Two fixations suffice in face recognition

Draft version of paper accepted for publication in *Psychological Science*

Janet Hui-wen Hsiao & Garrison Cottrell
Department of Computer Science & Engineering, University of California, San Diego, USA

Address for correspondence:
Janet Hui-wen Hsiao
Department of Computer Science & Engineering
University of California, San Diego
9500 Gillman Drive # 0404, La Jolla, CA 92093-0404, USA
Phone: 1-858-534-8603
Email: jhsiao@cs.ucsd.edu

Abstract

In visual word recognition, it is well-known that there exist preferred landing positions for eye fixations. However, the existence of preferred landing positions in face recognition is less well established. It is also unknown how many fixations are *required* to recognize a face. To investigate these questions, we recorded eye movements during face recognition. Subjects were allowed a variable number of fixations before the stimulus was masked during an otherwise standard face recognition task. We find that optimal recognition performance is achieved with two fixations – performance does not improve with additional fixations. The first fixation is just to the left of the center of nose, and the second on the center of the nose, suggesting preferred landing positions for face recognition. Furthermore, the fixations made during face learning are different in location from and more variable in duration than at face recognition time, suggesting that different strategies are used.

Two fixations suffice in face recognition

Introduction

In research on reading, the existence of preferred landing positions (PLP, Rayner, 1979) in sentence reading and optimal viewing positions (OVP, O'Regan et al., 1984) in isolated word recognition has been consistently reported. PLPs are where we fixate our eyes most often during reading; OVPs are where the initial fixation is directed to when the best recognition performance for isolated words is obtained. For English words, both the PLPs and OVPs have been shown to be to the left of the word center and argued to reflect the interplay of different variables, including the visual acuity difference between fovea and periphery, the information profile of words, perceptual learning, and hemispheric asymmetry (Brysbaert & Nazir, 2005).

Like reading, face recognition is an over-learned skill and is learned even earlier in life. However, it remains unclear whether PLPs or OVPs also exist in face recognition. Compared with words, faces are much larger in size, and thus more fixations may be required; nevertheless, eye movements may be thought to be unnecessary since faces are processed holistically (e.g., Farah et al., 1995). Yet, studies have suggested that face recognition performance is related to eye movement behavior (e.g., Althoff & Cohen, 1999). Henderson, Williams, and Falk (2005) restricted fixations during face learning and showed that eye movements during face recognition do not change due to this restriction. They concluded that eye movements have functional roles and are not just a recapitulation of those produced during learning (cf. Mäntylä & Holm, 2006). However, it remains unclear what the functional roles are: are all the fixations functionally significant in terms of their contribution to the recognition performance? Specifically, how many fixations do we really need to recognize a face, and where are they located?

Here we address these questions by manipulating the number of fixations that participants are allowed to make during face recognition. In contrast to Henderson et al.'s study (2005), in which the fixation during face learning was restricted to the center of the face, participants are able to move their eyes freely. During recognition, we restrict the maximum number of fixations to be one, two, three, or no restriction; the face will be covered by a mask after they have reached the maximum number. Thus, we are able to examine the influence of the number of fixations on face recognition performance with participants' natural eye movements. Also, while Henderson et al. (2005) analyzed fixation regions using the total fixation time and the number of trials with at least one fixation in the region, we analyze exact location and duration of the fixations. Previous studies show that the eyes, the nose, and the mouth are where participants look most often during face recognition (e.g., Barton et al., 2006). Studies using the Bubbles procedure show that the most diagnostic features for face identification are the eyes (e.g., Schyns, Bonnar, & Gosselin, 2002); thus, one would predict that three to four fixations may be required to recognize a face, and that the first two will be on the eyes. Recent computational models of face recognition have incorporated eye fixations (e.g., Lacroix et al., 2006). The NIMBLE model (Barrington, Marks, & Cottrell, 2007) achieves above chance performance with a single fixation (ROC area ~ 0.6; ROC area at chance level is 0.5), and performance improves and then levels off with an increasing number of fixations. We thus predict that participants will be above chance with a single fixation, and have better performance when more fixations are allowed, up to some limit.

Methods

Materials
The materials consisted of 16 male and 16 female 296×240 pixel grayscale, front-view face images, taken from the FERET database (Phillips et al., 2000). Another set of 16 male and 16 female face images were used as foils. All were Caucasians with neutral expressions and no facial hair or glasses. We aligned the faces without removing configural information by rotating and scaling the faces so that the triangle defined by the eyes and mouth was at a minimum sum squared distance to a predefined triangle (Zhang & Cottrell, 2006). The image size on the screen was 6.6 cm wide, and participants' viewing distance was 47 cm; thus, the face spanned about eight degrees of visual angle, equivalent to the size of a real face under the viewing distance of 100cm (about the distance between two persons during a normal conversation; cf. Henderson et al., 2005). Approximately one eye on the face may be foveated at a time.

Participants
We recruited eight male and eight female Caucasian UCSD students (mean age 22 years 9 months) for the study. They were all right handed according to the Edinburgh handedness inventory (Oldfield, 1971); all have normal or corrected to normal vision. They participated for course credit or received a small honorarium for their participation.

Apparatus
Eye movements were recorded with an EyeLink II eye tracker. Binocular vision was used; the data of the eye with less calibration error was used for analysis. The tracking mode was pupil only with a sample rate 500 Hz. A chinrest was used to reduce head movements. In data acquisition, saccade motion threshold was 0.1 degree of visual angle; saccade acceleration threshold was 8000 degree / square second; saccade velocity threshold was 30 degree / second. These are the EyeLink II defaults for cognitive research.

Design
The experiment consisted of a study and a test phase. In the study phase, participants saw the 32 faces, one at a time, for three seconds in random order. In the test phase, they saw the same 32 faces and 32 foils one at a time and were asked to recall the faces they saw at the study phase by pressing "YES" and "NO" buttons within three seconds. There was a 20-minute visual search task between phases that did not contain any face-like images.

The design had one independent variable: number of permissible fixations at test (one, two, three, and no restriction). The dependent variable was the discrimination performance measured by A', a bias-free nonparametric measure of sensitivity. The value of A' varies between 0.5 to 1.0; higher A' indicates better discrimination. Unlike d', A' can be computed when cells with zero responses are present[1]. In the eye movement data analysis, the independent variables were phase (study vs. test) and fixation (first, second, and third)[2]; the dependent variable was fixation location and duration. During the test phase, the 32 faces were divided into the four fixation conditions evenly, counterbalanced through a Latin square design. In order to counterbalance possible differences between the two sides of the faces, half of the participants were tested with mirror images of the original stimuli.

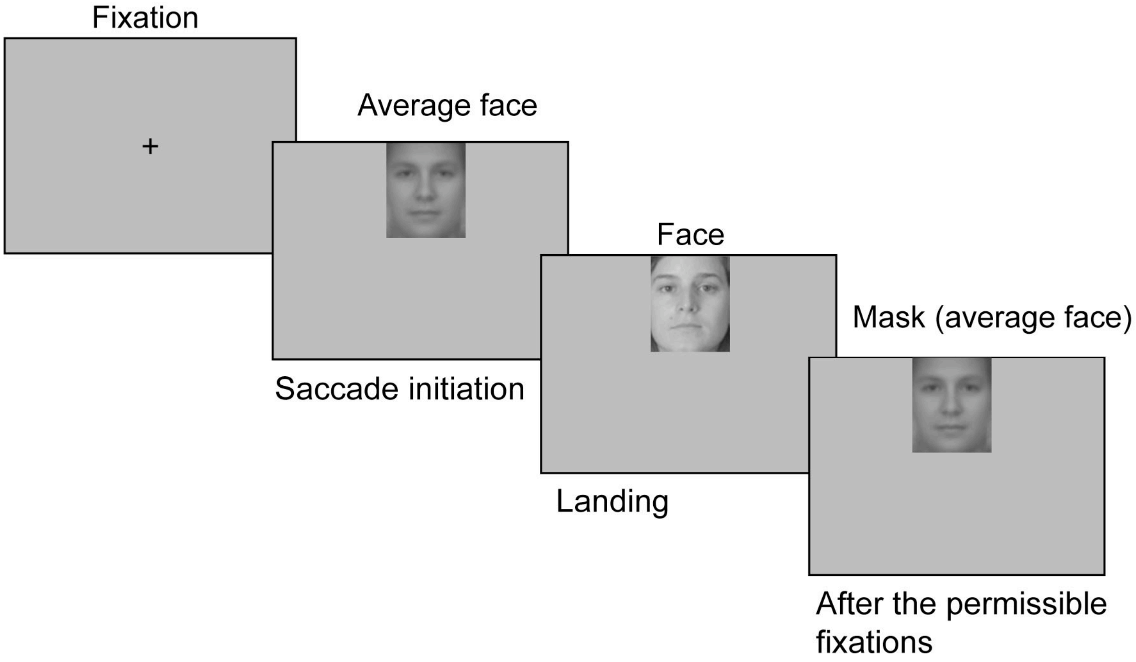

Figure 1. Experimental procedure. After the fixation cross, the average face image was presented either on the top or the bottom of the screen, and was replaced by the target image as soon as a saccade was detected from the fixation cross towards the image. During the test phase, when a restriction was imposed, the image stayed until either the participant's eyes moved awazy from the last permissible fixation, or the response, or the end of three seconds. The image was covered by the average face as a mask after the permissible fixations.

Procedure

The standard nine-point Eyelink II calibration procedure was administered in the beginning of both phases, and was repeated whenever the drift correction error was larger than one degree of visual angle. Each trial started with a solid circle at the center of the screen. Participants were asked to accurately fixate the circle for drift correction. The circle was then replaced by a fixation cross, which stayed for 500 ms or until the participant accurately fixated it. The average face (the pixel-wise average of all of the faces in the materials) was then presented either on the top or bottom of the screen, and was replaced by the target image as soon as a saccade was detected from the cross towards the image (Figure 1; the refresh rate of the monitor was 120 Hz). Thus participants only received reliable face identity information after the initial saccade. The initial saccade direction (up or down) was counterbalanced for each image across participants.

During the study phase, the target image stayed on the screen for three seconds. During the test phase, the image stayed until either the participant's eyes moved away from the last permissible fixation (if a restriction was imposed), or the response, or the end of three seconds. The image was covered by the average face as a mask after the permissible fixations; the mask stayed until the response (Figure 1). They were asked to respond as fast and accurately as possible. They were not told about the association between the mask and the number of fixations they made. The fixation conditions were randomized, so that even if they were aware of it, they were not able to anticipate the fixation condition in each trial.

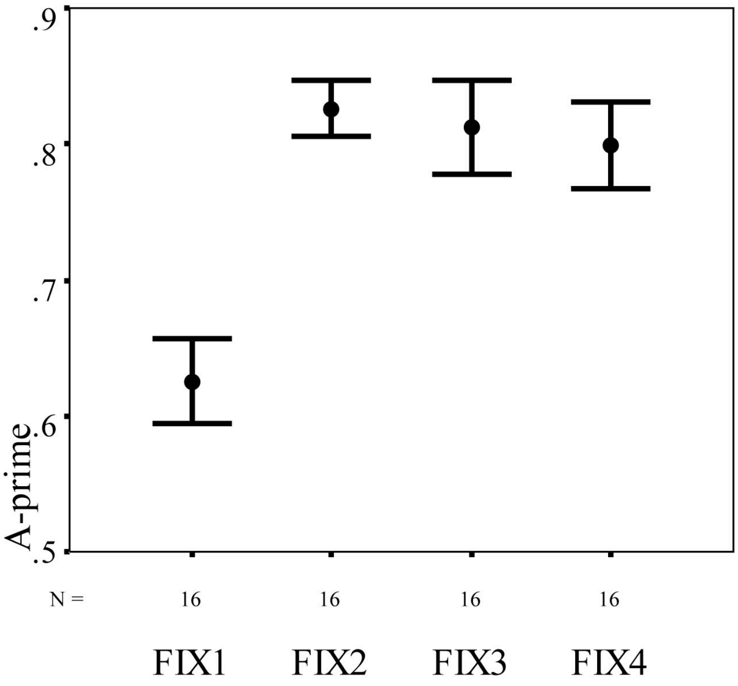

Figure 2. Participants' discrimination performance measured by A' in different fixation restriction conditions: one fixation (FIX1), two fixations (FIX2), three fixations (FIX3), and the free viewing condition (FIX 4+). Error bars show standard errors.

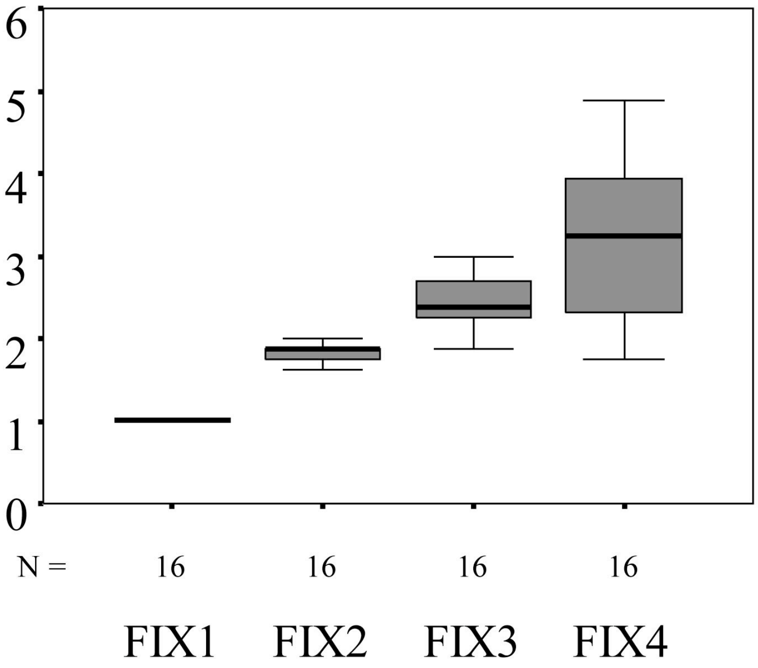

Figure 3. Boxplot of the average number of informative fixations, that is, fixations that landed on the face stimulus before the face image was covered by the average face, in different fixation restriction conditions before the response. Shaded area comprises 50% of the distribution.

Results

Repeated measures ANOVA was conducted for the analyses. The recognition performance measured by A' showed a fixation effect ($F(3, 45) = 11.722$, $p < 0.001$, $p_{rep} = 0.999$, $\eta_p^2 = 0.439$; Figure 2): the A' in the two-fixation condition was significantly better than the one-fixation condition ($F(1, 15) = 44.435$, $p < 0.001$, $p_{rep} = 0.999$, $\eta_p^2 = 0.748$); in contrast, A's in the two, three, and no-restriction conditions were not significantly different from each other. In the one-fixation condition, the participants had above-chance performance ($F(1, 15) = 16.029$, $p = 0.001$, $p_{rep} = 0.986$, $\eta_p^2 = 0.517$; the average A' was 0.63). Figure 3 shows a boxplot of the average number of informative fixations (i.e., fixations that landed on the face before the mask and the response; participants were allowed to respond before they reached the fixation limit) in different fixation conditions. Since the participants required at least one fixation to actually see the face, in the one-fixation condition, all participants made exactly one fixation. The variability in other conditions reflects that occasionally participants did not use all the fixations available to them. When the restriction was two, participants made 1.81 fixations on average; in contrast, without restriction, the average was 3.28. Nevertheless, their performance did not improve with more than 1.81 fixations on average.

The eye movements in the x-direction showed that the first two fixations were significantly different from each other ($F(1, 15) = 5.145$, $p = 0.039$, $p_{rep} = 0.894$, $\eta_p^2 = 0.255$); the first fixation was significantly to the left of the center ($\bar{x} = 113.3$ (3.2); x at the center = 120.5 since the image width was 240 in pixels; $F(1, 15) = 5.208$, $p = 0.037$, $p_{rep} = 0.897$, $\eta_p^2 = 0.258$), whereas the second fixation was not significantly away from the center ($\bar{x} = 118.7$ (3.2)). The first fixation during the study phase also had a leftward tendency ($\bar{x} = 115.6$ (2.6); $F(1, 15) = 3.511$, $p = 0.081$, $p_{rep} = 0.839$, $\eta_p^2 = 0.190$).

In y-direction, there was a phase effect ($F(1, 15) = 5.288$, $p = 0.036$, $p_{rep} = 0.898$, $\eta_p^2 = 0.261$): the fixations at test were lower in location than those during the study phase (Figure 4). In addition, during the study phase, the difference between the three fixations was significant ($F(2, 30) = 3.896$, $p = 0.040$, $p_{rep} = 0.892$, $\eta_p^2 = 0.206$), whereas during the test phase, the difference was not significant ($F = 2.494$); there was also a significant linear trend ($F(1, 15) = 7.185$, $p = 0.017$, $p_{rep} = 0.933$, $\eta_p^2 = 0.324$) moving upwards from the first to the third fixation. This result suggests slightly different eye movement strategies adopted between the two phases. In a separate analysis, we examined all (informative) fixations from all participants without averaging them by subject using a linear mixed model, and the same effects (in both x- and y-directions) held (Figure 5)[3].

The fixation duration data showed an interaction between phase and fixation ($F(2, 30) = 13.292$, $p < 0.001$, $p_{rep} = 0.994$, $\eta_p^2 = 0.470$): the fixation effect was significant during the study phase ($F(2, 30) = 21.940$, $p < 0.001$, $p_{rep} = 0.999$, $\eta_p^2 = 0.594$), but not at test (Figure 6). During the study phase, participants made a short fixation first and then gradually increased the duration for the subsequent fixations, whereas during the test phase, there was no significant difference among them. This again suggests different strategies adopted during the two phases.

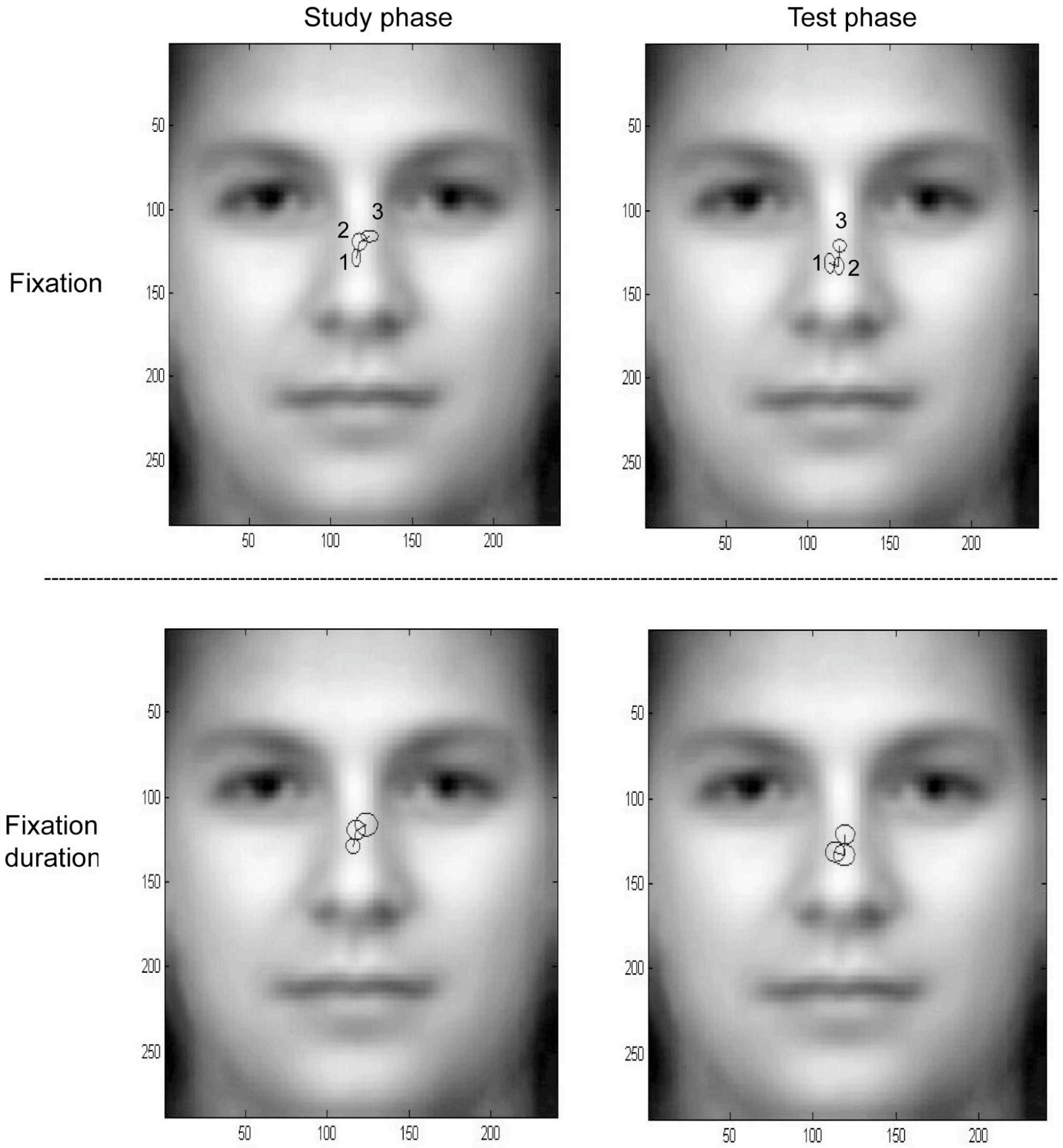

Figure 4.Top: Scan paths of the first three fixations during the study and the test phases; the radius of the ellipses on the faces show standard errors. Bottom: The radius of the circles show fixation duration (1 pixel = 50 ms).

The results showed that participants had better performance when given two fixations compared with one; however, there is a possibility that this improvement is simply due to longer viewing time in the two-fixation condition. To examine whether this is the case, we conducted a follow-up experiment comparing one- and two-fixation conditions given the same total fixation duration[4]. In each trial the total fixation duration was fixed to be 610 ms, which is the sum of the average duration of the first two fixations in the previous experiment. In the one-fixation condition, after the participants made the first fixation onto the face image, the image moved with their gaze (i.e., gaze contingent display); thus, they could only keep looking at the same location on the image as their first fixation. In the two fixation condition, the image became gaze contingent after a second fixation. They were only told that the image may move during the

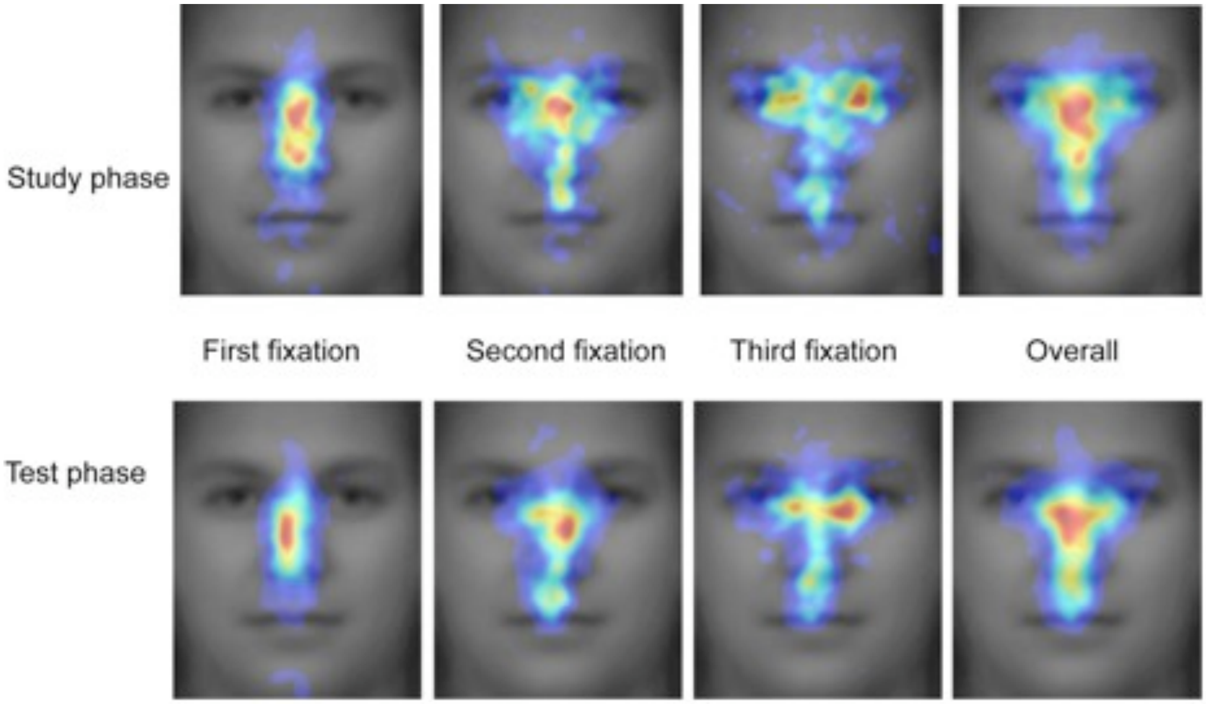

Figure 5. Distribution of the first three fixations in all trials and overall fixations from all subjects during the study and test phases. On each fixation a Gaussian distribution with a standard deviation equal to eight pixels is applied to smooth the distribution.

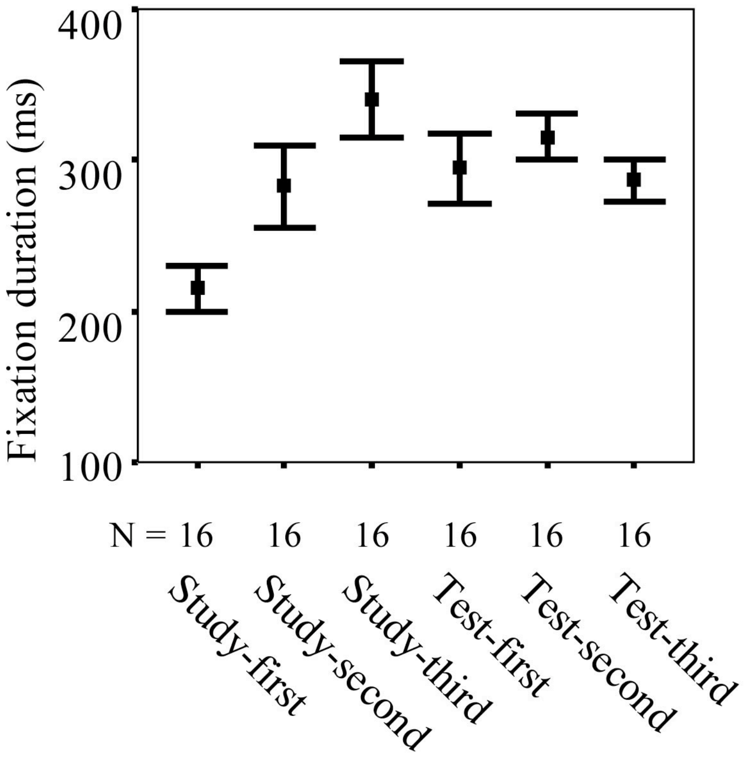

Figure 6. Duration of the first three fixations during the study and the test phases. Error bars show standard errors.

presentation, and we confirmed individually after the experiment that none of them were aware of the gaze contingent design during the experiment.

The results showed a fixation condition effect in A' ($F(1, 15) = 6.847$, $p < 0.05$, $p_{rep} = 0.929$, $\eta_p^2 = 0.313$; Figure 7): the A' in the two-fixation condition was significantly better than the one-fixation condition. This shows that given the same total fixation duration, the participants had better performance when they were allowed to make a second fixation to obtain information from a different location compared with the condition in which they could only look at the same location. This suggests that in the previous experiment, the advantage of the two-fixation condition was not purely due to the longer total fixation duration.

Discussion

We examined the influence of the number of eye fixations on face recognition performance. We showed that when the number of permissible fixations was one, the participants had above-chance performance, suggesting that we are able to recognize a face with one fixation. They had better performance when two fixations were allowed; there was no further performance improvement with more than two fixations, suggesting that two fixations suffice in face recognition.

The first two fixations are around the center of nose, with the first fixation being slightly to the left. Note that this result is different from our predictions according to the existing literature. A major difference between the current study and the existing literature is that previous studies start a trial from the center of the face, and hence the first saccade is usually

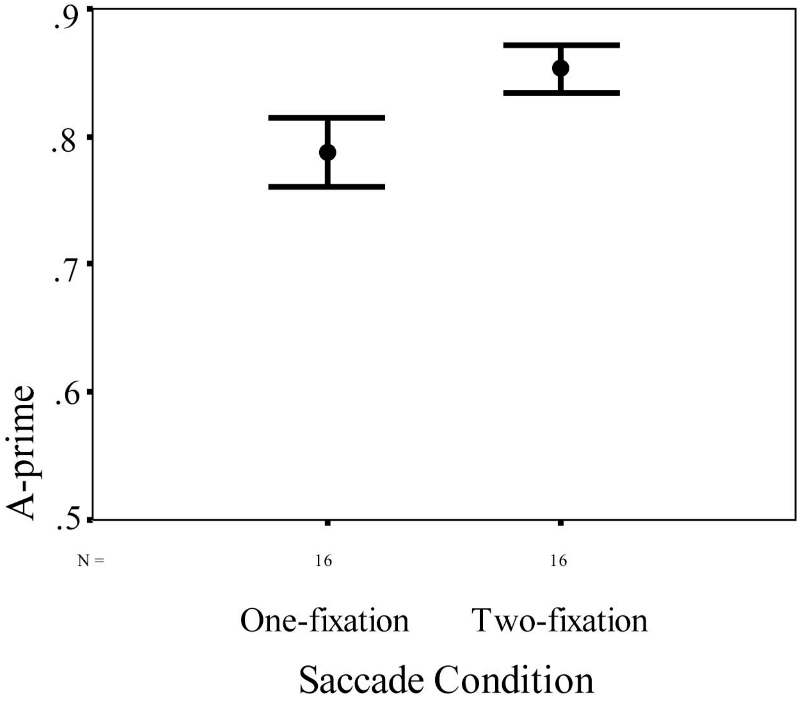

Figure 7. Participants' discrimination performance measured by A' in the follow-up experiment comparing one- and two- fixation conditions given the same total fixation duration (through

away from the center (mostly to the eyes; e.g., Henderson et al., 2005). In the current study, in order to examine the preferred landing position (PLP) of the first fixation, the face is initially presented parafoveally, so the participant has to make an initial saccade to the face. Thus we are able to show that the first two fixations, which are critical to face recognition performance, are around the center of the face. The participants do start to look at the eyes from the third fixation on (Figure 5), consistent with the existing literature.

Previous studies using the Bubbles procedure showed the most diagnostic features for face identification are the eyes (e.g., Schyns et al., 2002). Standard approaches to modeling eye fixation and visual attention are usually based on a saliency map, calculated according to biologically motivated feature selection or information maximization (e.g., Itti, Koch, & Niebur, 1998; Yamada & Cottrell, 1995). These models predict fixations on the eyes when viewing faces; our results showed this is not the case, suggesting eye movements in face recognition are different from those in scene viewing or visual search. Also, recent research has suggested a dissociation between face and object recognition: faces are represented and recognized holistically, involving less part-based shape representation compared with objects (e.g., Farah et al., 1995). The result that the first two fixations were around the center of nose instead of the eyes is consistent with this claim. It is also consistent with previous successful computational models that use a whole-face template-like representation (e.g., Dailey & Cottrell, 1999; O'Toole et al., 1988).

Study	Coordinate (pixel)				Duration (ms)		Saccade length from the previous fixation (pixel)	
	mean		standard error		mean	standard error	mean	standard error
	x	y	x	y				
1	115.6	128.8	2.6	5.3	235	15	n/a	n/a
2	117.6	119.3	4.6	5.1	283	28	55.7	3.0
3	123.8	116.0	5.4	3.3	340	26	49.0	2.2

Test	Coordinate (pixel)				Duration (ms)		Saccade length from the previous fixation (pixel)	
	mean		standard error		mean	standard error	mean	standard error
	x	y	x	y				
1	113.3	131.2	3.2	5.9	295	23	n/a	n/a
2	118.7	133.0	3.2	3.0	315	15	51.4	4.1
3	119.3	121	4.0	3.8	287	14	53.4	6.2

Table 1. Fixations and their duration during the study and the test phases. The center of the image is at (x, y) = (120.5, 148.5) in pixels; the size of the image is 240 pixels (width) x 296 pixels (height).

Our result is consistent with the view that the face-specific effects are in fact expertise-specific (e.g., Gauthier et al., 1999). Due to our familiarity with the information profile of faces, fixations on individual features may only generate redundant processes; instead, a more efficient strategy is to get as much information as possible with just one fixation. Given a perceptual span large enough to cover the whole stimulus and the fact that visual acuity drops from fovea to periphery, this fixation should be at the "center of the information," where the information is balanced in all directions; it may also be the optimal viewing position (OVP). Indeed, it has been shown that the OVP in word recognition can be modeled by an algorithm that calculates the "center of the information" (Shillcock, Ellison, & Monaghan, 2000). Our data showed that the first two fixations were indeed around the center of nose (Figure 4). Note that it is an artifact of averaging that they look very close to each other in Figure 4 (as can be seen in Figure 5); their locations were significantly different from each other. To further quantify this artifact, we compared the saccade lengths (in pixels) between the first two fixations during the study and test phases (Table 1), and found that they were not significantly different; the difference was in fixation duration - fixations were longer at test ($F(1, 15) = 18.352$, $p = 0.001$, $p_{rep} = 0.988$, $\eta_p^2 = 0.550$). Hence, even though two fixations suffice in face recognition, they are relatively long fixations. Our follow-up experiment also shows that, given the same total fixation duration, the participants have better performance when they are allowed to make two fixations compared with one. This suggests that the second fixation has functional significance: to obtain more information from a different location.

The phenomenon that the first fixation is to the left of the center is consistent with the left side bias effect in face perception (Gilbert & Bakan, 1973): a chimeric face made from two left half-faces from the viewer's perspective is judged more similar to the original face than that made from two right half-faces. It has been argued to be an indicator of right hemisphere (RH) involvement in face perception (Burt & Perrett, 1997; Rossion et al., 2003). Mertens, Siegmund, and Grusser (1993) reported in a visual memory task, the overall time that the fixation remained in the left gaze field was longer than the right for faces, but not for vases. Leonards and Scott-Samuel (2005) showed that participants have their initial saccade direction to one side, mostly left, for faces, but not for landscapes, fractals, or inverted faces. Vinette, Gosselin, and Schyns (2004) used the Bubbles procedure and showed that the earliest diagnostic feature used in face identification was the left eye. Joyce (2000) found that fixations during the first 250ms in face recognition tended to be on the left half-face. Our result thus is consistent with these previous studies.

Gosselin and Schyns (2001) argued that in their study the left eye was more informative in face identification because the left side of the images used had more shadows and thus was more informative as to face shape. Nevertheless, this artifact was not present here, since we mirror-reversed the images on half of the trials. Thus, it must be a subject-internal bias that drives the left side bias effect. It may be due to the importance of low spatial frequency information in face recognition (e.g., Whitman & Konarzewski-Nassau, 1997; Dailey & Cottrell, 1999), and the RH advantage in processing low spatial frequency information (Sergent, 1982; Ivry & Robertson, 1999). Due to the contralateral projection from the visual hemifields to the hemispheres, the left half-face from the viewer's perspective has direct access to the RH when the face is centrally fixated. It has been shown that each hemisphere plays a dominant role in the processing of the stimulus-half to which it has direct access (e.g., Hsiao, Shillcock, & Lavidor, 2006). Thus, the representation of the left half-face may be encoded by and processed in the RH, making it more informative than the right half-face.

There may be other factors that influence the OVP for face recognition. For example, they may be influenced by fundamental differences in the information profile normally portrayed as relevant to a given task. Thus, different tasks on the same stimuli may have different OVPs, especially when the distributions of information required are very different. The left side bias effect might also be due to a biologically based face asymmetry that is normally portrayed in daily life. In addition, Heath, Rouhana, and Ghanem (2005) showed that the left side bias effect in facial affect perception can be influenced by both laterality and script direction: right-handed Roman script readers demonstrated the greatest leftward bias, and Arabic script readers (i.e., scripts read from right to left) demonstrated a mixed or weak rightward bias (cf. Vaid & Singh, 1989). In our study the participants scanned from left to right, consistent with their reading direction (i.e. English readers). Further examinations are required to see whether Arabic readers have a different scan path from English readers.

In summary, we show that two fixations suffice in face recognition; both of them are around the center of nose, with the first one slightly but significantly to the left of the center. We argue that this location may be the "center of the information", or the OVP for face recognition. Different tasks on the same stimuli may have different OVPs and PLPs, since they may require different information from the stimuli. Future research is to examine whether the PLPs in other tasks can also be predicted from the claim about OVP being at the center of the information, and the factors that influence eye fixations during face recognition.

References

Althoff, R. R & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *25*, 997-1010.

Barrington, L., Marks, T., & Cottrell, G. W. (2007). NIMBLE: A kernel density model of saccade-based visual memory. In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Annual Cognitive Science Society* (pp. 77-82). Austin, TX: Cognitive Science Society.

Barton, J. J. S., Radcliffe, N, Cherkasova, M. V., Edelman, J., & Intriligator, J. M. (2006). Information processing during face recognition: The effects of familiarity, inversion, and morphing on scanning fixations. *Perception*, *35*, 1089-1105.

Burt, D. M. & Perrett, D. I. (1997). Perceptual asymmetries in judgements of facial attractiveness, age, gender, speech and expression. *Neuropsychologia*, *35*, 685-693.

Brysbaert, M. & Nazir, T. (2005) Visual constraints in written word recognition: evidence from the optimal viewing-position effect. *Journal of Research in Reading*, *28*, 216-228.

Dailey, Matthew N. & Cottrell, Garrison W. (1999) Organization of face and object recognition in modular neural networks. *Neural Networks*, *12(7-8)*, 1053-1074.

Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1995). What is "special" about face perception? *Psychological Review*, *105*, 482-498.

Gauthier, I., Tarr, M. J., Anderson, A. W., Skudlarski, P., & Gore, J. C. (1999). Activation of the middle fusiform "face area" increases with expertise in recognizing novel objects. *Nature Neuroscience*, *2*, 568–573.

Gilbert, C. & Bakan, P. (1973). Visual asymmetry in perception of faces. *Neuropsychologia*, *11*, 355-362.

Gosselin, F. & Schyns, P.G. (2001) Bubbles: a technique to reveal the use of information in recognition tasks. *Vision Research*, *41*, 2261-2271.

Heath, R. L., Rouhana, A., & Ghanem, D. A. (2005). Asymmetric bias in perception of facial affect among Roman and Arabic script readers. *Laterality*, *10*, 51-64.

Henderson, J. M., & Williams, C. C., & Falk, R. J. (2005). Eye movements are functional during face learning. *Memory & Cognition*, *33*, 98-106.

Hsiao, J. H., Shillcock, R., & Lavidor, M. (2006). A TMS examination of semantic radical combinability effects in Chinese character recognition. *Brain Research*, *1078*, 159-167.

Ivry, R. B. & Robertson, L. C. (1999). *The two sides of perception*. MIT Press.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine*, 20, 1254 – 1259.

Joyce, C.A. (2000). Saving Faces: Using Eye Movement, ERP, and SCR Measures of Face Processing and Recognition to Investigate Eyewitness Identification. (Ph.D. Thesis) University of California, San Diego, US, Dissertation Abstracts International: Section B: The Sciences & Engineering. Univ Microfilms International: US, 2001 Mar. 61 (8-B): p. 444

Lacroix, J. P. W., Murre, J. M. J., Postma, E. O., & Van den Herik, H. J. (2006). Modeling recognition memory using the similarity structure of natural input. *Cognitive Science*, *30*, 121-145.

Leonards, U. & Scott-Samuel N.E. (2005). Idiosyncratic initiation of saccadic face exploration in humans. *Vision Research*, *45*, 2677-2684.

Mäntylä, T. & Holm, L. (2006). Gaze control and recollective experience in face recognition. *Visual Cognition*, *13*, 365-386.

Mertens, I., Siegmund, H., & Grusser, O. J. (1993). Gaze motor asymmetries in the perception of faces during a memory task. *Neuropsychologia*, *31*, 989-998.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh Inventory. *Neuropsychologia*, *9*, 97-113.

O'Regan, J. K. (1990). Eye movements and reading. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp.395-453). Amsterdam: Elsevier.

O'Regan, J. K., Lévy-Schoen, A., Pynte, J., & Brugaillère, B. (1984). Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 250-257.

O'Toole, A. J., Millward, R. B., & Anderson, J. A. (1988). A physical system approach to recognition memory for spatially transformed faces. *Neural Networks*, *1*, 179-199.

Phillips, P. J., Moon, H., Rauss, P. J., & Rizvi, S. (2000). The FERET evaluation methodology for face recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*, 1090-1104.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8,* 21-30.

Rossion, B., Joyce, C. A., Cottrell, G. W., & Tarr, M. J. (2003). Early lateralization and orientation tuning for face, word, and object processing in the visual cortex. *Neuroimage*, *20*, 1609-1624.

Schyns, P. G., Bonnar, L., & Gosselin, F. (2002). Show me the features! Understanding recognition from the use of visual information. *Psychological Science*, *13*, 402-409.

Sergent, J. (1982) The cerebral balance of power: Confrontation or cooperation? *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 253-272

Shillcock, R., Ellison, T. M., & Monaghan, P. (2000). Eye-Fixation Behavior, Lexical Storage, and Visual Word Recognition in a Split Processing Model. *Psychological Review*, *107*, 824-851.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, *31*, 137-149.

Vaid, J. & Singh, M. (1989). Asymmetries in the perception of facial affect: is there an influence of reading habits? *Neuropsychologia*, *27*, 1277-1287.

Vinette, Gosselin, & Schyns. (2004). Spatio-temporal dynamics of face recognition in a flash: it's in the eyes. *Cognitive Science*, *28*, 289-301.

Whitman, D. & Konarzewski-Nassau, S. (1997). Lateralized Facial Recognition: Spatial Frequency and Masking Effects. *Archives of Clinical Neuropsychology*, *12*, 428-428(1).

Yamada, K. & Cottrell. G. W. (1995). A model of scan paths applied to face recognition. *Proceedings of the Seventeenth Annual Cognitive Science Conference*, *Pittsburgh, PA*, pp. 55-60, Mahwah: Lawrence Erlbaum.

Zhang, L. & Cottrell, G. W. (2006). Look Ma! No network: PCA of gabor filters models the development of face discrimination. *Proceedings of the Twenty-eighth Annual Cognitive Science Society Conference*.

Author Notes

Janet Hui-wen Hsiao and Garrison W. Cottrell, Department of Computer Science & Engineering, University of California, San Diego.

Correspondence should be addressed to Janet Hsiao, 9500 Gilman Drive, La Jolla, CA 92093-0404, or to the authors at jhsiao@cs.ucsd.edu or gary@ucsd.edu, respectively.

Acknowledgments

We are grateful to the NIH (grant #MH57075 to G.W. Cottrell), the JSMF (PEN, I. Gauthier, PI), and the NSF (grant #SBE-0542013 to the TDLC, G.W. Cottrell, PI). We thank the editor and two anonymous referees for helpful comments.

[1] $A' = 0.5 + \left[sign(H-F)\frac{(H-F)^2 + |H-F|}{4\max(H,F) - 4HF} \right]$, where H and F are the hit rate and false alarm rate respectively. The d’ measure may be affected by response bias when assumptions of normality and equal standard deviations are not met (Stanislaw & Todorov, 1999). In the current study, the criterion was negative; indeed, the percentage of “no” responses was marginally above chance ($p = 0.06$).

[2] We only analyzed the first three fixations since some participants did not make more than three fixations during test. Greenhouse-Geisser correction was applied whenever the test of sphericity did not reach significance.

[3] This analysis showed a fixation effect in the x-direction ($F(2, 1991.670) = 21.845$, $p < 0.001$, $p_{rep} > 0.999$): the participants scanned from left to right for both phases. However, this effect was not significant in the analysis in which data were averaged by subject. In the y-direction, there was a phase effect ($F(1, 1867.005) = 11.026$, $p = 0.001$, $p_{rep} = 0.986$) and a fixation effect ($F(2, 1933.095) = 12.633$, $p < 0.001$, $p_{rep} = 0.999$).

[4] We recruited six male and ten female Caucasian UCSD students (mean age 22); all right handed and all with normal or corrected to normal vision. The same apparatus, design, and procedure was used, except that there were only two fixation conditions at test.

[5] In the one-fixation condition, the average duration of the first fixation (before the participant moved their eyes away and the image moved with their eye gaze) was 308 ms, which is not significantly different from that in the two-fixation condition (311 ms; t-test, *n.s.*), or that in the previous experiment (295 ms; t-test, *n.s.*). This shows the participants did not attempt to make longer fixations because of the gaze contingent design.