

Task-driven visual exploration at the foveal scale

Natalya Shelchkova^a, Christie Tang^b, and Martina Poletti^{a,1}

^aDepartment of Neuroscience, University of Rochester Medical Center, Rochester, NY 14627; and ^bDepartment of Psychological and Brain Sciences, Boston University, Boston, MA 02215

Edited by Michael E. Goldberg, Columbia University College of Physicians, New York, NY, and approved January 31, 2019 (received for review July 18, 2018)

Humans use saccades to inspect objects of interest with the foveola, the small region of the retina with highest acuity. This process of visual exploration is normally studied over large scenes. However, in everyday tasks, the stimulus within the foveola is complex, and the need for visual exploration may extend to this smaller scale. We have previously shown that fixational eye movements, in particular microsaccades, play an important role in fine spatial vision. Here, we investigate whether task-driven visual exploration occurs during the fixation pauses in between large saccades. Observers judged the expression of faces covering approximately 1° , as if viewed from a distance of many meters. We use a custom system for accurately localizing the line of sight and continually track gaze position at high resolution. Our findings reveal that active spatial exploration, a process driven by the goals of the task, takes place at the foveal scale. The scanning strategies used at this scale resemble those used when examining larger scenes, with idiosyncrasies maintained across spatial scales. These findings suggest that the visual system possesses not only a coarser priority map of the extrafoveal space to guide saccades, but also a finer-grained priority map that is used to guide microsaccades once the region of interest is foveated.

microsaccades | ocular drift | fine spatial vision | face perception | priority maps

Visual exploration is traditionally studied with scenes that cover a relatively large portion of the visual field. In these conditions, saccades redirect the center of gaze toward interesting objects, so that they can be inspected with the high-acuity foveola. It is well established that humans tend to look at the most informative regions of the scene (1) and that this process is influenced by the goals of the task (2). The foveola covers only $\approx 1^\circ$ of visual angle, less than 0.1% of the visual field (3). Nevertheless, because of the fractal statistics of natural scenes and the scaling of retinal receptors, the input stimulus in this region is as complex as anywhere else on the retina. Can the concept of top-down, task-driven visual exploration extend to the much smaller scale of the foveola during the intersaccadic intervals?

During fixation the eyes are never at rest but continue to move with a jittery motion, known as ocular drift, and with microsaccades, small saccades ($<0.5^\circ$) that keep the stimulus within the foveola (4, 5). These eye movements are crucial for fine spatial vision (6, 7). In laboratory tasks, microsaccades are finely tuned to bring the preferred locus of fixation on fine spatial patterns (7). In this study, we investigated whether microsaccades, rather than being a simple recentering mechanism, are used to explore naturally complex foveal stimuli, in the same way humans use saccades to examine large visual scenes. To examine this we used human faces.

Appropriately interpreting facial expressions and gaze direction are fundamental human abilities, and the visuomotor system is highly specialized in extracting information from faces. Generally, humans scan faces using a “T” pattern (8, 9). When performing a facial recognition task, the first two saccades are the most relevant as performance saturates after two fixations (10). During this period the visual system optimizes the acquisition of information by looking at the most diagnostic features. As a result, when judging facial expression, humans tend to

look at the mouth region (9, 11), whereas scanning the upper part of the face is mostly associated with recognition tasks (10). When the face is presented at an eccentric location, the first saccade to the face is the most important for facial recognition (12). It normally brings the gaze close to the nose, but its exact landing location is also biased by the task demands (12). Therefore, by examining saccade landing positions and the scan paths of observers looking at faces it is possible to infer the task performed (13). Crucially, while these patterns of visual exploration are seen in most subjects, there are significant individual variations (13–15).

Visual exploration of faces, as visual exploration of scenes, has been primarily examined using stimuli spanning many degrees of visual angle. These stimuli cover not just the fovea, but also the parafovea and the visual periphery. However, humans view faces from a range of different distances, and the ability to recognize facial expressions extends to spatial scales much smaller than those typically studied. Humans can tell whether somebody is angry or happy or whether somebody is looking at them, even when the person is many meters away. In these circumstances, the face may cover less than 1° , and the distance between the different features may be in the order of arcminutes. There are mainly two reasons why visual exploration at the scale of the foveola has been little investigated. First, it is often implicitly assumed that the visual system simply needs to maintain fixation once a stimulus is foveated, and the need for further exploration is not immediately recognized.

Significance

Visual exploration is driven by saccades, which bring the objects of interest into the small high-acuity portion of the visual field, the foveola. While visual exploration is generally studied at large spatial scales, here we show that it also occurs at a much finer scale when examining complex foveal stimuli. By using high-precision methods for localizing the gaze, we show that, in the periods in between saccades, tiny gaze shifts actively inspect the stimulus based on the task demands. These visual scanning strategies closely resemble those adopted for exploring larger scenes, with individual differences maintained across spatial scales. Fine spatial vision, therefore, results from a synergy of visual processes, motor behaviors, and cognitive factors.

Author contributions: N.S., C.T., and M.P. designed research; N.S., C.T., and M.P. performed research; M.P. supervised the research; N.S., C.T., and M.P. analyzed data; and N.S. and M.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Data deposition: The data necessary to generate all the figures (containing data) in the main manuscript and the matlab scripts used to produce these figures have been deposited on the Open Science Framework repository (https://osf.io/tusgd/?view_only=ce1e605d448e4014b50e6a7548ae37ba).

¹To whom correspondence should be addressed. Email: martina.poletti@urmc.rochester.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1812222116/-DCSupplemental.

Second, whereas examining visuomotor scanning strategies over a large visual scene is relatively straightforward, being able to accurately localize gaze within a region as small as the foveola is a challenging task.

Here, we used high-resolution eye tracking and a state-of-the-art system for gaze-contingent control that enables more accurate localization of the line of sight compared with standard techniques (5). We examined the oculomotor behavior at fixation by precisely mapping gaze position onto the foveal stimulus. In two experiments, we first examined whether visual exploration at the foveal scale is top-down guided based on the task demands, while the physical stimulus remains unchanged. Then, we investigated how visual exploration at the scale of the foveola compares to the exploration at a larger scale.

Results

To explore whether task-driven visual exploration extends to the fine scale of the foveola we conducted a simplified version of a “Yarbus experiment”; subjects performed two different tasks with the same set of stimuli. In one task participants were asked to judge whether a face was looking at them and in another task whether the face was smiling at them (Fig. 1 *A–D*). Stimuli were presented foveally and covered approximately 1° of visual angle. The distance between the two task-relevant features, eyes and mouth, and the initial fixation location was the same (18′, with prime indicating unit of arcminutes; Fig. 1*C*). We classified gaze position based on where it was on the stimulus. Three main regions were identified: eyes, nose, and mouth (Fig. 1*E*). If the gaze was not in any of these regions, it was classified as being

on the background. If exploration of complex foveal stimuli is top-down driven, we expect the pattern of eye movements to systematically change in the two tasks. The pattern of eye movements on the stimulus was examined at high resolution while subjects performed the task.

Influence of the Task on the Examination of Foveal Stimuli. Our findings show that, despite the small size of the stimuli, and despite the fact that they were already ideally placed within the foveola to perform both tasks, subjects actively examined them using different scanning patterns in the two tasks. When asked to judge gaze direction, subjects’ gaze shifted toward the eyes region (Fig. 2 *A* and *B*); on the other hand, when judging facial expression, subjects spent more time on the mouth region (Fig. 2 *C* and *D* and *Movie S1*). Microsaccadic behavior was very consistent across subjects; most of the microsaccades landed on the eyes in the gaze direction task (0.70 ± 0.13 on the eyes vs. 0 microsaccades landing on the mouth; $P < 0.0001$, two-tailed paired *t* test), but this pattern flipped when judging facial expression, with most microsaccades landing on the nose and on the mouth (0.1 ± 0.10 on the eyes vs. 0.5 ± 0.33 on the mouth; $P = 0.02$, two-tailed paired *t* test, Fig. 2 *F* and *G*).

The oculomotor behavior in both tasks differed compared with the normal physiological fixational instability when subjects maintained fixation on a single point. When maintaining fixation, the amplitude of microsaccades was lower ($16' \pm 2'$ in the task vs. $13' \pm 3'$ during sustained fixation; $P = 0.007$, paired two-tailed *t* test), and most microsaccades maintained the gaze close to the center of the display, the spatial location corresponding to

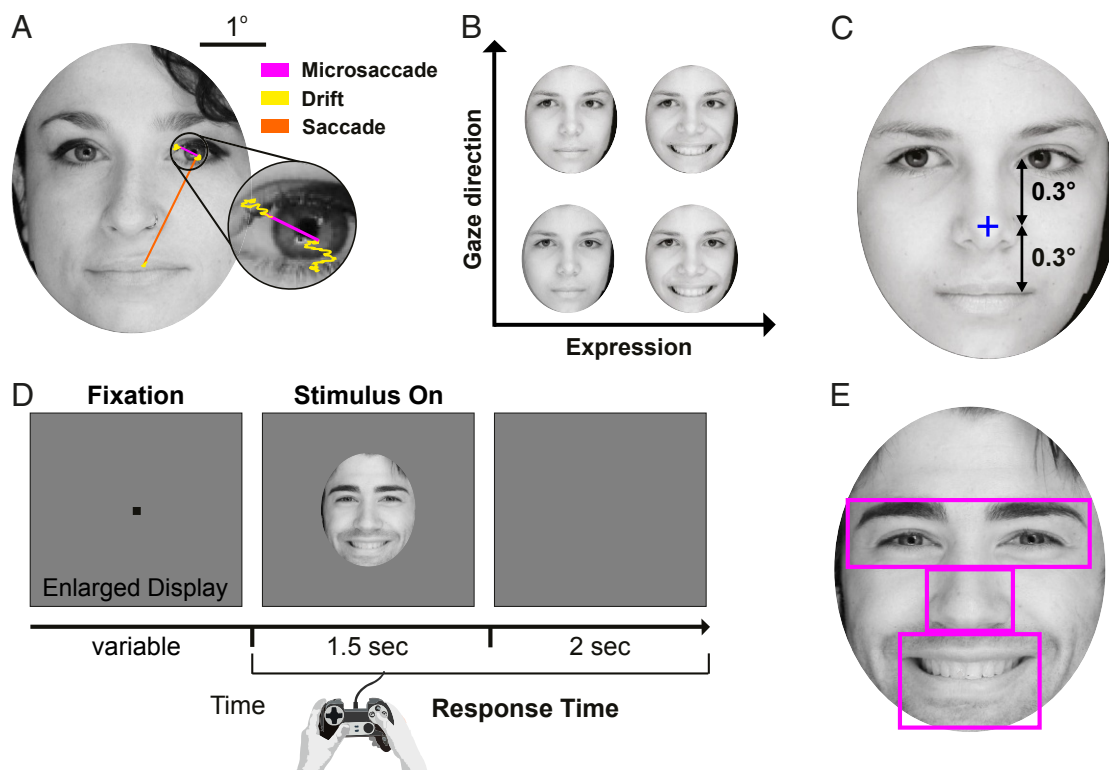


Fig. 1. Methods (experiment 1). (*A*) An example of eye movements recorded by means of a high-precision eye tracker. *Inset* shows eye movements during a fixation period. (*B*) Stimuli were generated by changing gaze direction and shape of the mouth. The same face was presented in four different versions: gaze looking straight or looking away and smiling or neutral expression. In the gaze direction task, subjects judged gaze direction, and in the expression task they judged whether or not the face was smiling. (*C*) The distance between the eyes/mouth and the initial fixation location (blue cross) was the same. The face covered 1.46° of visual angle in height. (*D*) Experimental paradigm. After a brief period of fixation a face was presented for 1.5 s at the center of the display. Subjects could respond at any time during the stimulus presentation and after its offset. (*E*) Gaze position on the stimulus was mapped at high resolution based on which feature the gaze was on. The feature regions used for data analysis are shown here delimited by a pink bounding box.

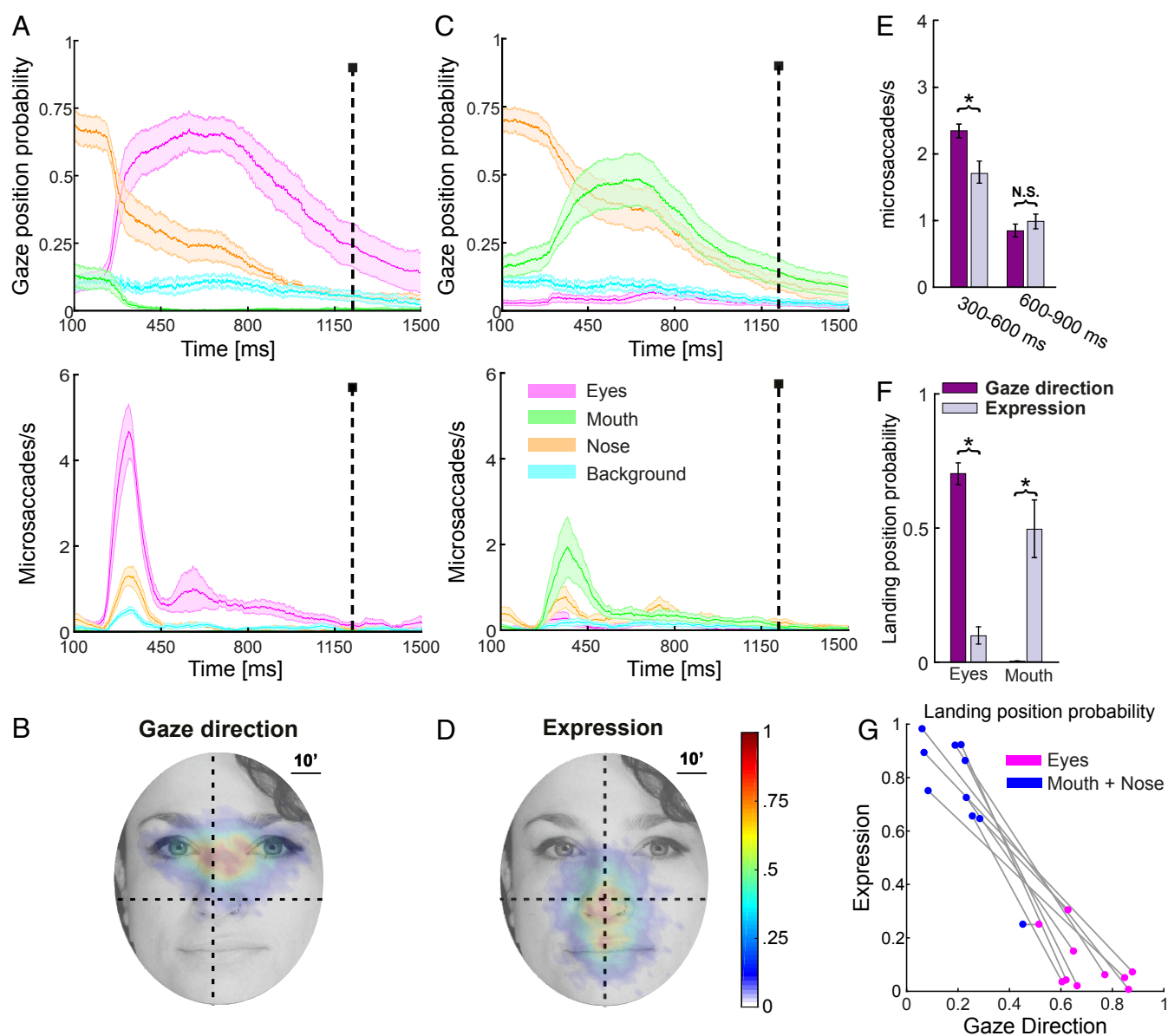


Fig. 2. Experiment 1 results. (A and C) Average probability of gaze position distribution (Top) and microsaccade landing position (Bottom) in the gaze direction (A) and the expression (C) tasks ($n = 10$). Data have been filtered using a running average with a 100-ms window. Dashed black lines mark the average response time. Shaded regions are SEM. (B and D) Average 2D normalized gaze distribution probability in the gaze direction (B) and the expression (D) tasks. (E) Average rate of microsaccades at the beginning and at the end of the trial for the two tasks. (F) Average probability of microsaccades landing on the eyes and on the mouth in the two tasks in the interval 300–600 ms from stimulus onset. Asterisks mark a statistically significant difference ($*P < 0.05$, two-tailed paired t test; n.s., not significant). (G) Single-subject probabilities of microsaccades landing on the mouth and nose vs. eyes in the two tasks. The lines connect the proportions of each single subject in both tasks.

the nose in the task (0.52 ± 0.2 vs. 0.23 ± 0.3 , 0.11 ± 0.1 , and 0.14 ± 0.07 , for nose, mouth, eyes, and background, respectively). These findings further show that, even during brief fixation periods, the visuomotor system does not simply maintain fixation on the foveated stimulus but engages in active exploration guided by the specific goals of the task.

Task-Driven Changes in the Rate and Time Course of Microsaccades. Furthermore, the results of experiment 1 show that, not only was the landing position of microsaccades different based on the task performed, but also their rate and time course varied systematically. The average rate of microsaccades was higher in the gaze direction task in the interval from 300 ms to 600 ms from stimulus onset ($2.4 \text{ ms/s} \pm 0.6 \text{ ms/s}$ and

$1.7 \text{ ms/s} \pm 1 \text{ ms/s}$ for gaze direction and expression, respectively; $P = 0.027$, two-tailed paired t test) (Fig. 2E), but was virtually the same in the two tasks during the rest of the trial (600–900 ms; $0.9 \text{ ms/s} \pm 0.6 \text{ ms/s}$ and $1.0 \text{ ms/s} \pm 0.7 \text{ ms/s}$ for gaze direction and expression, respectively; $P = 0.3$, two-tailed paired t test). Microsaccade time course was also modulated by the task. The rate of microsaccades peaked approximately 80 ms earlier in the gaze direction task ($327 \text{ ms} \pm 17 \text{ ms}$) compared with the expression task ($403 \text{ ms} \pm 80 \text{ ms}$; $P = 0.01$, two-tailed paired t test) and with a simple fixation ($391 \text{ ms} \pm 49 \text{ ms}$; $P = 0.005$, two-tailed paired t test) (Fig. 3).

Visual Scanning Strategies at Different Spatial Scales. In a second experiment we examined how the spatiotemporal pattern of

visual exploration at the foveal scale compares to that of visual exploration of larger stimuli. Subjects viewed human faces and judged whether or not the face's expression was neutral. In the parafovea condition, each face covered an area of 11.5 deg^2 , as if it was viewed from a distance of 3 m. In the foveola condition, instead, faces covered an area of 0.7 deg^2 , as if they were viewed from a distance of 13 m (Fig. 4A).

When the stimulus extended to the parafoveal region, almost all observers followed a very stereotyped scanning pattern (Fig. 4B and E). Immediately before the stimulus onset subjects fixated on a marker at the center of the display, so their initial gaze position upon stimulus presentation was on the upper part of the nose region, approximately at the center of the face. After a brief period of saccadic suppression following the presentation of the stimulus, the rate of saccades sharply increased. During this time most of the saccades landed on the mouth (Fig. 4B, Right) [0.77 ± 0.3 vs. 0.15 ± 0.3 , 0.05 ± 0.07 , and 0.03 ± 0.03 probability of landing on eyes, nose, and background, respectively; ANOVA $F(3,45) = 30.3$; $P < 0.0001$; Tukey's honestly significant differences (HSD) post hoc tests, mouth vs. eyes, $P < 0.0001$; mouth vs. nose, $P < 0.0001$; and mouth vs. background, $P < 0.0001$] (Movie S2). The rate of saccades then gradually decreased back to baseline. This pattern of visual exploration is expected when the area of the stimulus covers many degrees. A tendency to look over the mouth when judging facial expression has been reported by a number of studies (9, 11, 15–17). Moreover, a bias toward the lower part of the face when judging facial expression was also reported for the first saccade bringing a face, presented in the visual periphery, to the center of gaze (12).

In the foveola condition the exploratory behavior was driven by microsaccades (average amplitude $15' \pm 3'$; SI Appendix, Fig. S1). Similar to what happens in the parafovea condition for saccades, after an initial suppression period, the rate of microsaccades peaked at approximately 400 ms ($371 \text{ ms} \pm 65 \text{ ms}$ microsaccade rate peak time in the foveola condition vs. $403 \text{ ms} \pm 87 \text{ ms}$ saccade rate peak time in the parafovea condition; $P = 0.23$, two-tailed paired t test). During the period in which microsaccade rate reached a peak (300–600 ms), most microsaccades landed on the mouth region [0.40 ± 0.3 vs. $0.16 \pm$

0.2 , 0.25 ± 0.1 , and 0.20 ± 0.1 probability of microsaccades landing on eyes, nose, and background, respectively; ANOVA $F(3,45) = 3.2$; $P = 0.03$; Tukey's HSD post hoc tests, mouth vs. eyes, $P = 0.02$; mouth vs. nose, $P = 0.3$; mouth vs. background, $P = 0.09$] (Fig. 4C, D, and F; SI Appendix, Fig. S2; and Movie S2). Overall, microsaccadic behavior in this task was less precise than the saccadic behavior, both within and across subjects. This could be due to the fact that the stimuli used in experiment 2 were slightly smaller than those used in experiment 1; the distance between features ranged between $10'$ and $15'$. Critically, the decline in fine pattern vision reported across the foveola is less steep than the decline from the fovea to the visual periphery. As a result, in the foveola condition there is less of a drive to shift the gaze as precisely as in the parafovea condition. A small microsaccade landing on the lower part of the nose region, or a microsaccade landing into the background region adjacent to a feature, would still land less than $\approx 5'$ away from the target region and would still be precise enough for this task. However, a microsaccade landing on the eye region or on its surrounding background likely shifts the preferred fixational locus too far from the mouth, the most informative feature for this task. Consistent with this idea, our data show that most of the microsaccades landing on the background, or on the nose, landed primarily in the lower part of these features closer to the mouth region (0.65 ± 0.23 and 0.35 ± 0.23 probability of "nose" microsaccades landing on the lower and upper part of the nose, respectively, $P = 0.03$, paired two-tailed t test; 0.66 ± 0.22 and 0.34 ± 0.22 probability of "background" microsaccades landing on the lower and upper part of the background, respectively, $P = 0.01$, paired two-tailed t test) (Movie S2).

Crucially, microsaccades that brought the center of gaze closer to the task-relevant feature benefited performance in this task. The task was trivial, so to make sure that subjects remained engaged in the task and that performance did not saturate we lowered the contrast of the images and included a number of more ambiguous expressions. While the percentage of correct responses was well above chance for all subjects, there were some variations in performance across individuals. The rate of microsaccades landing on the mouth region was positively correlated with the performance in the task across subjects (Pearson correlation coefficient $r = 0.58$, $P = 0.02$; SI Appendix, Fig. S3); that is, subjects characterized by a higher rate of microsaccades landing on this task-relevant region also showed higher performance in the task. This improvement was associated only with microsaccades landing on the mouth; performance was not correlated with the global rate of microsaccades and with the rate of microsaccades landing on the eyes or background ($r = -0.14$, $P = 0.60$ for microsaccades landing on the eyes and $r = 0.05$, $P = 0.85$ for microsaccades landing on the background).

To ensure that the pattern of eye movement recorded when subjects performed the task was, indeed, the result of an active exploration and not the mere outcome of the physiological instability of the eye at fixation, similar to experiment 1, we examined fixational eye movements when subjects were required to keep their gaze on a marker at the center of the display. The rate of microsaccades was higher and the amplitude of microsaccades lower during fixation compared with when the subjects performed the task ($1.5 \text{ ms/s} \pm 0.8 \text{ m/s}$ and $1.2 \text{ ms/s} \pm 0.6 \text{ m/s}$ fixation and task, respectively, $P = 0.04$, paired two-tailed t test; and $13.6' \pm 3.6'$ and $15' \pm 3'$ fixation and task, respectively, $P = 0.04$, paired two-tailed t test) (SI Appendix, Fig. S1). Moreover, microsaccade landing position and the overall spatial distribution of gaze position differed between fixation and the task. As illustrated in Fig. 4D (red dashed line) and SI Appendix, Fig. S4, when subjects fixated on a central marker on a blank background, the probability of microsaccades landing on the spatial region corresponding to the mouth in the task was close to zero and it was lower than the probability of landing anywhere else (0.06 ± 0.06

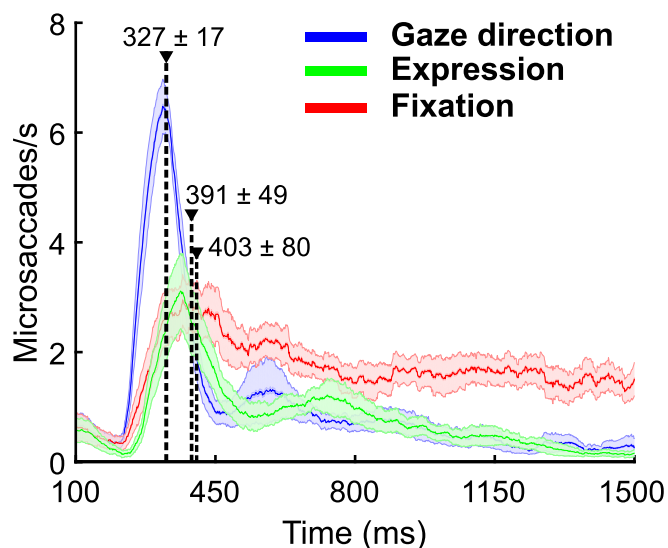


Fig. 3. Temporal occurrence of microsaccades. Shown is average microsaccade rate over time in experiment 1 and during sustained fixation. Data have been filtered using a running average with a 100-ms window. Dashed lines represent the average time when the rate of microsaccades reached a peak. Error bars represent SEM.

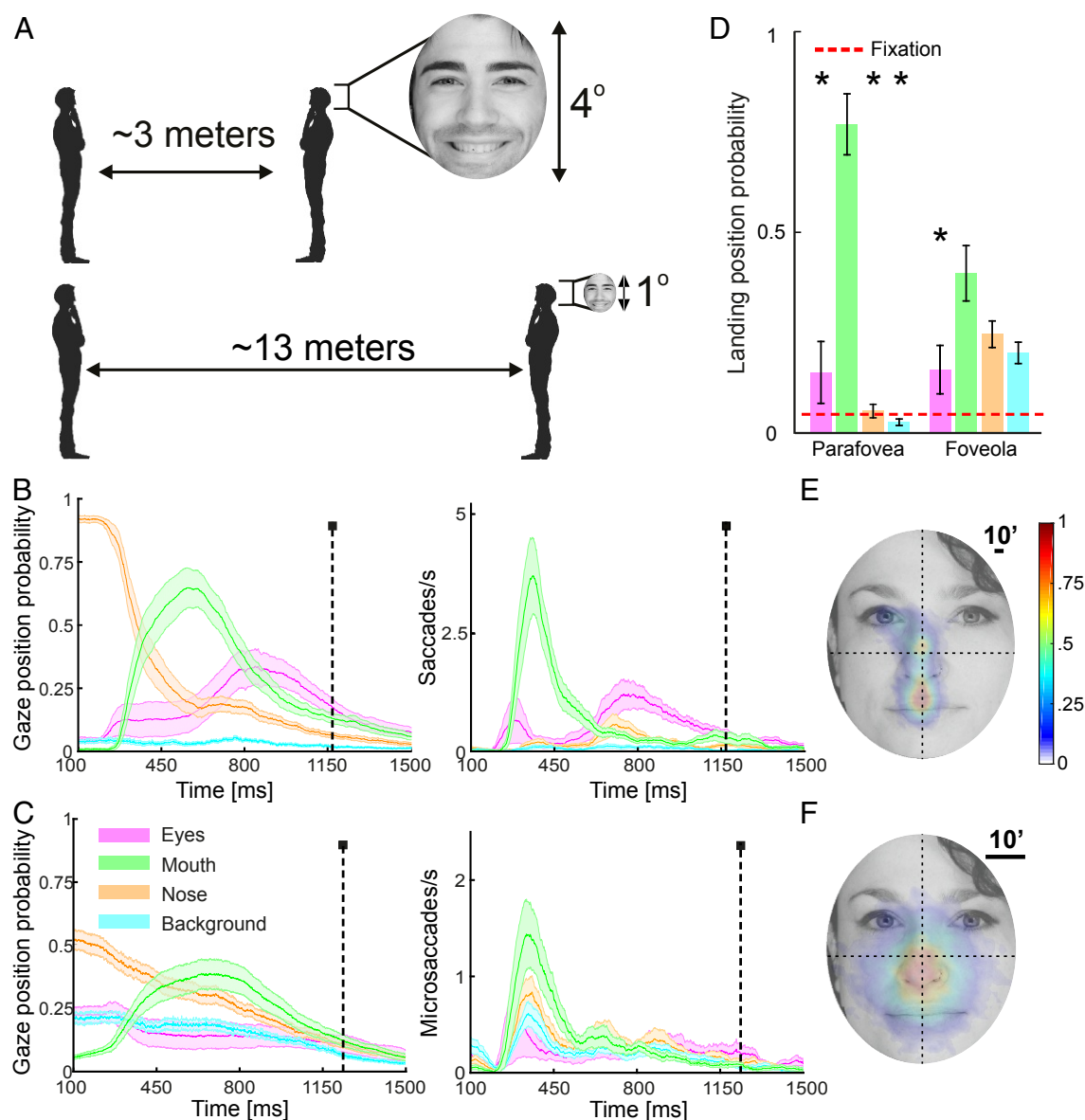


Fig. 4. Experiment 2 results. (A) Faces are normally viewed from different distances; the face of a person standing ≈ 3 m away spans 4° on the retina, but it spans only 1° when the observer is ≈ 13 m away. In experiment 2 faces covered either an area of 0.7 deg^2 (foveola condition, 1° height) or an area of 11.5 deg^2 (parafovea condition, 4.2° height). (B) Average distribution of gaze position (Left) and saccade landing position (Right) over time in the parafovea condition ($n = 16$). (C) Average distribution of gaze position (Left) and microsaccade landing position (Right) in the foveola condition ($n = 16$). Data have been filtered using a running average with a 100-ms window. Shaded regions are SEM. Dashed black lines mark the average response time. (D) Average probability of saccade (parafovea) and microsaccade (foveola) landing on different regions of the stimulus in the interval from 300 ms to 600 ms after the stimulus onset. For comparison, the average probability of microsaccade landing on the spatial region corresponding to the mouth is also shown when subjects maintained fixation on a marker in the absence of the stimulus (red dashed line). Asterisks mark a statistically significant difference (* $P < 0.05$, Tukey's HSD post hoc tests). Error bars represent SEM. The same color code (key in C) applies to B and D. (E and F) Average 2D normalized gaze distribution probability in the parafovea (E) and the foveola (F) conditions.

vs. 0.35 ± 0.1 , 0.32 ± 0.2 , and 0.27 ± 0.1 , for mouth, eyes, nose, and background, respectively, $P < 0.0001$; Tukey's HSD post hoc tests, mouth vs. eyes, $P < 0.0001$; mouth vs. nose, $P < 0.0001$; and mouth vs. background, $P = 0.0006$). Similar to experiment 1, these findings show that the motor behavior during the task differed from the physiological pattern of fixational eye movements when simply maintaining fixation, and it was modulated by the task performed.

Interestingly, not only were microsaccades modulated by the task, but also intersaccadic eye movements changed in the foveola condition. Ocular drift, the incessant jitter of the eye, was

characterized by a smaller diffusion coefficient when subjects performed the task with foveal stimuli compared with when they simply maintained fixation on a single point (diffusion coefficient at fixation $17 \text{ arcmin}^2 \pm 5 \text{ arcmin}^2$ vs. $14 \text{ arcmin}^2 \pm 4.3 \text{ arcmin}^2$ in the foveola condition, $P = 0.009$; *SI Appendix, Fig. S5*). Reducing the amount of displacement introduced by ocular drift may be beneficial in this task as it further enhances the high spatial frequency content of the stimulus (6, 18). These findings suggest that intersaccadic drift may be actively modulated either by the task or by the spatial characteristics of the visual stimulus.

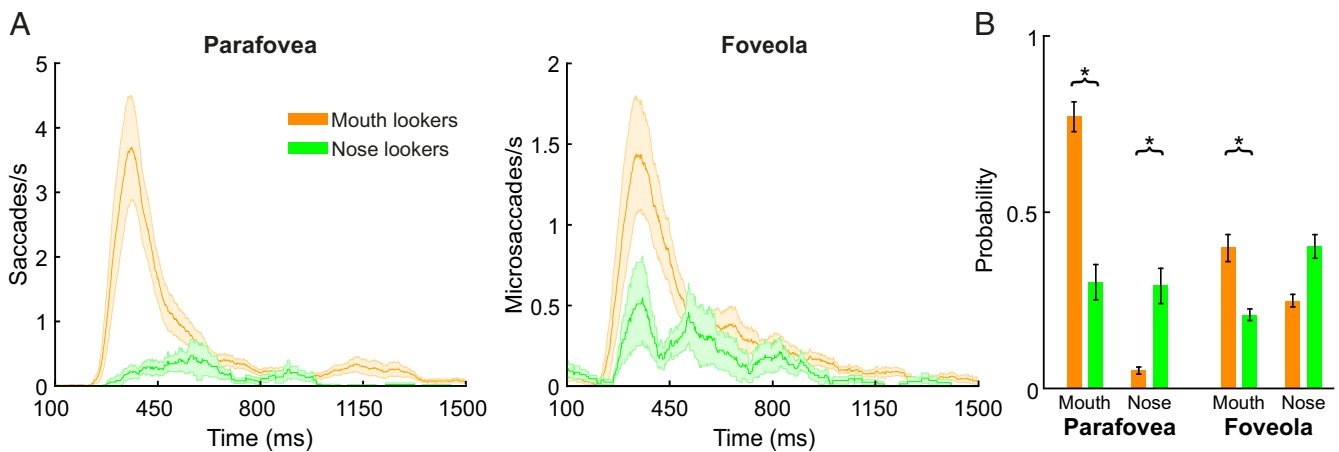


Fig. 5. Individual differences are maintained across scales. (A) Average rate of saccades (parafovea, *Left*) and microsaccades (foveola, *Right*) landing on the mouth during the course of the trial for nose lookers ($n = 5$) and mouth lookers ($n = 16$). (B) Probability of microsaccade and saccade landing over different regions of the stimulus for nose and mouth lookers. Probabilities are calculated in the interval from 300 ms to 600 ms from the stimulus onset. Asterisks mark a statistically significant difference ($*P < 0.05$, two-tailed t test). Error bars represent SEM.

Individual Differences Are Maintained Across Scales. It has been previously reported that the pattern of eye movements when viewing faces varies significantly across observers (13–15, 19, 20). Similarly, here we found that in the parafovea condition a small percentage of subjects (24% of the total, five subjects) maintained fixation around the center of the display (at the nose location) for the entire duration of the stimulus presentation (0.29 ± 0.23 probability of saccades landing on the nose and 0.30 ± 0.2 on the mouth for nose lookers vs. 0.05 ± 0.07 and 0.77 ± 0.3 for the mouth lookers; nose vs. mouth lookers, $P = 0.001$ and $P = 0.005$ for nose and mouth, respectively, two-tailed t test) (Fig. 5A and B). Although the nose lookers did not explore the face, their performance in the task was as good as that of the other subjects (88.7 ± 2.2 for nose lookers vs. 85.5 ± 6.4 for mouth lookers; $P = 0.3$, two-tailed t test). Because of their markedly different behavior, these subjects were removed from the main analysis. Notably, however, our data show that these individual differences were maintained across scales; the nose lookers showed a similar behavior in the foveola condition (0.40 ± 0.2 probability of microsaccades landing on the nose and 0.21 ± 0.08 on the mouth for nose lookers vs. 0.25 ± 0.13 and 0.40 ± 0.28 for the mouth lookers; $P = 0.04$, for mouth vs. nose lookers microsaccades landing on the mouth, two-tailed t test) (Fig. 5B). Similarly, also in the foveola condition the performance in the task was the same for nose and mouth lookers (78.4 ± 5 for nose lookers vs. 79.8 ± 7 for mouth lookers; $P = 0.7$, two-tailed t test).

Furthermore, even across the mouth lookers there were significant variations in the proportion of microsaccades landing on the eyes vs. those landing on the mouth. These differences, however, were also preserved across scales; the difference in the proportion of saccades/microsaccades landing on the eyes vs. on the mouth was highly correlated across subjects in the parafovea and in the foveola condition ($r = 0.77$, $P = 0.0005$). These findings show that idiosyncrasies in the visual scanning patterns are preserved across scales.

Discussion

Previous studies have shown that microsaccades precisely position a preferred foveal locus of fixation in high-acuity tasks (7, 21). This observation raises the question of whether scanning and exploration of visual objects and scenes, which have traditionally been ascribed to large saccades, also apply to microsaccades at a finer spatial scale. This question has so far remained unan-

swered primarily because of the challenges inherent in precisely determining the portion of the scene covered by the foveola. Our approach for accurate gaze localization has enabled us to circumvent these limitations and investigate this issue. Here, we show that task-driven visual exploration extends to the scale of the foveola. This process follows scanning strategies qualitatively similar, but smaller in scale, to those occurring during saccadic exploration. Microsaccades are modulated by the task demands both in space and in time: They consistently target task-relevant locations within the fovea, and, for a given stimulus, their rate and dynamics vary with the task at hand. These findings complement our previous work on microsaccades. They show that this oculomotor behavior is not the outcome of purely bottom-up recentering mechanisms, but is the manifestation of active, top-down-driven, visual scanning strategies.

The results reported here have important implications for the study of priority maps. In our experimental paradigm, exposure to the stimulus was relatively long, and subjects were left free to perform multiple saccades. However, in experiment 1, subjects delivered their response about 200 ms before the offset of the stimulus and, in most trials, after only one microsaccade (1.5 ± 0.6 and 1.2 ± 0.7 microsaccades in the gaze direction and expression tasks, respectively). Thus, the first microsaccade following stimulus onset appeared to be critical for performing the task. This first microsaccade, which generally occurred within the initial 350 ms of exposure, was clearly driven by the task. Later microsaccades were not as strongly directed toward a specific facial feature. These findings strongly suggest that the first microsaccade was driven by a priority representation of the foveal input. Thus, contrary to the general idea that the function of priority maps is to represent the relevance of stimuli outside the fovea, our results show that these maps must also include foveal representations.

The existence of a foveal priority map is supported by the notion that visual functions are not uniform across the foveola (7). This map enables selection of the most relevant regions of the foveal landscape to guide visual exploration and is responsible for directing the first microsaccade following stimulus presentation. Priority maps of the extrafoveal space are known to contribute to driving different effectors and behaviors, from eye movements to reaching (22, 23). However, microsaccades appear to be the only motor behavior that can be controlled at the fine scale of the foveola. This raises the question of whether this finer-grain priority map of the foveola is specifically

restricted to the guidance of microsaccades or can also be accessed by other systems. Further work is necessary to investigate this question.

Our work also shows that individual differences in visual exploration are maintained across spatial scales. During viewing of a face, significant variations in visual scanning strategies occur across individuals (13–15, 20). These idiosyncrasies are maintained over time (14), and they do not change even when central vision is blocked using an artificial scotoma (24). A difference between nose lookers and mouth lookers similar to the one we observed has already been reported in the landing position of the first saccade toward a face presented peripherally (14, 20). Although there is a general tendency of the first saccade to land just below the eyes, some individuals exhibit strong biases toward the nose or eye regions. The experimental paradigm used here differs in a number of ways from the paradigms of these previous studies. Our stimuli were not just considerably smaller (also in parafoveal condition); they were also presented centrally for a relatively long period. Yet individual differences in eye movements closely resemble those of these previous reports. While most of the observers primarily looked at the mouth, others kept fixation on the nose. It is possible that these different strategies reflect variations in perceptual sensitivity and retinal anatomy. Indeed, strong individual differences in the shape of the retina have been reported not only in the parafovea (25), but also in the foveola, with changes in cone density (25, 26) and the size of the foveal pit (27).

Previous research on cognitive/attentional influences on microsaccades has mostly focused on how microsaccade patterns are affected by the peripheral allocation of covert attention (28–30). These findings have emphasized the importance of controlling for these small gaze shifts when manipulating attention. However, in contrast with natural viewing conditions, the spatial cuing paradigms used by these studies provide minimal visual stimulation at the center of gaze. This observation prompts two important questions. First, whether microsaccades continue to be modulated by the peripheral allocation of attention in more natural conditions, when foveal stimulation is rich in details. Second, whether allocating attention far from the fovea leads by itself to a suppression of the visuomotor scanning associated with foveal exploration. Addressing these questions is fundamental for a better understanding of the interplay of attention and eye movements in more ecological conditions, when both foveal processing and peripheral processing are required during the time frame of one fixation. Previous work showed that analysis of foveal stimuli and the selection of the next saccade target proceed in parallel and independently (31), suggesting that allocating attention peripherally may not necessarily interfere with the foveal exploration.

In sum, our work shows that fine oculomotor behavior is more complex than commonly assumed. Foveating the stimulus of interest is necessary but not sufficient for fully examining the stimulus. During fixation, humans engage in subtle visuomotor explorative strategies to inspect fine spatial patterns. Microsaccades are the main motor component of these strategies. They efficiently explore the stimulus already falling within the foveola by sampling the most informative foveal locations with the preferred locus of fixation.

Materials and Methods

Observers. A total of 31 emmetropic human observers, all naive about the purpose of the study, participated in the experiments (age range 18–25 y). Twenty-one observers (17 males and 4 females) took part in experiment 2 (Fig. 4) and 10 observers (4 males and 6 females) in experiment 1 (Fig. 2). All experiments were approved by the Boston University Charles River Campus Institutional Review Board. Informed consent was obtained from all participants. Before conducting any experiment the experimenter reviewed and explained to the participant the material in the consent form. The form was

signed only after the subject fully understood the material and voluntarily agreed to take part in the study.

Stimuli and Apparatus. Stimuli were displayed on a fast-phosphor CRT monitor (Iiyama HM204DT) at a vertical refresh rate of 85 Hz and spatial resolution of $2,048 \times 1,536$ pixels (1 pixel = $0.53'$). Observers performed the task monocularly with their right eye while the left eye was patched. A dental-imprint bite bar and a headrest prevented head movements. The movements of the right eye were measured by means of a Generation 6 Dual Purkinje Image (DPI) eye tracker (Fourward Technologies), a system with an internal noise of $\sim 20''$ and a spatial resolution of $1'$ (5, 32). Vertical and horizontal eye positions were sampled at 1 kHz and recorded for subsequent analysis.

Stimuli were rendered by means of EyeRIS (33), a custom-developed system based on a digital signal processor, which allows flexible gaze-contingent display control. This system acquires eye movement signals from the eye tracker, processes them in real time, and updates the stimulus on the display according to the desired combination of estimated oculomotor variables.

Stimuli were generated by using images of faces taken from online databases (34, 35). The images used were pre-labeled according to their expression. In experiment 2 we grouped the faces into two main categories, neutral faces ($n = 125$) and faces expressing an emotion ($n = 125$). All faces were frontal views of either white males or females who had minimal facial hair or makeup. All of the images were converted to grayscale and faces were cropped to fit within an oval mask. The faces were chosen so that the difference between expressions was not too obvious and some faces were more ambiguous than others. Furthermore, in experiment 2 the contrast of the stimuli was lowered to increase the difficulty of the task. A subset of the neutral faces of experiment 2 was used to create a new database of images for experiment 1. The eyes and the mouth of these images were manipulated so that each face was presented in four different versions: looking straight and smiling or neutral and looking away and smiling or neutral. White noise was added to the images to increase the difficulty of the task. A total of 186 faces were used in experiment 1.

Procedure and Experimental Tasks. Every session started with preliminary setup operations that lasted a few minutes. The subject was positioned optimally and comfortably in the apparatus. Subsequently, a calibration procedure was performed in two phases. In the first phase, subjects sequentially fixated on each of the nine points of a 3×3 grid, as is customary in oculomotor experiments. These points were located 1.32° apart on the horizontal and vertical axes. In the second phase, subjects confirmed or refined the voltage-to-pixel mapping given by the automatic calibration. In this phase, they fixated again on each of the nine points of the grid while the location of the line of sight estimated on the basis of the automatic calibration was displayed in real time on the screen. Subjects used a joystick to correct the predicted gaze location, if necessary. These corrections were then incorporated into the voltage-to-pixel transformation. This dual-step calibration allows a more accurate localization of gaze position than standard single-step procedures, improving 2D localization of the line of sight by approximately one order of magnitude (5, 7). The manual calibration procedure was repeated for the central position before each trial to compensate for possible drifts in the electronics as well as microscopic head movements that may occur even on a bite bar.

Experiment 1. Subjects were instructed to perform two different tasks. In one task they were asked whether a face was looking straight ahead or away, whereas in the other task they were asked to judge whether a face was smiling or not. The height of the face measured 1.46° , and mouth and eyes were approximately at the same distance from the initially fixated location at the center of the display. The same set of stimuli was presented in both conditions. The two tasks were run in blocks. The blocks' presentation order was randomized. The same images were presented in both conditions and the order of images presentation was randomized for each task and subject.

Experiment 2. Subjects were instructed to judge whether a face expression was neutral or not. In the parafovea condition, the height of the face measured 4.2° , whereas in the foveola condition it measured 1° . The two conditions were run in blocks. The blocks' presentation order was randomized. The same images were presented in both conditions and the order of images presentation was randomized for each condition and subject.

In both experiments stimuli were presented for 1.5 s and subjects responded by pressing a button on a remote controller at any time during stimulus presentation and for a period of 4 s after the stimulus was turned off. The 1.5-s fixation trials were interleaved during the experiment. In these

trials observers were instructed to fixate on a marker at the center of the display.

Data Analysis. Recorded eye movement traces were segmented into separate periods of drift and saccades. Classification of eye movements was performed automatically and then validated by trained laboratory personnel with extensive experience in classifying eye movements. Periods of blinks were automatically detected by the DPI eye tracker and removed from data analysis. Only trials with optimal, uninterrupted tracking, in which the fourth Purkinje image was never eclipsed by the pupil margin, were selected for data analysis. Eye movements with minimal amplitude of 3' and peak velocity higher than 3°/s were selected as saccadic events. Saccades with an amplitude of less than 0.5° (30') were defined as microsaccades. Consecutive events closer than 15 ms were merged together into a single saccade to automatically exclude postsaccadic overshoots (36, 37). Saccade amplitude was defined as the vector connecting the point where the speed of the gaze shift grew greater than 3°/s (saccade onset) and the point where it became less than 3°/s (saccade offset). Periods that were not classified as saccades or blinks were labeled as drifts.

Trials with blinks/loss of tracks (3.2%, 3.2%, and 4.9% of the total trials for parafoveal condition, foveal condition, and experiment 2, respectively) and trials with early responses (<700 ms, 6% of the total trials) were discarded. To categorize gaze position during the task three regions were identified on the stimulus: nose, eyes, and mouth. If the gaze was not in any of these regions, it was categorized as being on the background. Averages across observers in different conditions and tasks were examined by means of one-way within-subjects ANOVAs followed by Tukey post hoc tests. Comparisons between two conditions and tasks across observers were tested using two-tailed paired *t* tests.

On average, performance was evaluated over 153 trials per condition per observer. Figs. 2–5 show summary statistics across observers, and Fig. 2G shows average values for each individual observer. The data necessary to generate all the figures (containing data) in the main manuscript and the matlab scripts used to produce these figures have been deposited on the Open Science Framework repository (https://osf.io/tusgd/?view_only=ce1e605d448e4014b50e6a7548ae37ba).

ACKNOWLEDGMENTS. This work has been supported by National Science Foundation Grant NSF-BCS-1534932 (to M.P.).

- Buswell GT (1935) *How People Look at Pictures* (Univ of Chicago Press, Chicago).
- Yarbus AL (1967) *Eye Movements and Vision* (Plenum Press, New York).
- Azzopardi P, Cowey A (1993) Preferential representation of the fovea in the primary visual cortex. *Nature* 361:719–721.
- Rolfs M (2009) Microsaccades: Small steps on a long way. *Vis Res* 49:2415–2441.
- Poletti M, Rucci M (2016) A compact field guide to the study of microsaccades: Challenges and functions. *Vis Res* 118:83–97.
- Kuang X, Poletti M, Victor JD, Rucci M (2012) Temporal encoding of spatial information during active visual fixation. *Curr Biol* 20:510–514.
- Poletti M, Listorti C, Rucci M (2013) Microscopic eye movements compensate for nonhomogeneous vision within the fovea. *Curr Biol* 23:1691–1695.
- Williams CC, Henderson JM (2007) The face inversion effect is not a consequence of aberrant eye movements. *Mem Cogn* 35:1977–1985.
- Schurgin MW, et al. (2014) Eye movements during emotion recognition in faces. *J Vis* 14:1–16.
- Hsiao JH, Cottrell G (2008) Two fixations suffice in face recognition. *Psychol Sci* 19:998–1006.
- Smith ML, Cottrell GW, Gosselin F, Schyns PG (2006) Transmitting and decoding facial expressions. *Psychol Sci* 16:184–189.
- Peterson MF, Eckstein MP (2012) Looking just below the eyes is optimal across face recognition tasks. *Proc Natl Acad Sci USA* 109:E3314–E3323.
- Kanan C, Bseiso DNF, Ray NA, Hsiao JH, Cottrell GW (2015) Humans have idiosyncratic and task-specific scanpaths for judging faces. *Vis Res* 108:67–76.
- Peterson MF, Eckstein MP (2013) Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. *Psychol Sci* 33:1216–1225.
- Mehoudar E, Arizpe J, Baker CI, Yovel G (2014) Faces in the eye of the beholder: Unique and stable eye scanning patterns of individual observers. *J Vis* 14:1–10.
- Arizpe J, Kravitz DJ, Yovel G, Baker CI (2012) Start position strongly influences fixation patterns during face processing: Difficulties with eye movements as a measure of information use. *PLoS One* 7:e31106.
- Blais C, Jack RE, Scheepers C, Fiset D, Caldara R (2008) Culture shapes how we look at faces. *PLoS One* 3:e3022.
- Rucci M, Iovin R, Poletti M, Santini F (2007) Miniature eye movements enhance fine spatial detail. *Nature* 447:852–855.
- Perlman SB, et al. (2009) Individual differences in personality predict how people look at faces. *PLoS One* 4:1–6.
- Peterson MF, Lin J, Zaun I, Kanwisher N (2016) Individual differences in face-looking behavior generalize from the lab to the world. *J Vis* 16:1–18.
- Ko HK, Poletti M, Rucci M (2010) Microsaccades precisely relocate gaze in a high visual acuity task. *Nat Neurosci* 13:1549–1553.
- Bisley JW, Goldberg ME (2010) Attention, intention, and priority in the parietal lobe. *Annu Rev Neurosci* 9:1–21.
- Zelinsky GJ, Bisley J (2015) The what, where, and why of priority maps and their interactions with visual working memory. *Ann N Y Acad Sci* 1339:154–164.
- Tsank Y, Eckstein MP (2017) Domain specificity of oculomotor learning after changes in sensory processing. *J Neurosci* 37:11469–11484.
- Curcio CA, Sloan KR, Kalina RE, Hendrickson AE (1990) Human photoreceptor topography. *J Comp Neurol* 292:497–523.
- Li KY, Tiruveedhula P, Roorda A (2010) Intersubject variability of foveal cone photoreceptor density in relation to eye length. *Invest Ophthalmol Vis Sci* 51:6858–6867.
- Tick S, et al. (2011) Foveal shape and structure in a normal population. *Invest Ophthalmol Vis Sci* 52:5105–5110.
- Hafed ZM, Clark JJ (2002) Microsaccades as an overt measure of covert attention shifts. *Vis Res* 42:2533–2545.
- Engbert R, Kliegl R (2003) Microsaccades uncover the orientation of covert attention. *Vis Res* 43:1035–1045.
- Yuval-Greenberg S, Merriam EP, Heeger DJ (2014) Spontaneous microsaccades reflect shifts in covert attention. *J Neurosci* 34:13693–13700.
- Ludwig CJ, Davies JR, Eckstein MP (2014) Foveal analysis and peripheral selection during active visual sampling. *Proc Natl Acad Sci USA* 111:E291–E299.
- Ko H-k, Snodderly DM, Poletti M (2016) Eye movements between saccades: Measuring ocular drift and tremor. *Vis Res* 122:93–104.
- Santini F, Redner G, Iovin R, Rucci M (2007) EyeRIS: A general-purpose system for eye movement contingent display control. *Behav Res Methods* 39:350–364.
- Minear M, Park DC (2004) A lifespan database of adult facial stimuli. *Behav Res Methods Instr Comput* 36:630–633.
- Vieira TF, Bottino A, Laurentini A, Simone MD (2014) Detecting siblings in image pairs. *Vis Comput* 30:1333–1345.
- Deubel H, Bridgeman B (1995) Fourth Purkinje image signals reveal eye-lens deviations and retinal image distortions during saccades. *Vis Res* 35:529–538.
- Stevenson SB, Roorda A (2005) Correcting for miniature eye movements in high resolution scanning laser ophthalmoscopy. *Ophthalmic Technologies XV*, Proceedings of SPIE, eds Manns F, Söderberg PG, Ho A, Stuck BE, Belkin M (SPIE, Bellingham, WA), Vol 5688, pp 145–151.