an overt viewing strategy to follow (e.g., reading strategy: left to right), eye movement behavior resembled that of unfamiliar faces.

Taken together, the extant data on face processing argues for qualitatively different sampling from famous and unfamiliar faces; specifically, famous face recognition is heavily reliant on the eyes and the eye region of the image. Consistent with these results, the eyes and the eye region have been recognized as containing the most informative pixels for facial identification (e.g., Gold, Sekuler, & Bennett, 2004; Vinette, Gosselin, & Schyns, 2004). The faces used in these studies receive thousands of exposures across the experimental sessions. Given the differences between overexposed (e.g., famous facial images) and the processing of unfamiliar faces, there must be a transition in processing style that should be apparent during face learning.

Eye movements have also been shown to play a functional role in learning new faces (Henderson, Williams, & Falk, 2005). Observers in this study learned new faces under a free-viewing condition or a restricted central fixation condition. During the recognition task, performance was better for faces learned with free viewing. Additionally, within the free-viewing condition, eye movements were more restricted during the recognition task than at learning, such that at recognition, observers sampled only from internal features. Interestingly, there were no differences in eye movements reported for new and old faces, suggesting that differences between learning and recognition are task specific, and not due to prior exposure as previously proposed (cf. Althoff & Cohen, 1999).

Some prior exposure does seem to change face-processing abilities as moderately familiar faces appear to be processed in a similar way to famous faces. One study used an identity-matching task in which observers were tested on newly learned faces over three consecutive days (Bonner, Burton, & Bruce, 2003). Their task required observers to judge whether the identities of two different facial images were the same or different; one of the facial images contained the whole face and the other contained either external or internal features only. On the first 2 days, matching of external features with whole faces was performed successfully; however, matching of internal features with whole faces was impaired. By the third day, matching of both internal and external features with whole faces was performed successfully. Sensitivity to internal features is typically seen for famous faces, but not for unfamiliar faces (O’Donnell & Bruce, 2001; Stacey et al., 2005). Based on these data, it appears that prior exposure modifies face-processing sensitivity and these changes occur early in face learning.

**Scope of the present study**

We measured eye movements across multiple exposures of newly learned faces. The experiment was conducted over four consecutive days. On the first 3 days, 10 novel faces were introduced to observers at an individual level (i.e., by name). Recall tests were performed on each day for the faces learned up to that day (i.e., on Day 1 there were 10 faces, on Day 2 there were 20 faces, and on Day 3 there were 30 faces). Feedback was provided after each trial to encourage learning. On the fourth day, observers performed an old/new recognition task followed by a recall task. Eye movement behavior was compared across face exposures to examine changes as a function of prior exposure. In addition, eye movement behavior was compared across the recall and recognition tasks—recall tasks are thought to tap into conscious recollection of
episodic information, whereas recognition tasks are thought to reflect strength of familiarity in the absence of conscious recollection (for a review, see Yonelinas, 2002).

Given the differences between famous and unfamiliar faces reviewed above, we expected to see a qualitative shift in face scanning as the faces became more familiar. Specifically, there should be an increase in the use of information around the eyes as familiarity increases.

Method

Observers

Eleven volunteers (all female, mean age 18.7 years, one left-handed, all right eye-dominant) from the McMaster University community participated in the study. Males were not tested because of their qualitatively different face processing.

All subjects reported normal or corrected-to-normal vision. Informed consent was obtained from each observer. Eligible observers received course credit plus $20.00 for their participation, and the remainder received $40.00 compensation. All procedures complied with the tri-council policy on ethics (Canada) and were approved by the McMaster Ethics Research Board.

Apparatus and stimuli

A Power Mac G4 computer was connected to a ViewSonic Professional Series P220f monitor for presentation of the stimuli using the Psychophysics Toolbox (Version 2.55; Pelli, 1997) running within the MatLab interpreter (Version 5.2.1; Mathworks Inc.). An additional Dell computer was used to collect eye movement data using the EyeLink II system (Version 1.1, 2002).

The face stimuli were 92 black-and-white pictures of Caucasian female faces with neutral expressions. Stimuli were adapted from a larger set of stimulus photographs courtesy of Dr. Daphne Maurer’s Visual Development Lab, Department of Psychology, McMaster University, originally acquired and processed as described in Mondloch, Geldart, Maurer, and Le Grand (2003). All the faces were unknown to the subjects, and the faces were without glasses, jewelry, or other extraneous items. An elliptical mask was used to isolate each face from mid forehead to lower chin (including eyebrows and outer margins of the eyes). The 8-bit (256-level) gray scale images had an average luminance value of approximately 5.5 cd m^-2. Faces were presented at the center of the display. With the constant viewing distance of 80 cm, face stimuli were approximately 7.9 degrees of visual angle high and 5.7 degrees of visual angle wide.

Names were selected from the US Census Bureau-Documentation and Methodology for Frequently Occurring Names in the US, circa 1990; the first 30 names were selected from the list of female names. Names were presented using a computer-generated voice (Mac OS 9.2 “Victoria”).

For each participant, 60 of the 92 faces were chosen and randomly assigned to the 60 names. These 60 faces were then assigned to be introduced on one of the 3 days, or to be used as novel faces on Day 4 (see below). This procedure ensured that unique attributes about any one face could not influence the average results since that one face could appear randomly on any day, or not at all.

Procedure

The experiment was conducted across four consecutive days.

Day 1

The observer was introduced to the different tasks across the 4 days, given a 10-item Edinburgh handedness questionnaire (Oldfield, 1971), tested for eye dominance using the hole-in-the-card technique (cf. Leonards & Scott-Samuel, 2005), and was asked to sign an informed consent form. They were then introduced to the eye-tracking facility and the head-mounted eye tracker.

After calibrating and validating the tracker (which was performed prior to every introduction, recall, and recognition task), the observer was introduced to 10 novel faces by name. Introduction trials were initiated once the observer had achieved central fixation. A computer-generated voice introduced the observer to the novel face stimulus with the statement “Observer’s name, this is face’s name” (e.g., “Sharon, this is Joan”). Immediately after the introduction, a novel face stimulus was presented at the center of the display for 5 seconds. Observers were instructed to learn the face by name for recall throughout the rest of the experiment.

Immediately following, observers were tested on their ability to identify the 10 faces by a recall test. During each recall test trial, a previously learned face was presented at the center of the display until the observer vocally generated a name for the face. In the same room, the experimenter entered the observer’s response into the computer, initiating the removal of the face stimulus from the display. Immediately after removal of the face from the display, the observer received auditory feedback of the true name of the face.

Day 2

The observer was retested on their ability to identify the 10 faces by a recall test. During each recall test trial, a previously learned face was presented at the center of the display until the observer vocally generated a name for the face. In the same room, the experimenter entered the observer’s response into the computer, initiating the removal of the face stimulus from the display. Immediately after removal of the face from the display, the observer received auditory feedback of the true name of the face.
which was the same as the test at the end of the first day. The observer was then introduced to 10 novel faces by name and was immediately tested on their ability to identify the 10 newly learned faces as well as the 10 faces learned the previous day (20 faces total) by a recall test.

**Day 3**

The observer was retesting on their ability to identify the 20 faces learned on Days 1 and 2 by a recall test. Observers’ were then introduced to 10 novel faces by name and immediately tested on their ability to identify the 10 newly learned faces as well as the 20 faces learned on Days 1 and 2 by a recall test (30 faces in total).

**Day 4**

The observer performed an old/new recognition test, which included 60 face stimuli: 30 previously viewed faces (from Days 1 to 3) and 30 novel faces. The observers’ task was to determine whether the face stimuli had been previously learned or not, using a simple button response “x” or “y” on the keyboard pad, which was counter balanced across subjects. A trial was initiated once the observer had achieved central fixation. For each trial, a face was presented at the center of the display until response or 3 seconds, whichever came first. Observers did not receive response feedback but they were asked to rate their confidence level regarding there response (on a percentage scale). This was followed by a recall test on the 30 previously learned faces.

**Data analysis**

Eye-data analysis included all saccades made after the fixation starting the trial. Due to the variable viewing times across subjects and conditions, we analyzed the first 2 seconds of recall trials (99.7% of recall trials lasted 2 seconds or longer) and the first second of recognition trials (99.3% of recognition trials lasted 1 second or longer). Fixations made outside of the facial image were excluded from analysis, which included a total of 0.7% of all fixations. Two percent of all trials were excluded due to calibration issues.

Areas of interest (eyes, nose, mouth, forehead, chin, and cheeks) were defined using a similar template as used in Henderson, Falk, Minut, Dyer, and Mahadevan (2000) in which non-overlapping rectangular sections enclosed the feature of interest. Three area-of-interest templates were used across the 60 images to accommodate low, medium, and high feature placement. All three templates covered the same area, differing mainly in the size of the forehead and the chin regions. The medium feature template was used for 75% of the stimuli; both high and low feature templates were used for 13% of the stimuli. Importantly, faces were randomly assigned to the different conditions between subjects.

Mean fixation counts and mean fixation durations were computed from the eye movement data. In addition, proportion of fixations and proportion dwell time were computed for each area of interest. Eye movement analysis was based on trials with correct and incorrect responses; separate analyses of correct and incorrect trials yielded similar results. Trials in both recognition and recall tasks were defined according to the number of prior exposures, which allows for a direct comparison between tasks. In addition, recall tasks performed on the same day were collapsed; t tests between same day recall tasks revealed no significant differences.

**Results**

**Recognition task**

Figure 1A represents d' scores for performance in the old/new recognition for faces with 2, 4, and 6 prior exposures. Performance accuracy improved with number of exposures to a particular face. This observation was supported by a significant main effect of exposure, \( F(2, 20) = 16.752, p < 0.001 \), and a significant linear trend, \( F(1, 10) = 22.965, p < 0.01 \).

Figure 1B represents mean fixation count in the old/new recognition test for faces with 0, 2, 4, and 6 prior exposures. Mean fixation count decreased with number of prior exposures. This observation was supported by a significant main effect of exposure, \( F(3, 30) = 8.777, p < 0.01 \), and a significant linear trend, \( F(1, 10) = 10.934, p < 0.01 \). Mean fixation duration increased with number of prior exposure but this effect was not significant, \( F(3, 30) = 1.984, p = n.s. \).

Figures 2A and 2B illustrate proportion fixations and proportion dwell time, respectively, at the eyes, nose, mouth, and other regions (forehead, chin, and cheeks combined) of facial images with 0, 2, 4, and 6 prior exposures. Overall, observers looked longer and more often at the eyes than any other region of the face. In addition, the nose and the mouth regions had a higher proportion of fixations and a greater proportion dwell time than the other regions. These observations were supported by a significant main effect of feature; proportion fixation count: \( F(3, 30) = 23.065, p < 0.001 \); proportion dwell time: \( F(3, 30) = 19.741, p < 0.001 \). The effect of prior exposure and the interaction between exposure and feature were not significant, \( F < 1 \).

**Recall tasks**

Figure 1C represents mean performance accuracy in the recall task for faces with 1, 2 (or 3), 4 (or 5), and 7 prior
exposures. The double digits stand for the two recall tests performed on the same day; data from these tests were collapsed. Mean accuracy increased with number of prior exposures. This observation was supported by a significant main effect of exposure, $F(3, 30) = 13.469, p < 0.001$, and a significant linear trend, $F(1, 10) = 25.339, p < 0.001$.

Figure 1D represents mean fixation count in recall tests for faces with 1, 2 (or 3), 4 (or 5), and 7 prior exposures. Mean fixation count decreased with prior exposure. This observation was supported by a significant main effect of exposure, $F(3, 30) = 6.908, p < 0.01$, and a significant linear trend, $F(1, 10) = 10.177, p < 0.05$. Mean fixation duration increased with number of prior exposure but this effect was not significant, $F(3, 30) = 1.978, p = ns$.

Figures 2C and 2D illustrate proportion fixation count and proportion dwell time, respectively, at the eyes, nose, mouth, and other regions of facial images with 1, 2 (or 3), 4 (or 5), and 7 prior exposures. Overall, observers looked longer and more often at the eyes than any other region of the face. In addition, the nose and the mouth regions had a higher proportion of fixations and longer dwell time than the other regions. These observations were supported by a significant main effect of feature; proportion fixation count: $F(3, 30) = 91.618, p < 0.001$; proportion dwell time: $F(3, 30) = 66.470, p < 0.001$. In addition, proportion fixations and proportion dwell time at eyes regions increased with prior exposures, whereas proportion fixation and proportion dwell time at nose, mouth, and other regions decreased with prior exposures. These observations were supported by a significant linear trend of Feature $\times$ Exposure for proportion dwell time, $F(1, 10) = 7.421, p < 0.05$, and a significant of linear trend of Feature $\times$ Exposure for proportion fixation, $F(1, 10) = 8.372, p < 0.05$.

Figure 3 illustrates the change in fixation pattern that occurs with increased exposure: with only one prior exposure to a face image, the observer samples information from the entire face, whereas after seven exposures to
a face the observer bases their judgment on information from the eye region. To further investigate this pattern of data, we performed additional one-way repeated measures ANOVA on a difference score comparing the last exposure minus the first exposure for proportion of dwell time and proportion of fixations with the factor of feature (eyes, nose, mouth, other). The analysis was only performed for faces with more than four exposures (i.e., faces learned on Days 1 and 2). Figure 4 illustrates the difference scores for last minus first exposure for proportion fixation count and proportion dwell time at the eyes, nose, mouth, and other regions of the facial image. For proportion dwell time, there was a significant main effect of feature, $F(3, 30) = 5.370, p < 0.05$, such that proportion of dwell time to the eyes increased from first to last exposure, whereas proportion of dwell time to nose, mouth, and other regions decreased from first to last exposure, $t(10) = 3.355, p < 0.01$, for eyes versus other regions. The same pattern of results was observed for proportion fixations, main effect of feature, $F(3, 30) = 6.428, p < 0.05$, such that proportion of fixations to eyes increased from first to last exposure, whereas proportion of fixations to nose, mouth, and other regions decreased from first to last exposure, $t(10) = 3.536, p < 0.01$, for eyes versus other regions.

**Discussion**

The present study examined how eye movements change as newly learned faces become more familiar. Over four consecutive days, observers were exposed to newly learned faces. Recall tasks were performed on all 4 days and on the fourth day observers performed an old/new recognition task. Eye movement behavior was compared across face exposures and task type. Overall, eye movements changed as a function of face familiarity. In both recall and recognition tasks, performance accuracy...
improved with exposure. In turn, mean fixation count decreased with exposure. The eyes and the eye region were viewed for longer and more often than any other region of the facial image, regardless of face familiarity. But as faces became familiar, observers sampled more from the eyes and sampled less from the nose, mouth, forehead, chin, and cheek regions (see Figure 4). Interestingly, this observation was seen for recall tasks only and not for the recognition task.

This was the first study to measure eye movement changes across exposures as new faces become familiar. With increased exposure to a particular face, observers required fewer fixations for identification. In addition, as a facial image became more familiar, observers changed the way they sampled a facial image; when viewing the image for the first time, observers sampled information from the whole face, and after multiple exposures to the same image, observers sampled information mostly from the eyes and the eye region. Based on these data, it appears that identification of unfamiliar and familiar faces requires different processing strategies; more specifically, that whole face processing is required for unfamiliar face identification whereas part-based face processing is sufficient for familiar face identification. Indeed, it has been demonstrated that eye movements and therefore processing strategy differ between famous and unfamiliar faces (Althoff & Cohen, 1999; Stacey et al., 2005), but in the present study, we demonstrate these processing differences in the same facial image. These results shed light on the
current debate in the face-processing literature regarding whole versus part-based face processing (cf. Maurer, Le Grand, & Mondlock, 2002). Accordingly, these differences should be considered when comparing face processing of single versus multiple trial exposures.

In addition, we have demonstrated that eye movement patterns change as a function of prior exposure. After multiple exposures to a newly learned face, observers required fewer fixations for identification; they sampled less from the whole face and focused more on particular areas, such as eyes. Comparable results were reported in a study contrasting eye movements for famous versus unfamiliar faces (Althoff & Cohen, 1999). Together, these findings contradict the hypothesis that eye movement differences between learning and test of an old/new recognition tasks are merely due to task demands. Henderson et al. (2005) concluded this after observing eye movement differences between learning and test phases of an old/new recognition task for old faces, but not for new faces. However, it is important to note that all faces used by Henderson et al. were initially unfamiliar and faces that were considered “old” received only one prior exposure. In the present study, we observed similar null effects when comparing faces with 0 and 2 prior exposures (see Figure 1B); differences in eye movement behavior were not observed until the fourth exposure. These findings indicate that eye movements change as a result of prior exposure, and these changes occur gradually.

Different effects of exposure were observed in recognition and recall tasks. In the last facial exposure, relative to the first, observers sampled more from the eyes and sampled less from the nose, mouth, and other regions (see Figures 2 and 3). This pattern of results was observed for the recall tasks, but not for the recognition task. Given the distinct nature of recall and recognition tasks, the findings are interesting but not surprising (for a review, see Yonelinas, 2002). Recall tasks capture conscious recollections of episodic information, whereas recognition tasks measure strength of the memory trace in the absence of conscious recollection. Related differences between individual identification tasks and more shallow tasks have been observed in the effect of face inversion and direction of lighting (Enns & Shore, 1997; McMullen, Shore, & Henderson, 2000). In general, familiarity judgments are faster than recollection (Yonelinas & Jacoby, 1994), and distinct neural regions support these tasks (Ranganath et al., 2003). Moreover, judgments of familiarity seem to reflect automatic processes, whereas recollection reflects more controlled processes (Jacoby, 1991; Toth, 1996). These processing differences may be manifested in differences seen in observers’ scanning behavior. When recalling information about familiar individuals perhaps, we seek out the most informative facial regions (e.g., eyes) to provide an effective retrieval cue. In contrast, recognition judgments of a familiar individual might be made by the re-instantiation of the whole face image via quick and automatic scanning.

In summary, we have demonstrated that eye movements change as a function of prior exposure. As a face becomes more familiar, observers made fewer fixations and sampled more information from the eyes. These eye movement changes occurred gradually and seem to be more apparent in tasks that require overt recollection as opposed to recognition. Future research should focus on understanding the relevance of these eye movement changes to face learning. One approach may be to compare eye movement behavior during face learning of populations with face processing deficits, such as individuals with autism and prosopagnosia.

Acknowledgments

The authors would like to thank Dr. Daphne Maurer and the Vision Development Lab at McMaster University’s Department of Psychology, Neuroscience & Behaviour, for use of their face photographs, from which the stimuli were constructed. The authors would also like to thank Craig Wilson for his programming expertise. This Research was supported by the Natural Sciences and Engineering Research Council of Canada through a Canadian Graduate Scholarship-M to JJH and a Discovery Grant to DIS. Further support was supplied through a Premier’s Research Excellence Award and a CFI/OIT New Opportunities Award to DIS.

Commercial relationships: none.
Corresponding author: Dr. David I. Shore.
Email: dshore@mcmaster.ca.
Address: Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, Ontario, L8S 4K1, Canada.

References


Downloaded From: http://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/932849/ on 01/06/2018


