Task-driven visual exploration at the foveal scale

Natalya Shelchkova[a], **Christie Tang**[b], **and Martina Poletti**[a]

[a]Department of Neuroscience, University of Rochester Medical Center, Meliora Hall, Rochester, NY 14627 ; [b]Department of Psychological and Brain Sciences, Boston University, 64 Cummington Mall, Boston, MA 02215

This manuscript was compiled on November 19, 2018

Saccades enable visual exploration by bringing objects of interest in the foveola, the retinal region of highest visual acuity. Visual exploration is normally investigated with scenes spanning many degrees, yet, in everyday tasks the visual input falling within the one-degree foveola is often complex and composed of multiple parts. Here we examine whether task-driven visual exploration extends also at the foveal scale, during the fixation pauses in between saccades. We have previously shown that fixational eye movements, in particular microsaccades, play an important role in fine spatial vision, and that the visuomotor system is capable of fine-tuning these small eye movements. Using a custom-made system enabling accurate localization of the line of sight within one-degree of visual angle, we mapped gaze position at high-resolution during fixation on complex foveal stimuli. Observers judged facial expression of faces as if viewed from a distance of many meters, so that they covered approximately 1 deg of visual angle. Our findings reveal that active spatial exploration takes place also at the foveal level, and that it is driven by the goals of the task. The scanning strategies used at this scale resemble those used when examining larger scenes, and idiosyncrasies in the scanning pattern are maintained across scales. These findings strongly suggest that the visual system possesses not only a coarser priority map of the extrafoveal space to guide saccades, but also a finer grain priority map that is used to guide microsaccades once the region of interest is foveated.

Microsaccades | Spatial Attention | Ocular Drift | Fine Spatial Vision | Parafovea | Face Perception

Visual exploration has been traditionally studied using scenes extending to a relatively large portion of the visual field. In this context it is well established that humans tend to look at the regions of the scene containing the most meaningful information[1], and that visual examination is influenced by the goals of the task[2]. Saccades are instrumental in visual exploration as they bring interesting objects at the center of gaze, in the foveola, the retinal locus where visual resolution is highest. But can the concept of top-down task-driven visual exploration be applied also at the much smaller scale of the foveola during fixation periods? The foveola covers only $\approx 1^o$ 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 retinal projection falling onto this region is as complex as anywhere else on the rest of the retina.

During fixation the eyes are never at rest but continue to move with a jittery motion, known as ocular drift, and with microsaccades, saccades smaller than half a degree[4, 5]. These eye movements are crucial for fine spatial vision[6, 7]. In particular, microsaccades are finely tuned to bring the preferred locus of fixation on high-acuity stimuli[7]. Are microsaccades merely a refined re-centering mechanism triggered by visual offsets and driven only by low level factors? In this study we investigated whether microsaccades can be used to visually explore complex foveal stimuli based on the task goals, in the same way humans use saccades to examine large scenes. To address this question, as more complex stimuli we used human faces. The visuomotor system is, indeed, highly specialized in extracting information from faces, directing the gaze to the most diagnostic regions of the face based on the task[8–10].

Appropriately interpreting facial expressions and gaze direction are fundamental human abilities. By examining the scan paths of observers looking at faces it is possible to determine what are the attended regions and to infer the specific task performed[10]. With long enough presentation times, generally subjects scan faces using a "T" pattern[11, 12]. Yet, the first two saccades are the most relevant in facial recognition tasks as performance saturates after 2 fixations[9]. If, on the other hand, faces are presented only for a brief period of time, the visual system needs to optimize the extraction of information. As a result, the examined features depend on the task's goals; when judging facial expression humans tend to look at the mouth region[12, 13], while scanning the upper part of the face is mostly associated with recognition tasks[9]. When the face is presented at an eccentric location, the first saccade to the face is the most important for facial recognition[8]. It normally brings the gaze close to the nose, and its exact landing location is biased by the task demands[8]. Crucially, despite this pattern of visual exploration is seen in most subjects, there are significant individual variations[10, 14, 15].

Visual exploration of faces, as visual exploration of scenes, has always been examined using stimuli spanning many degrees of visual angle, therefore extending to the parafovea and the visual periphery. Yet, humans view faces from a range of different distances and the ability to recognize facial

Significance Statement

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 across the extrafoveal space, here we show that it can also be carried out at a finer scale when examining complex foveal stimuli. Thanks to high-precision techniques for localizing the gaze, our work revealed that during brief fixation periods tiny gaze shifts actively examine the foveal input based on the task goals. The visual scanning strategies implemented at this minute scale resemble those used when exploring larger visual scenes, with individual differences being maintained across scales. Fine spatial vision, therefore, results from a synergy of cognitive and motor factors.

.

MP devised the study, NS, CT and MP collected and analyzed the data, MP and NS wrote the manuscript.

[2]To whom correspondence should be addressed. E-mail: martina_poletti@urmc.rochester.edu

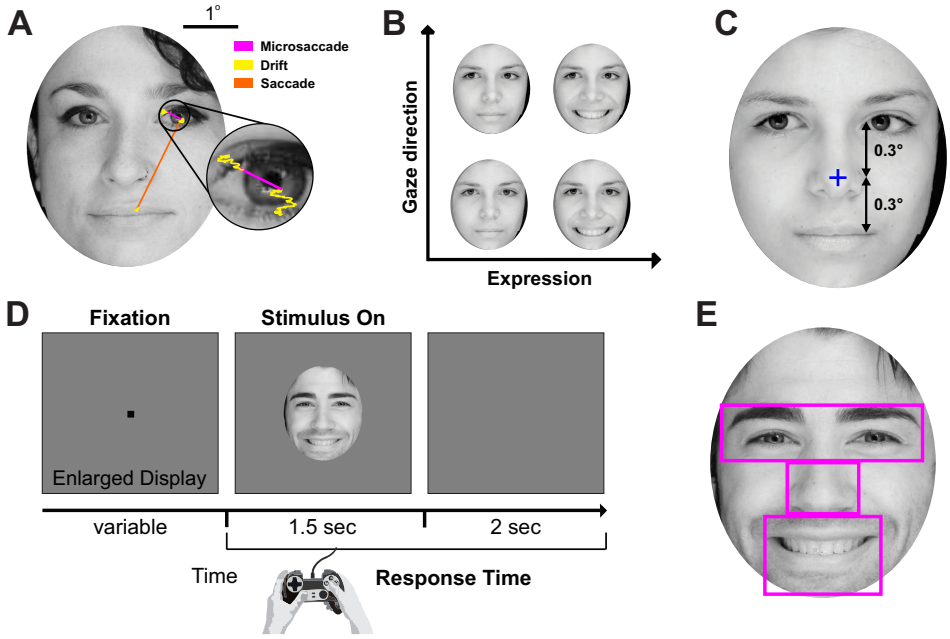

Fig. 1. Methods (Experiment 1). **A**, an example of eye movements recorded by means of a high-precision eyetracker. The enlargement shows eye movements during a fixation period. **B**, stimuli were generated by changing the face's gaze direction and shape of the mouth. The same face was presented in 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 approximately 1 degree of visual angle. **D**, experimental paradigm. After a brief period of fixation a face was presented for 1.5 sec 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.

expressions extends to spatial scales much smaller than those normally studied; human can tell whether somebody is angry or happy, or whether somebody is looking at them, even when a face is viewed from a distance of many meters. In this condition the face may cover approximately 1 deg of visual angle and the distance between the different features is in the order of arcminutes. Are all the details falling in the foveola processed simultaneously, or does the visual system guide sequential exploration of the facial features based on the task requests also at this scale? This issue has never been investigated. First, it is often implicitly assumed that the visual system simply needs to maintain fixation once a stimulus is foveated. Second, whereas examining eye movement scanning patterns over a large visual scene is relatively trivial, being able to accurately localize the gaze within a one-degree region of the visual field is a challenging task.

Using high-resolution eyetracking and a state-of-the-art gaze contingent display system, allowing for a more accurate localization of the center of gaze compared to standard techniques[5], we examined the oculomotor behavior at fixation by precisely mapping gaze position on the foveal stimulus. We first examined whether visual exploration at the foveal scale can be guided by top-down factors based on the request of the task 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-B-D*). Stimuli were presented foveally, and covered approximately 1 degree of visual angle. The distance between the two task-relevant features, eyes and mouth, and the initial fixation location was the same (18′; Fig. 1*C*). This raises the question of whether simply maintaining fixation at the center of the face is sufficient to perform both tasks, or if humans visually explore even such small stimuli. 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. We classified gaze position based on where it was on the stimulus. Three main regions were identified (Fig. 1*E*), eyes, nose and mouth. If the gaze was not in any of these regions it was classified as being on the background.

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 the stimuli were already ideally placed within the foveola to perform both tasks, subjects actively examined these fine stimuli using different scanning patterns in the two tasks. When asked to judge gaze direction the gaze shifted toward the eyes region (Fig. 2*A*,*B*), on the other hand, when judging facial expression, subjects spent more time on the mouth region (Fig. 4*C*,*D*, and Supplementary Video 2). Microsaccadic behavior was very consistent across subject; most of microsaccades landed on the eyes in the gaze direction task (0.70±0.13 on the eyes vs. 0

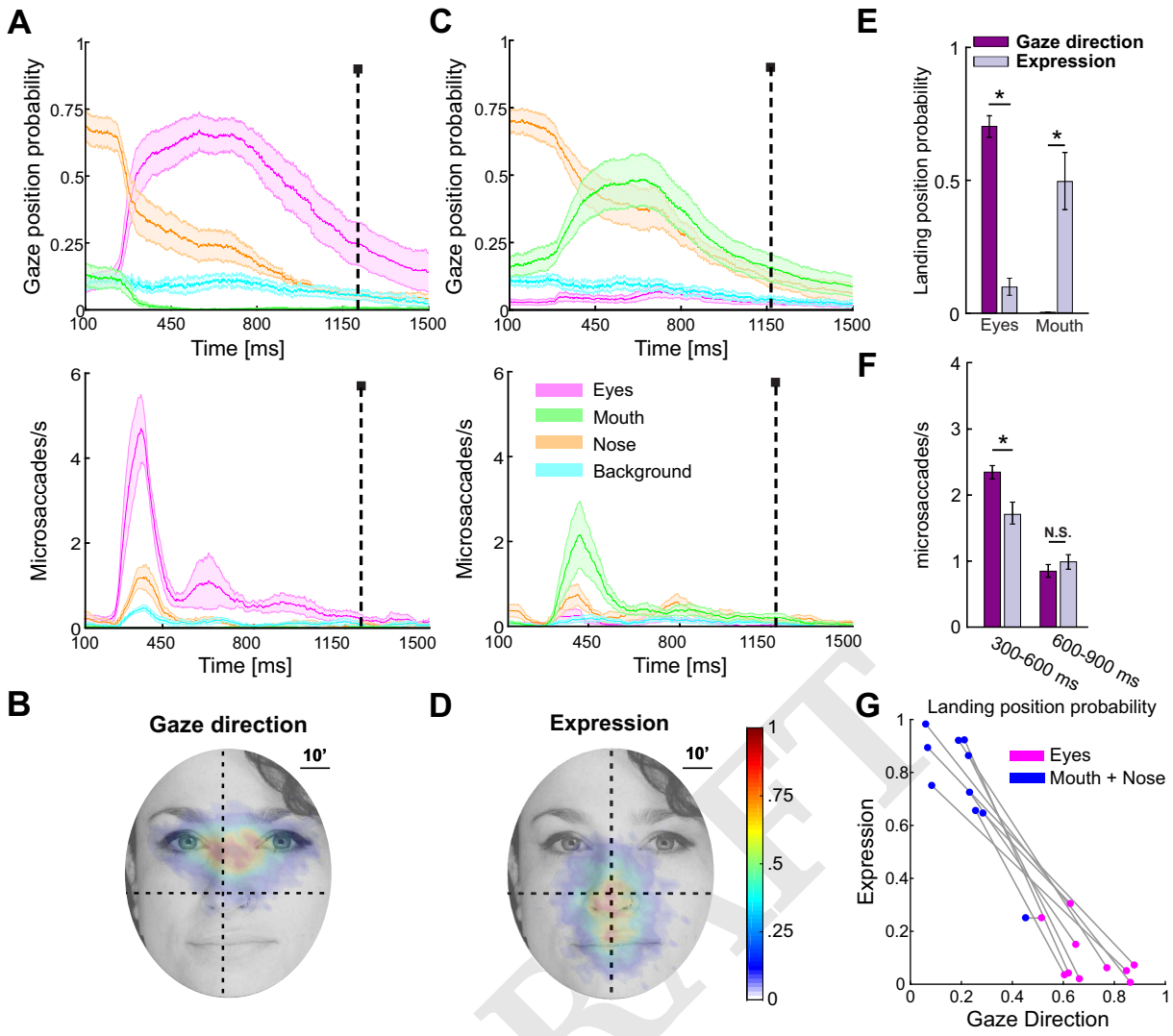

Fig. 2. Experiment 2 results. Average probability of gaze position distribution (*left*) and microsaccade landing position (*right*) in the gaze direction (**A**) and the expression (**C**) task (N=10). Data have been filtered using a running average with a 100 ms window. Dashed black lines mark the average response time. Shaded regions are s.e.m.. Average 2D normalized gaze distribution probability in the gaze direction (**B**) and in the expression (**D**) tasks. **E**, average probability of microsaccades landing on the eyes and on the mouth in the two tasks in the interval 300 ms to 600 ms from stimulus onset. **F**, average rate of microsaccades at the beginning and at the end of the trial for the two tasks. Asterisks mark a statistically significant difference ($p<0.05$, two-tailed paired t-test). **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 task.

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*E*,*G*).

The oculomotor behavior in both tasks differed compared to the normal physiological fixational instability when subjects maintained fixation on a single point. When maintaining fixation the amplitude of microsaccades was lower (16′±2′ in the task vs. 13′±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 the nose in the task (0.52±0.2 vs 0.2±0.3, 0.10±0.1, 0.14±0.07, for nose, mouth, eyes and background respectively). These findings show that, even during brief fixation periods, the visuomotor system does not simply maintain fixation on the foveated stimulus but it 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 the landing position of microsaccades was different based on the task performed, but their rate and time course also 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 ms/s ±0.6 ms/s, 1.7 ms/s ± 1 ms/s for gaze direction and expression respectively; $p=0.027$, two-tailed paired t-test. Fig 2*F*), but was virtually the same in the two tasks during the rest of the trial (600 ms-900 ms, 0.9 ms/s ±0.6 ms/s, 1.0 ms/s ± 0.7 ms/s for gaze direction and expression respectively; $p=0.3$, two-tailed paired t-test). Microsaccade time course was also modulated by the the task. The rate of microsaccades peaked approximately 90 ms earlier in the gaze direction task (327 ms±17 ms) compared to the face expression task (403 ms±80

ms, p=0.01, two-tailed paired t-test) and to a simple fixation (391 ms±49 ms, p=0.005, two-tailed paired t-test. Fig. 3).

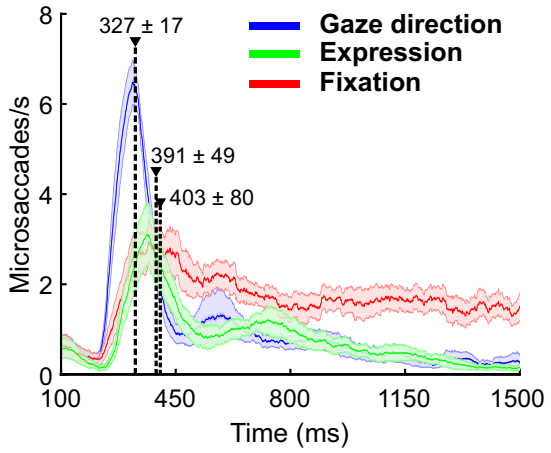

Fig. 3. Temporal occurrence of microsaccades. Average microsaccade rate over time in experiment 2 and during sustained fixation. Data have been filtered using a running average with a 100 ms window. Dashed lines reperesent the average time when the rate of microsaccades reached a peak. Error bars represent s.e.m..

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 visual exploration of larger stimuli. Subjects viewed human faces and judged whether or not the face's expression was neutral. To increase the difficulty of the task several facial expressions were ambiguous and the contrast of the stimuli was lowered. In the parafovea condition, each face covered an area of 11.5 deg^2, as if they were viewed from a distance of 3 meters. In the foveola condition, instead, faces covered an area of 0.7 deg^2, as if they were viewed from a distance of 13 meters (Fig. 4*A*).

When the stimulus extended to the parafoveal region, almost all observers followed a very stereotyped scanning pattern (Fig. 4*C*,*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. 4*C*) (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 HSD *post hoc* tests: mouth vs. eyes, p<0.0001, mouth vs. nose: p<0.0001, mouth vs. background: p<0.0001. See Supplementary Video 2). 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[12, 13, 15–17]. Moreover, a bias toward the lower part of the face when judging facial expression was also reported for the first saccade bringing the face within at the center of gaze[8].

In the foveola condition the exploratory behavior was driven by microsaccades (average amplitude 15′±3′, Supplementary Fig. 1). 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 ms±65 ms microsaccade rate peak time in the foveola condition vs. 403 ms±87 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 ms - 600 ms), most microsaccades landed on the mouth region (0.40±0.3 vs. 0.16±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 HSD *post hoc* tests: mouth vs. eyes, p=0.02, mouth vs. nose: p=0.3, mouth vs. background: p=0.09. Fig. 4*D*,*F*, Fig. 4*B*,*G* and Supplementary Video 2). 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 in fine spatial vision 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 ≈ 5′ away from the target feature, 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 to perform this task. Consistently 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.).

Crucially, microsaccades bringing 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; Supplementary Figure 2), 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

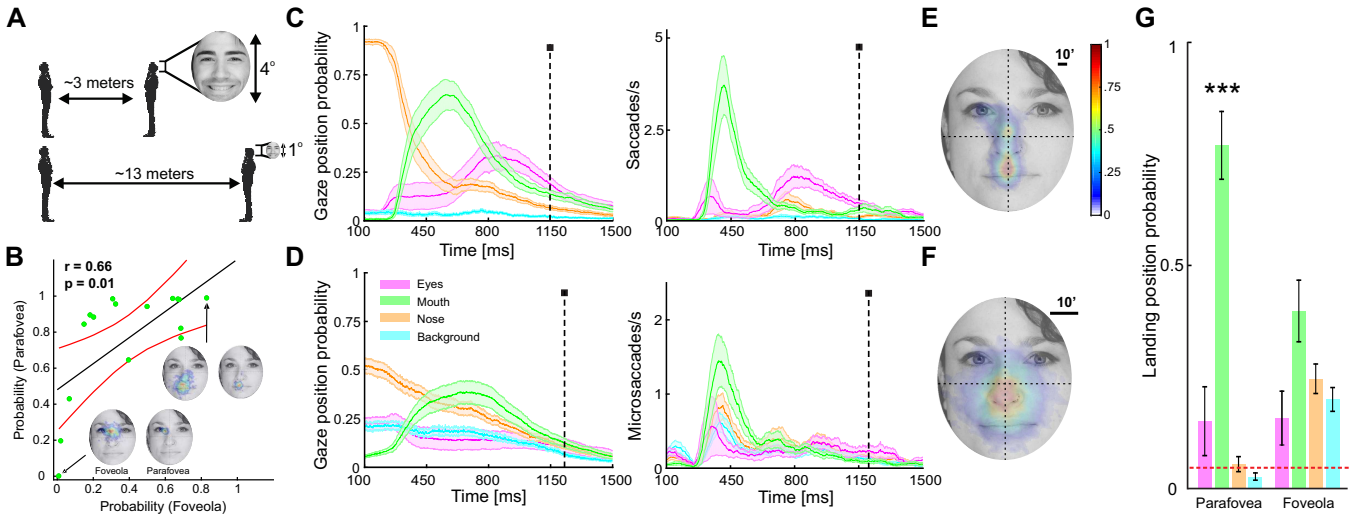

Fig. 4. A, faces are normally viewed from different distances; the face of a person standing ≈ 3 meters away and spans 4deg on the retina, but it spans only 1 deg when the observer is ≈ 13 meters away. In experiment 2 faces covered either an area of $0.7deg^2$ (foveola condition, 1 deg height), or they covered an area of $11.5deg^2$ (parafovea condition, 4.2 deg height). **B**, single subject probabilities of microsaccades (foveola) and saccades (parafovea) landing on the mouth region in the two conditions. Probabilities refer to the interval from 300 to 600 ms after the stimulus onset. The black line represents the linear fit of the data, and the red lines mark the 95% confidence intervals of the fit. The correlation value r and the p value of the correlation are also shown in the graph. Examples of gaze distribution in the two conditions are shown for two subjects. **C**, average distribution of gaze position (*left*) and saccade landing position (*right*) over time in the parafovea condition (N=16). **D**, 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. Black lines represent the average response time. Shaded regions are s.e.m.. Dashed black lines mark the average response time. Average 2D normalized gaze distribution probability in the parafovea (**E**) and in the foveola (**F**) condition. **G**, average probability of saccade (parafovea) and microsaccade (foveola) landing on different regions of the stimulus in the interval from 300 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 HSD *post hoc* tests). Error bars represent s.e.m..

instability of the eye at fixation, 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 to when the subjects performed the task (1.5 ms/s±0.8 m/s and 1.2ms/s±0.6 ms/s fixation and task respectively; p=0.04, paired two-tailed t-test. 13.6′±3.6′ and 15′±3′ fixation and task respectively; p=0.04, paired two-tailed t-test. Supplementary Figure 1). Moreover, microsaccades landing position and the overall spatial distribution of gaze position differed across the two conditions. As illustrated in Fig. 4*G* (red dashed line)and Supplementary Figure 3, 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 vs 0.35±0.1, 0.32±0.2, 0.27±0.1, for mouth, eyes, nose and background respectively. $p<0.0001$, Tukey HSD *post hoc* tests: mouth vs. eyes, $p<0.0001$, mouth vs. nose, $p<0.0001$, mouth vs. background, p=0.0006). Similarly 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 actively modulated by the task performed.

Interestingly, not only microsaccades were 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 to when they simply maintained fixation on a single point (diffusion coefficient at fixation 17 arcmin^2 ± 5 arcmin^2 vs 14 arcmin^2 ± 4.3 arcmin^2 in the foveola condition, p=0.009; Supplementary Figure 4). 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.

Individual differences are maintained across scales. It has been previously reported that the pattern of eye movements when viewing faces varies significantly across observers[10, 14, 15, 19, 20]. Similarly, here we found that in the parafovea condition a small percentage of subjects (24% of the total, 5 subjects) maintained fixation around the center of the display 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. 5*A*,*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. ??*B*). Similarly, also in the foveola con-

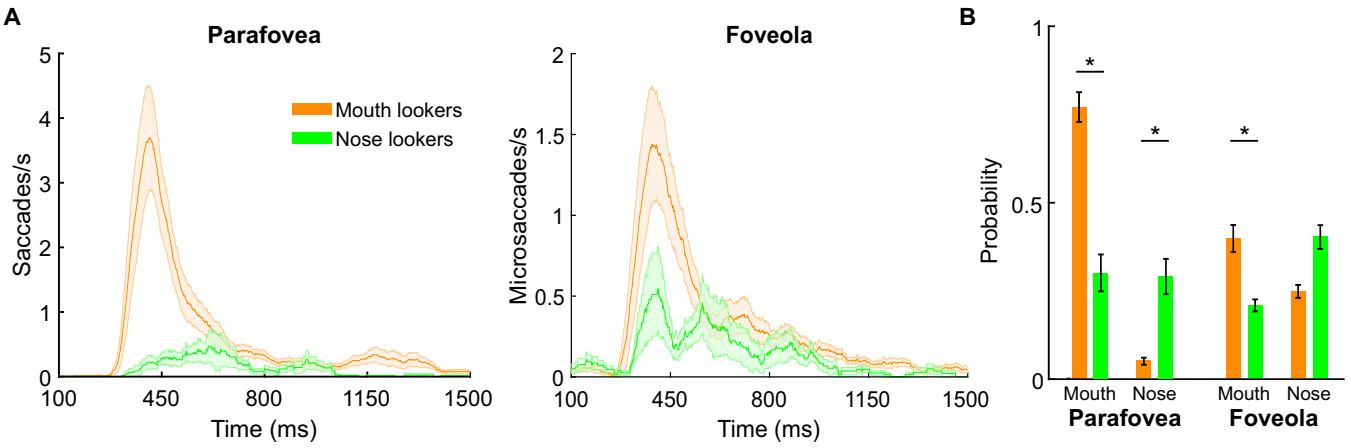

Fig. 5. Individual differences are maintained across scales. A, average rate of saccades (parafovea, *left*) and microsaccades (foveola, *right*) landing on the mouth during the course of the trial for nose lookers (N=5) and mouth lookers (N=16). **B**, probability of microsaccade and saccade landing 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 s.e.m..

dition 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 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

The existence of a continuum between saccades and microsaccades[21–24] raises the question of whether scanning and exploration of visual objects and scenes, which has traditionally been ascribed to large saccades, also applies to microsaccades at a smaller spatial scale. This question has, however, remained unanswered due to technical limitations; localizing the gaze with high precision within the small portion of the visual field projecting onto the foveola is extremely challenging. Our work circumvented these limitations and addressed this open issue. Here we show that visual exploration extends to the scale of the foveola. More specifically, visual exploration of complex foveal stimuli follows the same patterns unfolding when examining scenes at a larger scale. Microsaccades consistently target foveal locations containing task-relevant information, and their rate and temporal dynamics are modulated by the goals of the task. Importantly, this study complements the findings of our previous work showing that microsaccades are finely tuned to precisely re-center high-acuity stimuli on a preferred locus of fixation where fine pattern vision is highest[7]; this oculomotor behavior is not simply the outcome of a purely bottom-up driven re-centering mechanism, but it is the manifestation of active, top-down driven, visual scanning strategies.

When examining whether microsaccades are influenced by cognitive/attentional factors, previous research has mostly focused on how the pattern of microsaccades changes based on the peripheral allocation of covert attention[25–27]. Crucially, these findings pointed out the need to control for these small gaze shifts when manipulating covert attention. However, differently from normal viewing conditions, in the spatial cuing paradigms used to study covert attention, visual stimulation at the center of gaze is minimal. This prompts at least two questions; first whether in more natural conditions, when foveal stimulation is rich of details, microsaccades are still modulated by the peripheral allocation of attention. Second, whether allocating attention peripherally inevitably leads to a suppression of visual scanning behavior of foveal detail. Previous work indicated that analysis of foveal stimuli proceeds in parallel and independently from selection of the next saccade target[28], suggesting that allocating attention peripherally may not necessarily interfere with the examination of complex foveal stimuli. Addressing these questions is fundamental for a better understanding of the interplay of attention and eye movements in more ecological conditions, when both foveal and peripheral processing is required during the timeframe of one fixation.

Our work has important implications for the study of priority maps. In the experimental paradigm used here subjects were free to perform multiple saccades and stimulus presentation was relatively long. Yet, in experiment 1 subjects delivered their response about 200 ms before the offset of the stimulus, and in most trials subjects performed only one microsaccade before responding (1.5±0/.6 microsaccades, and 1.2±0.7 microsaccades in the gaze direction and expression task respectively). Thus, the first microsaccade after the onset of the stimulus was the most critical for performing the task and the facial expression judgment was formed shortly afterwards. The first microsaccade, generally happening within the first 350 ms after the stimulus onset, was also clearly driven by the goal of the task. On the other hand, microsaccades occurring after ≈500 ms were much less pulled toward one single feature of the face. Whereas priority maps are generally thought to represent the relevance of stimuli in the extrafoveal space, our findings strongly suggest that the first microsaccade executed after the stimulus onset was driven by a priority representation of the foveal input. As soon as a stimulus is presented

foveally, the system determines what region in the foveal space contains the most relevant information to perform the task and it guides microsaccades over that region. Because visual capabilities are not uniformly distributed across the foveola[7], establishing a high-resolution priority map of the most relevant parts of the foveal landscape to guide visual exploration at this scale could be beneficial. Notably, priority maps are used to drive different effectors[29, 30], from eye movements to reaching, however, saccades are the only motor behavior that can be controlled at a fine scale. This raises the question of whether a finer grain priority map of the foveola would be specifically limited to the guidance of microsaccades, while coarser priority maps of the extrafoveal space can, instead, be used to guide multiple effectors. Further work is necessary to investigate foveal priority maps, their influences on attention and visual perception at the foveal scale.

Our work also shows that individual differences in visual exploration are maintained across spatial scales. The existence of individual differences in visual scanning strategies is known; visual scanning strategies used to explore faces do not only change according to the task goal, a number of studies have reported the existence of significant variations in these strategies across individuals[10, 14, 15, 20]. These idiosyncrasies are maintained over time[14], and they do not change even when central vision is blocked using an artificial scotoma[31]. More specifically, a difference between nose lookers and mouth lookers has been reported previously when examining 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 land closer to the nose region while other land closer to the eye region. The experimental paradigm used here differs in a number of ways from the paradigm used in these studies; our stimuli were considerably smaller both in the foveola and in the parafovea condition, and were presented centrally for a relatively long period of time. Yet, we reported a similar difference in the oculomotor behavior across subjects; while most of the observers explored the face, primarily looking at the mouth, some observers kept fixation on the nose. These findings show that idiosyncrasies are also maintained across different spatial scales. It is possible that these strategies reflect anatomical differences in the structure of the retina and the foveola itself. Indeed, it has been shown that cone density across the foveola[32, 33] and the size of the foveal pit vary greatly across subjects[34].

Ultimately this work shows that fine oculomotor behavior is much more complex than previously thought. Contrary to the common assumption, foveating the stimulus of interest is necessary but not sufficient. During fixation the visuomotor system engages in a subtler level of visual examination. Microsaccades are efficiently used to guide visual exploration of the foveal landscape, sampling with the preferred locus of fixation the most informative foveal locations.

Materials and Methods

Observers: A total of 31 emmetropic human observers, all naïve about the purpose of the study, participated in the experiments (age range 18-25). 21 observers (17 males and 4 females) took part in Experiment 2 (Fig. 4), and 10 (4 males and 6 females) in Experiment 1 (Fig. 2). Informed consent was obtained from all participants following procedures approved by the Boston University Charles River Campus Institutional Review Board.

Stimuli and apparatus: Stimuli were displayed on a fast-phosphor CRT monitor (Iyama HM204DT) at a vertical refresh rate of 85 Hz and spatial resolution of 2048×1536 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) eyetracker (Fourward Technologies), a system with an internal noise of ∼20” and a spatial resolution of 1′[5, 35]. Vertical and horizontal eye positions were sampled at 1 kHz and recorded for subsequent analysis.

Stimuli were rendered by means of EyeRIS[36], 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 eyetracker, 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[37, 38]. The images used were prelabeled 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 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 to 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 it is customary in oculomotor experiments. These points were located 1.32^o 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 joypad 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^o, 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 were 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^o, whereas in the foveola condition it measured 1^o. 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 seconds and subjects responded by pressing a button on a remote controller at any time during stimulus presentation and for a period of 4 seconds after the stimulus was turned off. 1.5 seconds 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 lab personnel with extensive experience in classifying eye movements. Periods of blinks were automatically detected by the DPI eyetracker 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^o/s were selected as saccadic events. Saccades with an amplitude of less than half a degree ($30'$) were defined as microsaccades. Consecutive events closer than 15 ms were merged together into a single saccade in order to automatically exclude post-saccadic overshoots[39, 40]. Saccade amplitude was defined as the vector connecting the point where the speed of the gaze shift grew greater than 3^o/s (saccade onset) and the point where it became less than 3^o/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%, 4.9% of the total trials for parafovea 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. All figures show average values for each individual observer and summary statistics across observers. All data will be made available upon reasonable request.

ACKNOWLEDGMENTS. This work has been supported by the National Science Foundation grant NSF-BCS-1534932 to M.P..

1. Buswell GT (1935) *How people look at pictures.* (The University of Chicago Press, Illinois).
2. Yarbus AL (1967) *Eye Movements and Vision.* (Plenum Press, New York).
3. Azzopardi P, Cowey A (1193) Preferential representation of the fovea in the primary visual cortex. *Nature* 361(6414):719–721.
4. Rolfs M (2009) Microsaccades: Small steps on a long way. *Vision Research* 49(20):2415–2441.
5. Poletti M, Rucci M (2016) A compact field guide to the study of microsaccades: Challenges and functions. *Vision Research* 118:83–97.
6. Kuang X, Poletti M, Victor JD, Rucci M (2012) Temporal encoding of spatial information during active visual fixation. *Current Biology* 20(6):510–514.
7. Poletti M, Listorti C, Rucci M (2013) Microscopic eye movements compensate for nonhomogeneous vision within the fovea. *Current Biology* 23(17):1691–1695.
8. Peterson MF, Eckstein MP (2012) Looking just below the eyes is optimal across face recognition tasks. *Proceedings of the National Academy of Sciences of the United States of America* 109(48):E3314–23.
9. Hsiao JH, Cottrell G (2008) Two fixations suffice in face recognition. *Psychological Science* 19(40):998–1006.
10. Kanan C, N.F BD, Ray NA, Hsiao JH, Cottrell GW (2015) Humans have idiosyncratic and task-specific scanpaths for judging faces. *Vision Research* 108:67–76.
11. Williams CC, Henderson JM (2007) The face inversion effect is not a consequence of aberrant eye movements. *Memory Cognition* 35(8):1977–1985.
12. Schurgin MW, et al. (2014) Eye movements during emotion recognition in faces. *Journal of Vision* 14(13):1–16.
13. Smith ML, Cottrell GW, Gosselin F, Schyns PG (2006) Transmitting and decoding facial expressions. *Psychological Science* 16(3):184–189.
14. Peterson MF, Eckstein MP (2013) Individual differences in eye movements during face identification reflect observer-specific optimal points of fixation. *Psychological Science* 33:1216–1225.
15. 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. *Journal of Vision* 14(7):1–10.
16. 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(2):e31106.
17. Blais C, Jack RE, Scheepers C, Fiset D, Caldara R (2008) Culture shapes how we look at faces. *PLoS One* 3(8):e3022.
18. Michele R, Iovin R, Poletti M, Santini F (2007) Miniature eye movements enhance fine spatial detail. *Nature* 447(7146):852–855.
19. Perlman SB, et al. (2009) Individual differences in personality predict how people look at faces. *PLoS One* 4(6):1–6.
20. Peterson MF, Lin J, Zaun I, Kanwisher N (2016) Individual differences in face-looking behavior generalize from the lab to the world. *Journal of Vision* 16(7):1–18.
21. Otero-Millan J, Macknik SL, Langston RE, Martinez-Conde S (2013) An oculomotor continuum from exploration to fixation. *Proceedings of the National Academy of Sciences of the United States of America* 110(15):6175–6180.
22. Zuber BL, Stark L, Cook G (1965) Microsaccades and the velocity-amplitude relationship for saccadic eye movements. *Science* 150(3702):1459–1460.
23. Ko HK, Poletti M, Rucci M (2010) Microsaccades precisely relocate gaze in a high visual acuity task. *Nature Neuroscience* 13(12):1549–1553.
24. Hafed ZM, Goffart L, Krauzlis RJ (2009) A neural mechanism for microsaccade generation in the primate superior colliculus. *Science* 323(5916):940–943.
25. Hafed ZM, Clark JJ (2002) Microsaccades as an overt measure of covert attention shifts. *Vision Research* 42(22):2533–2545.
26. Engbert R, Kliegl R (2003) Microsaccades uncover the orientation of covert attention. *Vision Research* 43(9):1035–1045.
27. Yuval-Greenberg S, Merriam EP, Heeger DJ (2014) Spontaneous microsaccades reflect shifts in covert attention. *Journal of Neuroscience* 34(41):13693–13700.
28. Ludwig CJ, Davies JR, Eckstein MP (2014) Foveal analysis and peripheral selection during active visual sampling. *Proceedings of the National Academy of Sciences of the United States of America* 111(2):E291–299.
29. Bisley JW, Goldberg ME (2010) Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience* 9(3):1–21.
30. Zelinsky GJ, Bisley J (2015) The what, where, and why of priority maps and their interactions with visual working memory. *Annals of the New York Academy of Sciences* 1339(1):154–164.
31. Tsank Y, Eckstein MP (2017) Domain specificity of oculomotor learning after changes in sensory processing. *Journal of Neuroscience* 37(47):11469–11484.
32. Curcio CA, Sloan KR, Kalina RE, Hendrickson AE (1990) Human photoreceptor topography. *Journal of Comparative Neurology* 292(4):497–523.
33. Li KY, Tiruveedhula P, Roorda A (2010) Intersubject variability of foveal cone photoreceptor density in relation to eye length. *Investigative Ophthalmology & Visual Science* 51(12):6858–6867.
34. Tick S, et al. (2011) Foveal shape and structure in a normal population. *Investigative Ophthalmology & Visual Science* 52(8):5105–5110.
35. Ko Hk, Snodderly DM, Poletti M (2016) Eye movements between saccades: Measuring ocular drift and tremor. *Vision Research* 122:93–104.
36. Santini F, Redner G, Iovin R, Rucci M (2007) EyeRIS: A general-purpose system for eye movement contingent display control. *Behavioral Research Methods* 39(3):350–364.
37. Minear M, Park DC (2004) A lifespan database of adult facial stimuli. *Behavior Research Methods, Instruments, & Computers* 36(4):630–633.
38. Vieira TF, Bottino A, Laurentini A, Simone MD (2014) Detecting siblings in image pairs. *The Visual Computer* 30(12):1333–1345.
39. Deubel H, Bridgeman B (1995) Fourth purkinje image signals reveal eye-lens deviations and retinal image distortions during saccades. *Vision Research* 35:529–538.
40. Stevenson SB, Roorda A (2005) Correcting for miniature eye movements in high resolution scanning laser ophthalmoscopy in *Ophthalmic Technologies XV*. pp. 145–151.