Perceptual saccadic suppression starts in the retina

Saad Idrees[1]†, Matthias-Philipp Baumann[1,2]†, Felix Franke[3], Thomas A. Münch[1,4]∗, Ziad M. Hafed[1,2]∗

[1]Werner Reichardt Centre for Integrative Neuroscience, Tübingen University, Tübingen, Germany 72076
[2]Hertie Institute for Clinical Brain Research, Tübingen University, Tübingen, Germany 72076
[3]Bio Engineering Laboratory, ETH Zürich, Basel, Switzerland 4058
[4]Institute for Ophthalmic Research, Tübingen University, Tübingen, Germany 72076

∗Correspondence to:
ziad.m.hafed@cin.uni-tuebingen.de and thomas.muench@uni-tuebingen.de

†Equal contributions

∗Co-corresponding authors

Corresponding author addresses:

Ziad M. Hafed
Werner Reichardt Centre for Integrative Neuroscience
and
Hertie Institute for Clinical Brain Research
Otfried-Mueller Str. 25
Tuebingen, 72076
Germany
Phone: +49 7071 29 88819

and

Thomas A. Münch
Werner Reichardt Centre for Integrative Neuroscience
Otfried-Mueller Str. 25
Tuebingen, 72076
Germany

Abstract

Visual sensitivity is strongly impaired around the time of rapid eye movements. This robust perceptual phenomenon, called saccadic suppression, is frequently attributed to active suppressive signals that are directly derived from eye movement commands. Here we show instead that visual-only mechanisms, activated by saccade-induced image shifts, can account for all perceptual properties of saccadic suppression that we have investigated. Such mechanisms start at the very first stage of visual processing in the brain, the retina. Critically, suppression originating in the retina outlasts perceptual suppression around the time of saccades, suggesting that extra-retinal movement-related signals, rather than causing suppression, may instead act to shorten it. Our results demonstrate a far-reaching contribution of visual processing mechanisms to perceptual saccadic suppression, starting in the retina, without the need to invoke explicit motor-based suppression commands.

Introduction

Saccadic eye movements are a prominent feature of visual behavior because they allow successive sampling of visual information from the environment. However, from the perspective of the flow of visual information into the brain, these rapid eye movements constitute highly disruptive visual events, introducing spurious motions that should normally go unnoticed, or get canceled, at the level of perception. The question of how and why such perceptual cancelation takes place across saccades has intrigued philosophers and scientists for many decades[1–4]. Indeed, visual sensitivity to brief visual probes is strongly impaired around the time of saccades, in a phenomenon known as saccadic suppression that has repeatedly been demonstrated in a multitude of experiments[5–15].

Despite the robustness of saccadic suppression as a perceptual phenomenon, the mechanisms behind it remain highly controversial. On the one hand, perceptual saccadic suppression may arise through internal knowledge of planned eye movements and their associated motor commands[5,12,16–19]. According to this popular view, internal knowledge of eye movement commands is a necessary prerequisite for saccadic suppression: a movement-related signal[17,18], such as corollary discharge from (pre-)motor areas, may act as a suppressive command for visual neurons to cause perceptual suppression, and maybe even in a pathway-selective manner[10].

On the other hand, perceptual saccadic suppression could alternatively, or additionally, arise as a result of the visual consequences of retinal image shifts caused by eyeball rotations[2,20-31]. After all, the early visual system, including the retina, is a highly sensitive light sensing device, and it therefore ought to detect visual transients associated with saccade-induced retinal-image shifts. Such early detection

of visual transients could jumpstart an image processing cascade to mediate perceptual saccadic suppression.

In this study, rather than arguing either strictly for or strictly against one of the above two seemingly contrasting hypotheses, we instead asked to what extent they might interact with and support each other for the ultimate service of perception. We were specifically motivated by the fact that the very first visual processing stage in the brain, the retina, is not only sensitive to visual transients (such as saccade-induced image shifts), but it also possesses rich image processing circuitry that is capable, in principle, of regularizing the visual disruptions[32–37] caused by saccades. We therefore asked: how much of the characteristics of perceptual saccadic suppression can be explained by visual-only mechanisms? And, to the extent that there are visual-only mechanisms underlying perceptual saccadic suppression, would the first neural locus for such visual-only mechanisms indeed be the very first stage of visual processing in the brain, the retina?

We used a multi-disciplinary approach in which we experimentally mimicked the visual consequences of saccades and recorded neural activity from *ex vivo* retinae of different animal models. We also measured perceptual reports in humans using both real saccades as well as simulated saccade-like image displacements. We found a surprisingly far-reaching contribution of visual processing mechanisms to perceptual saccadic suppression, starting in the retina, without the need to invoke explicit motor-based suppression commands. Intriguingly, the role of motor-based commands seems to be the opposite of what has been proposed before. Rather than sending an explicit suppressive command to reduce the sensitivity of the visual system, motor-

based commands instead seem to minimize the duration of visually-derived saccadic suppression.

Results

Perceptual saccadic suppression depends on image content

We first asked human subjects to generate saccades across textured backgrounds, akin to how saccades may be made in real life. Subjects viewed coarse or fine textures (Fig. 1a, Methods and Supplementary Fig. 1). Starting from one of four locations on the display, subjects made 4.8 deg saccades towards display center (Fig. 1a, left). We varied saccade onset and endpoint locations, as well as texture images, across trials to avoid subjects remembering specific texture patterns (Methods). At a random time, a luminance pedestal (probe flash) was added to the texture background, for one display frame (approximately 12 ms; Methods), at one of four locations relative to the saccade endpoint (7 deg eccentricity; Fig. 1a, right). At trial end, the subjects were asked to localize the probe flash, and we analyzed how well they did so. We took care to ensure that the retinal region of flash location was stimulated with the background texture (rather than the edge of the monitor or the black surround of the dark laboratory) throughout any given trial (Methods). We also ensured that flash size was larger than the image blobs in the coarse texture, such that average luminance variation within each flash was matched across trials and textures. Coarse and fine textures had blobs that approximated the sizes of retinal ganglion cell (RGC) or retinal bipolar cell receptive fields, respectively, at the retinal flash locations[38] (Methods).

For both coarse and fine textures, subjects could not localize flashes presented peri-saccadically, thus experiencing strong perceptual saccadic suppression (Fig. 1b, c).

However, there was a clear dependence of the suppression on the background visual image: saccadic suppression started significantly earlier and recovered significantly later with saccades across coarse rather than fine textures (Fig. 1d; the highlighted time intervals show significant differences between coarse and fine textures at a p-value of $p<0.001$, cluster-based random permutation test[39,40]; Methods). Moreover, the peak amount of suppression was stronger with the coarse than the fine textures (Fig. 1d), although performance reached a floor effect in both conditions in this version of the experiment, masking a larger difference (see below and Fig. 2). This dependence of perceptual saccadic suppression on background texture was robust across individual subjects (Supplementary Fig. 2a; also see Supplementary Fig. 4 for further individual subject effects).

To rule out the possibility that the difference in perceptual saccadic suppression profiles between the coarse and fine textures was due to the flashes being simply easier to see over the fine texture, we performed a control experiment in which we collected full psychometric curves of perceptual performance during simple fixation. We found that, without any saccades, the visibility of the probe flashes was identical over coarse and fine background textures (Supplementary Fig. 3a, b). Therefore, the image dependence of the results of Fig. 1 was related to saccadic suppression itself and not to the baseline visibility of brief flashes over the different textures. Similarly, we carefully analyzed eye movement properties, and we found that the results of Fig. 1 were also not due to different saccade kinematics for the different textures (Supplementary Fig. 3c, d).

To further explore the differences in suppression profiles that we observed in Fig. 1, we next employed a more sensitive procedure to evaluate perceptual thresholds.

Specifically, we repeated the same experiment of Fig. 1 on five subjects (three of whom being the same as those who participated in the experiment of Fig. 1). This time, we collected full psychometric curves of perceptual performance (Methods; similar to Supplementary Fig. 3a, b). Because collecting full psychometric curves for each texture and each time point relative to saccade onset would be a very data-intensive endeavor, we reduced the number of time points relative to saccade onset at which we probed perception. We also expedited the data collection by implementing a real-time saccade detection algorithm, described by Chen and Hafed[41], and we presented the probe flash at four distinct times after online saccade detection. The four flash times were strategically chosen to evaluate peak suppression (shortly after saccade onset) as well as the time course to recovery long after a saccade. We used an adaptive QUEST[42] procedure to estimate the perceptual threshold per condition and flash time (Methods), with the perceptual threshold being defined as the flash contrast value resulting in 62.5% correct performance. Besides the QUEST procedure, we also collected more trials showing different flash contrast levels relative to perceptual threshold, in order to obtain full psychometric curves. The results are shown in Fig. 2, and they match those of Fig. 1: relative to the baseline psychometric curves of flash visibility long after saccades (dashed curves), peri-saccadic psychometric curves were clearly shifted towards higher contrast thresholds (Fig. 2a-d), consistent with Fig. 1. More importantly, with the more sensitive approach of full psychometric curves, we could clearly see that perceptual saccadic suppression was much stronger for coarse than fine textures at peak suppression; that is, perceptual thresholds near peak suppression were higher for coarse than fine textures (Fig. 2e). Supplementary Fig. 4 shows the corresponding individual subject psychometric curves and perceptual thresholds.

In summary, we found that perceptual saccadic suppression was associated with a visual component directly influencing its strength and time course: saccades across coarse textures were associated with both stronger and longer-lasting perceptual suppression than saccades across fine textures, even when the kinematics of the eye movements (and thus the underlying motor commands) did not differ across the two conditions.

Perceptual saccadic suppression originates in the retina

To test if this visual component of perceptual saccadic suppression originates in the retina, we isolated mouse and pig retinae and performed multi-electrode array recordings (Methods). We visually exposed each retina to coarse and fine textures, matched to ganglion and bipolar cell receptive field sizes in the recorded species (Methods, Supplementary Fig. 1). We then rapidly translated the textures globally to simulate saccade-like image displacements (Fig. 3a, Methods). Such displacements can robustly activate RGCs, as is evident from the example mouse RGC shown in Fig. 3b. In fact, most recorded RGCs (mouse: 83% of 1,423 cells, pig: 73% of 394 cells) responded robustly to texture displacements, indicating that saccade-induced visual transients during active gaze behavior can constitute strong signals to the retina. Next, at different times relative to texture displacements, we introduced a luminance pedestal (probe flash) to the entire texture for 16 or 33 ms, similar in principle to the perceptual experiments of Figs. 1, 2. Such flashes, when presented in isolation (that is, temporally removed from the texture displacement; baseline), elicited responses in a sizable fraction of RGCs (mouse: 688 of 1,423 RGCs; pig: 228 of 394 RGCs). This allowed us to evaluate the consequences of texture displacements on flash responses in these cells, in a way that is conceptually similar to the experiments in Figs. 1, 2, in which we evaluated the consequences of

saccades on flash perception. For example, the same RGC of Fig. 3b showed much suppressed neural responses to the flash when it was presented immediately after texture displacements compared to the baseline condition (Fig. 3c, d). These suppressed flash-induced responses (Fig. 3d) look remarkably similar to suppressed visual responses in, say, macaque superior colliculus for stimuli presented after real saccades [13,15,43]. Thus, neuronally, there does exist "saccadic suppression" of visual sensitivity at the very first stage of visual processing in the brain, the retina, and it looks qualitatively indistinguishable from saccadic suppression at downstream neural sites[13,15,43] and, indeed, perception (Figs. 1, 2).

Importantly, retinal "saccadic suppression" strongly depended on background texture (Fig. 3e), exactly like in our human experiments (Figs. 1, 2). Specifically, we quantified retinal "saccadic suppression" by calculating a neuronal modulation index, defined as $(r_d - r_b)/(r_d + r_b)$. Here, r_d is the response strength to the probe flash presented with a delay d relative to the texture displacement onset, and r_b is the response strength in baseline (Methods). This modulation index is, by definition, negative for suppressed flash-induced responses. The great majority of RGCs were strongly suppressed during and after texture displacements, with gradual recovery afterwards (Fig. 3e; Supplementary Fig. 5 shows the underlying population data), and the suppression was more pronounced for coarse than fine textures (Fig. 3e; Supplementary Fig. 5). These results are consistent with the dependence of human perceptual saccadic suppression on background texture statistics shown above (Figs. 1, 2), suggesting that this dependence starts already in the retina.

We also found that retinal "saccadic suppression" was a robust phenomenon across many different RGC cell types (Supplementary Fig. 6). Further, it occurred both in

mouse (Fig. 3e, left) and pig (Fig. 3e, right) retinae, two mammalian species with different native oculomotor behavior, different lifestyles, and different eye sizes. Thus, our results so far suggest that perceptual saccadic suppression (Figs. 1, 2), including its dependence on background texture statistics, most likely originates in the retina (Fig. 3), being the outcome of very general retinal-circuit mechanisms that are conserved across species.

Stimulus-stimulus interactions underlie retinal suppression

To understand the underlying mechanisms for retinal “saccadic suppression” in more detail, we explored its properties using different analyses and additional stimulus manipulations. First, we wondered about neural activity saturation, given that saccade-like texture displacements before flash onset could activate RGCs (e.g. Fig. 3b). Specifically, if RGC activity is elevated by the texture displacement alone (because it was a visual transient), then any subsequent flash-induced response could have caused the cell to reach activity saturation. However, this was not sufficient to fully explain our results. For example, we observed that suppression often also occurred in RGCs that did not respond strongly to the texture displacements in the first place (Fig. 4a).

Second, we investigated whether retinal “saccadic suppression” critically depended on particular saccade-like profile speeds. In the original experiments of Fig. 3, we simulated saccade-induced image translation speeds to the best of our abilities (given the sampling rate of our display; Methods). However, if we replaced the original translation over 100 ms with a sudden texture jump from the start- to end-position in one display update (an infinite-speed texture jump), then the same suppression took place, with similar dependence on texture statistics (Fig. 4b).

Similarly, in yet another manipulation, when we presented first a flash and then a texture displacement, then the second response (now to the texture displacement) was suppressed (Fig. 4c). This suggests that retinal “saccadic suppression” can be explained by general stimulus-stimulus interaction effects in the retina. As a result, it is a phenomenon that is unlikely to critically depend, at least qualitatively, on the specific oculomotor repertoire of either mice, pigs, or humans.

The most compelling evidence for stimulus-stimulus interactions underlying retinal “saccadic suppression” came from experiments when we replaced the texture displacements with a structure-free contrast step (Fig. 4d). Specifically, instead of a background texture and a displacement of this texture, we exposed the retina to a uniform gray background and introduced a sudden uniform luminance increase or decrease as the visual transient. This luminance step was either of high contrast (+/- 0.20 to 0.40 Michelson contrast) or low contrast (+/- 0.03 to 0.15 Michelson contrast) (Methods). The probe flash then followed the contrast step as in the original experiments. Our prediction was that a high-contrast visual transient would be equivalent to the displacement of a coarse texture across a given RGC receptive field (e.g. Fig. 3), while a low-contrast transient would be equivalent to the displacement of the fine texture across the same receptive field (simply due to the motion energy expected with each type of texture displacement). This was indeed the case (Fig. 4d). Both the time courses and profiles of retinal suppression effects were similar for low-contrast steps and displacements of fine textures. Similarly, the suppression effects were similar for high-contrast steps and displacements of coarse textures.

Therefore, the most likely mechanism for our retinal “saccadic suppression” effect is that such suppression emerges as a result of specific spatio-temporal retinal-circuit

image processing that is initiated by visual transients; whether they be through texture displacements, infinite-speed texture jumps, or contrast steps (Fig. 4e). It is very intriguing that such stimulus-stimulus retinal effects may be inherited all the way deep into the brain's visual processing hierarchy, including cortical (frontal eye field) and subcortical (superior colliculus) areas[44] that are also implicated in saccadic suppression[15,43,45,46].

Motor-related signals shorten visually-derived suppression

In retina, we not only observed similarities to perceptual saccadic suppression (the presence of retinal suppression, and its dependence on texture statistics), but we additionally noticed that retinal "saccadic suppression" was particularly long lasting (e.g. Fig. 3e). To explore the potential perceptual implications of this observation, we then asked our human subjects to maintain fixation while we introduced saccade-like texture displacements in a manner similar to the retinal experiments of Fig. 3 (Fig. 5a, Methods); brief flashes occurred around the time of these "simulated saccades" exactly like in the first experiment (Fig. 1). This time, due to the absence of real saccades (trials with microsaccades were excluded; Methods), non-visual (motor-related) components could not influence flash-induced neural responses and their perception. Still, given the retinal results of Figs. 3, 4, we had three hypotheses on what to expect under these conditions, all of which we validated: (1) strong perceptual suppression still occurred regardless of texture details (Fig. 5b, c); (2) suppression strength and duration depended on texture statistics (Fig. 5d); and (3) suppression outlasted the suppression with real saccades (Fig. 5e, f). This last point, in particular, suggests that motor-related signals associated with real saccades may act to shorten the perceptual interruption resulting from visually-induced saccadic suppression, while maintaining the putatively retinally-determined (Figs. 3, 4)

dependence on image statistics. Note also that the first and third points above are consistent with earlier perceptual results shown by Diamond et al[17], although these earlier observations were later interpreted in the literature[18] as suggesting motor- rather than visually-derived saccadic suppression.

In humans, we observed perceptual suppression also prior to saccade-like texture displacements[20,24] (Fig. 5). This was again consistently dependent on texture statistics (Fig. 5b-d; also see Fig. 6 below for additional evidence). Critically, compared to simulated saccades, making real saccades also shortened such pre-saccadic perceptual suppression (Fig. 5e). Even in our retinal data, we found very slight “pre-saccadic” suppression. However, the effect size of suppression before texture displacements was much smaller than after texture displacements: the strongest “pre-saccadic” retinal effect occurred at -67 ms with a median population modulation index of -0.024 ($p = 6 \times 10^{-8}$, Wilcoxon signed-rank test) compared to -0.55 ($p = 3 \times 10^{-82}$) for “post-saccadic” suppression at 150 ms delay (Fig. 3e, Supplementary Fig. 5b). It is therefore likely that this particular phenomenon of perceptual pre-saccadic suppression (Fig. 5b-f) arises not in the retina, but from visual (not movement-command-related) processing further downstream, perhaps through backwards masking[26,47].

Since the results of Fig. 5 did not explicitly report perceptual thresholds, we also repeated the same experiment again, but this time using the same QUEST and full psychometric curve procedures described above for Fig. 2. In the current experiment, we again picked 4 specific time points relative to texture displacement onset for calculating perceptual thresholds (Methods). Like in the case of Fig. 2, we chose these 4 time points strategically to highlight perceptual threshold elevations at

maximal suppression and also to highlight differences between coarse and fine textures. We also explicitly sampled a negative time point close to texture displacement onset, such that we could fill in the gap in the negative time courses shown in Fig. 5. The net conclusion (Fig. 6) was the same as that in Fig. 5. There was robust elevation of perceptual thresholds before, during, and after the texture displacements. Most importantly, the elevation was much stronger and longer-lasting (both before and after texture displacements) for coarse than for fine textures. The effect was also robust across individual subjects (Supplementary Fig. 7).

Therefore, the long-lasting suppression effects that we observed in RGCs (Figs. 3, 4) were not an idiosyncrasy of the *ex vivo* electrophysiological procedures that we used, but they were reflected in the longer duration of perceptual suppression after simulated saccades. Importantly, they were indicative of a potential shortening of visually-derived suppression in association with real saccades.

Visually-derived suppression underlies even more phenomena

Our results so far suggest that visual contributions can go a very long way in explaining properties of perceptual saccadic suppression, without the need for invoking mechanisms related to motor commands. We therefore wondered whether such contributions can also explain classic suppression phenomena in experiments when uniform, rather than textured, backgrounds are used. One such robust phenomenon has been the selective suppression of low spatial frequencies. In a classic study by Burr et al[10], subjects viewed briefly flashed Gabor gratings over a uniform background. Around the time of saccades, visibility of low-spatial frequency gratings was suppressed much more strongly than of high-frequency gratings, and

this was interpreted as a motor-related influence on magnocellular pathways[17,18]. Still, convincing neural mechanisms for this phenomenon remain elusive[15,27,28,30,48–53]. Can the strong prominence of visual contributions to saccadic suppression revealed in our results so far also be extended to account for this classic phenomenon? In other words, is the selective suppression of low spatial frequencies around the time of saccades[10] intrinsically a visual, rather than motor, phenomenon?

The answer lies in considering this phenomenon from the perspective of visual input during such experiments: saccades across a uniform background invariably involve moving the image of the video monitor (or other form of display) in visual coordinates. Therefore, the image of any edge discontinuity associated with the display monitor (or with the surrounding cardboard paper around it[10]) will invariably move across the retina because of the saccade. This allows us to ask if one can replicate selective suppression of low spatial frequencies[10] without any saccades at all, solely based on the visual flow experienced during such experiments.

We first replicated the classic phenomenon itself. Our subjects localized briefly flashed vertical Gabor gratings with different spatial frequencies (Methods); the flashes occurred peri-saccadically as in Fig. 1a. Here, however, the screen was a homogeneous gray, like in the classic experiment, with the exception of a surround region showing a stationary texture (the coarse texture used in our earlier experiments, Fig. 7a). We call the large homogeneous central region of the screen (diameter: 20 deg) the “virtual monitor”. The outcome confirmed the classic findings: Fig. 7b (left) shows localization performance for flashed gratings around saccade onset, compared to flashes without saccades (and without any other display transients; Methods), and Fig. 7b (right) plots the ratio of those percepts as a

visualization aid. Perception of low spatial frequency gratings was selectively suppressed (relevant statistics are shown in Fig. 7; full time courses of these effects are shown in Supplementary Figs. 8, 9). These results are consistent with the classic phenomenon[10].

The presence of the textured surround allowed us to next isolate the effects of visual flow during these experiments. In separate trials, we asked subjects to fixate, and we presented saccade-like image motion. For example, the virtual monitor moved together with its textured surround from the top left corner towards display center (Fig. 7c), in order to simulate a real saccade from the lower right corner to display center (Fig. 7a). We then briefly presented the same Gabor gratings as in Fig. 7a, b. Relative to fixation position, this experiment was comparable to the situation with real saccades: there was a uniform background against which a brief Gabor grating was flashed. And, indeed, we observed the same selective suppression of low spatial frequencies despite the absence of saccades (Fig. 7d). Moreover, again consistent with our results from Figs. 1-6, the suppression with simulated saccades lasted longer than with real saccades (robust selective suppression in Fig. 7d occurred even 84 ms after simulated saccades; Supplementary Figs. 8, 9). Similar results were obtained with a uniform black surround around the virtual monitor, as might be the case in typical laboratory settings (Supplementary Fig. 10). Therefore, visual mechanisms account even for the results of Burr et al[10] and similar experiments[15] using uniform backgrounds, without the need to invoke non-visual (motor-related) mechanisms.

Motivated by the differences between coarse and fine textures in Figs. 1-6, we next replaced the coarse texture around the virtual monitor (Fig. 7c) with a fine texture,

and we repeated the experiments with simulated saccades (Fig. 7f). In this case, surprisingly, we observed uniform suppression of gratings of all spatial frequencies (Fig. 7f). This led us to make a strong prediction: if saccadic suppression properties do indeed rely on visual processing, then suppression during real saccades should depend mainly on visual context, and one should be able to easily violate the classic phenomenon (namely, the specific suppression of low spatial frequencies[10]). This is exactly what we found (Fig. 7e): for real saccades across the virtual monitor, and with the surrounding visual context being a fine rather than coarse texture, we observed perceptual suppression for all gratings, abolishing suppression selectivity for low spatial frequencies. In all cases, the effects were not explained by motor variability across surround texture conditions (Supplementary Fig. 3e, f).

In summary, perceptual saccadic suppression occurred in all of our experiments, either with or without real saccades, simply as a function of visual flow (Figs. 1, 2, 5-7). Such suppression quantitatively depended on scene statistics, both for full-field textures (Figs. 1, 2, 5, 6) in a manner predicted by retinal processing (Figs. 3, 4), and for textures limited to the surround (Fig. 7). Even the suppression selectivity of low spatial frequency probes[10] was determined by visual context (Fig. 7).

Discussion

We found that visual image processing accounts for a large component of classic perceptual demonstrations of saccadic suppression, and that such image processing occurs as early as in the very first stage of visual processing, the retina. Key to all of our observations is the single insight that, from the perspective of visual image processing, a saccade is itself a potent stimulus to the visual system. For example, our RGCs often responded vigorously to saccade-like image displacements (Fig. 3b).

Therefore, when probing perceptual sensitivity around the time of saccades using brief flashes, as in classic studies of perceptual saccadic suppression, the visual system is not only responding to the externally provided brief flashes, but it is also responding to the self-induced visual flows caused by eyeball rotations. This means that perceptual saccadic suppression experiments are essentially no different from a wide class of stimulus-stimulus paradigms in which suppression of neural responses to a second stimulus becomes expected, due to so-called neural adaptation after a first stimulus[44]. In fact, neural adaptation to successive stimulus presentations has been observed in many cortical and subcortical areas, as reviewed by Mayo and Sommer[44], where it is often discussed from the perspective of short term plasticity in extra-retinal brain sites. Our results demonstrate that such adaptation, in the context of saccadic eye movement behavior, may underlie the perceptual phenomenon of saccadic suppression, and that it is already jumpstarted as early as in the retina.

Such early neural implementation is interesting because it suggests that the image dependence of perceptual saccadic suppression that we observed (Figs. 1, 2) is derived, at least in part, from visual image processing starting in the retina. In fact, we found remarkable congruence between the image dependence of three phenomena: perceptual suppression with real saccades (Figs. 1, 2), perceptual suppression with simulated saccades (texture displacements, Figs. 5, 6), and neural suppression patterns in RGCs (Figs. 3, 4). In all cases, modifying the background texture statistics resulted in highly predictable changes in suppression profiles. Therefore, perceptual saccadic suppression paradigms are essentially stimulus-stimulus paradigms from the perspective of visual flow, and they can, by consequence, reflect the substantial image processing capabilities already present in the retina.

Additional support for the above sentiment also emerges from the time courses of stimulus-stimulus neural adaptation effects in areas like the frontal eye field and superior colliculus[44]. These time courses are particularly intriguing to us, primarily because they agree with our observations that retinal (Figs. 3, 4) and perceptual (Figs. 5, 6) suppression with simulated saccades had longer suppression time courses than observed with real saccades (Figs. 1, 2). Indeed, the time courses of the neural adaptation effects in the frontal eye field and superior colliculus[44], and related brain areas, are similar to our observed perceptual time courses in the absence of real saccades. Given that both the frontal eye field and superior colliculus have previously been implicated in suppression with real saccades[15,43,45,46], it is thus conceivable that saccadic suppression in these areas is inherited, at least partially, from the retina.

Looking forward, we believe that it is imperative to also investigate the neural mechanisms behind visual masking in much more detail. In our perceptual experiments with simulated saccades (Figs. 5, 6), we saw clear suppression of perceptual localization performance even when the probe flashes appeared before texture displacement. That is, perceptual localization of the probes was masked, backwards in time, by the subsequent texture displacement. In the past, pre-saccadic suppression with real saccades (e.g. Fig. 1) has sometimes been taken as evidence that perceptual saccadic suppression is fundamentally driven by motor-related signals like corollary discharge. However, our results (Fig. 5, 6) show that motor activity is not required. As a perceptual phenomenon, this has been described as backwards visual masking[47], but what are the underlying neural mechanisms? Such backwards masking was not present in our retinal results, certainly not as clearly as

in perception, so it must emerge through visual mechanisms in other brain structures. It would be interesting to explore the neural mechanisms of backwards masking in more detail.

More generally, our results suggest that visual flow does indeed matter a great deal in perceptual saccadic suppression, even in paradigms that have often been taken as indication for motor-related top-down suppression (Fig. 7). It would be interesting in the future to further test the generalizability of this notion. We were indeed greatly surprised when we performed the experiments of Fig. 7, and found that the classic selective suppression of low spatial frequencies in perception around the time of saccades[10] can be violated with a simple change of visual context. More importantly, we found that we could obtain selective suppression of low spatial frequencies even without saccades at all. Thus, with or without saccades, either selective or unselective suppression could occur as a function of visual flow. In hindsight, this might shed light on a somewhat surprising recent finding in superior colliculus neurons[15]. Specifically, using essentially the same paradigms, it was found that only one type of superior colliculus visually-responsive neurons (so-called visual-motor neurons) exhibited selective suppression of low spatial frequency sensitivity as in the classic perceptual phenomenon[15]. The other type of visually-responsive superior colliculus neurons (visual-only neurons) showed mild suppression but, critically, no selectivity for spatial frequency[15]. These two types of neurons occupy different laminae of the superior colliculus and have different patterns of lateral interactions from across the visual field representation of this structure[54]. It is now very conceivable, in light of our current results (Fig. 7), that both patterns of suppression (selective or not) may be embedded simultaneously in these different neuronal populations with specific circuitry and tuning for visual peripheral contexts.

Finally, it should be emphasized that motor-related mechanisms still likely play an important role in perceptual saccadic suppression. In fact, such mechanisms seem to be equally important as the visual mechanisms, since motor-related mechanisms appear to shorten pre- and post-saccadic suppression originating from visual processing (Fig. 5), and might therefore minimize the duration of saccade-induced disruptions. Indeed, there is evidence for post-saccadic enhancement of excitability in a variety of cortical areas[55–57]. It would be interesting to further investigate how such neural enhancement may contribute to the shortened time courses of perceptual saccadic suppression that we observed (e.g. Fig. 5e, f). Furthermore, besides just suppression, saccades are also associated with spatial updating of visual reference frames (due to the image shifts that they cause). Information contained in the motor command itself is likely critical for adjustments of spatial receptive fields across saccades, which have been observed in parietal and frontal cortices[58,59]. Our findings leave open the possibility, however, that trans-saccadic image flow might play a role in this phenomenon as well.

Methods

Ethics approvals

We performed electrophysiological experiments on *ex vivo* mouse and pig retinae as well as non-invasive perceptual experiments on human subjects.

Animal use was in accordance with German and European regulations, and animal experiments were approved by the Regierungspräsidium Tübingen.

Human subjects provided written, informed consent, and they were paid 8 Euros per session of 45-60 minutes each. Depending on the experiment, each subject was measured for 2-10 sessions (detailed trial and session numbers are provided below). Human experiments were approved by ethics committees at the Medical Faculty of Tübingen University, and they were in accordance with the Declaration of Helsinki.

Retina electrophysiology laboratory setup

We used retinae extracted from *PV-Cre x Thy-S-Y* mice (*B6;129P2-Pvalb*$^{tm1(cre)Arbr}$*/J × C57BL/6-tg (ThystopYFPJS))*, which are functionally wild type[60–62]. 23 retinae from 7 male and 15 female mice (3-12 months old) were used. We also replicated experiments on pig retinae obtained from domestic pigs after they had been sacrificed during independent studies at the Department of Experimental Surgery in our Medical Faculty. We used 9 pig retinae.

We housed mice on a 12/12 h light/dark cycle, and we dark adapted them for 4-16 h before experiments. We then sacrificed them under dim red light, removed the eyes, and placed eyecups in Ringer solution (in mM: 110 NaCl, 2.5 KCl, 1 $CaCl_2$, 1.6

$MgCl_2$, 10 D-glucose, and 22 $NaHCO_3$) bubbled with 5% CO_2 and 95% O_2. We removed the retina from the pigment epithelium and sclera while in Ringer solution.

Pigs were anesthetized using atropine, azaperone, benzodiazepine (midazolam), and ketamine, and then sacrificed with embutramide (T61). Before embutramide administration, heparin was injected. The pigs were dark-adapted for 15-20 min before sacrifice. Immediately after sacrifice, the eyes were enucleated under dim red light, and the cornea, lens, and vitreous were removed. Eyecups were kept in CO_2-independent culture medium (Gibco) and protected from light. We transported eyecups to our laboratory and cut pieces from mid-peripheral or peripheral retinae.

We recorded retinal ganglion cell (RGC) activity using either low or high-density multi-electrode arrays (MEAs). The low-density setup consisted of a perforated 60-electrode MEA (60pMEA200/30ir-Ti-gt, Multichannel Systems, Reutlingen, Germany) having a square grid arrangement and 200 μm inter-electrode distance. We mounted an isolated retina on a nitrocellulose filter (Millipore) with a central 2 x 2 mm hole. The mounted retina was placed with the RGC side down into the recording chamber, and good electrode contact was achieved by negative pressure through the MEA perforation. We superfused the tissue with Ringer solution at 30-34 °C during recordings, and we recorded extracellular activity at 25 kHz using a USB-MEA-system (USB-MEA 1060, Multichannel Systems) or a memory-card based system (MEA1060, Multichannel Systems). More details are provided in Reinhard et al[63].

The high-density MEA setup consisted of either a HiDens CMOS MEA[64] (developed by the lab of Andreas Hierlemann, Basel, Switzerland) or a MaxOne system[65]

(Maxwell Biosystems, Basel, Switzerland). The HiDens CMOS MEA featured 11,011 metal electrodes with inter-electrode (center-to-center) spacing of 18 μm placed in a honeycomb pattern over an area of 2 x 1.75 mm. Any combination of 126 electrodes could be selected for simultaneous recording. The MaxOne MEA featured 26,400 metal electrodes with center-to-center spacing of 17.5 μm over an area of 3.85 x 2.1 mm. In this system, up to 1,024 electrodes could be selected for simultaneous recordings. For each experiment, a piece of isolated retina covering almost the entire electrode array was cut and placed RGC-side down in the recording chamber. We achieved good electrode contact by applying pressure on the photoreceptor side of the retina by carefully lowering a transparent permeable membrane (Corning Transwell polyester membrane, 10 μm thick, 0.4 μm pore diameter) with the aid of a micromanipulator. The membrane was drilled with 200 μm holes, with center-center distance of 400 μm, to improve access of the Ringer solution to the retina. We recorded extracellular activity at 20 kHz using FPGA signal processing hardware and custom data acquisition software.

In total, we performed 36 recordings, 24 from mouse and 12 from pig retina. 15 of the 36 recordings were done using low-density MEAs. Once a basic experimental protocol was established, we shifted to HiDens CMOS MEA providing much higher throughput. 12 experiments were done using this setup. We upgraded to the MaxOne MEA for even higher throughput and did our final 9 recordings using this setup.

We presented light stimuli to the retinal piece that was placed on the MEA using a DLP projector running at 60 Hz (Acer K11 for low-density MEA experiments and Lightcrafter 4500 for high-density MEA experiments). 60 Hz is above the flicker

fusion frequency of both mouse and pig retinae; therefore, the framerate of these projectors was adequate for our purposes. The Acer K11 projector had a resolution of 800 x 600 pixels covering 3 x 2.25 mm on the retinal surface. Lightcrafter 4500 had a resolution of 1280 x 800 pixels, extending 3.072 x 1.92 mm on the retinal surface. We focused images onto the photoreceptors using a condenser (low-density MEA recordings, illumination from below) or a 5x objective (high-density MEAs, illumination from above). In each case, the light path contained a shutter and two motorized filter wheels with a set of neutral density (ND) filters (Thorlabs NE10B-A to NE50B-A), having optical densities from 1 (ND1) to 5 (ND5). Light intensity was adjusted to be in the mesopic range.

We measured the spectral intensity profile (in μW cm^{-2} nm^{-1}) of our light stimuli with a calibrated USB2000+ spectrophotometer (Ocean Optics) and converted the physical intensity into a biological equivalent of photoisomerizations per rod photoreceptor per second ($R^{*}rod^{-1}s^{-1}$), as described before[62]. Light intensities of the projector output covered a range of 3 log units (i.e. 1,000-fold difference between black and white pixels, over an 8-bit range). We linearized the projector output, and we used only grayscale images of limited contrast, spanning at most the range from 0 to 120 in the 8-bit range of the projector (see stimulus description below for details). Absolute light intensities were set to the mesopic level, where a stimulus intensity of ‘30’ in our 8-bit DLP projector scale (0-255) corresponded to 225 to 425 $R^{*}rod^{-1}s^{-1}$, depending on the experimental rig used for the experiment (i.e. different DLP projectors and MEAs). We pooled all data from the different rigs because separate individual analyses from the individual setups revealed no effects of recording conditions in the different setups.

Human psychophysics laboratory setup

We used a similar laboratory setup to our recent experiments[40,66,67]. Briefly, subjects sat in a dark room 57 cm in front of a CRT monitor (85 Hz refresh rate; 41 pixels per deg resolution) spanning 34.1 x 25.6 deg (horizontal x vertical). Head fixation was achieved with a custom head, forehead, and chin rest[66], and we tracked eye movements (of the left eye) at 1 kHz using a video-based eye tracker (EyeLink 1000, SR Research Ltd, Canada). Gray and texture backgrounds (e.g. Figs. 1, 5, 7) were always presented at an average luminance of 22.15 cd m^{-2}, and the monitor was linearized (8-bit resolution) such that equal luminance increments and decrements were possible around this average for textures and gratings.

Human Experiment 1 (Fig. 1) was performed by eight subjects (two female) who were 21-25 years old. All subjects were naïve to the purposes of the experiment, except for subject MB (an author). For Human Experiment 2, the “simulated saccade” version of Human Experiment 1 (Fig. 5), six of the same subjects participated. A control experiment for testing visibility of flashes without saccades and without saccade-like texture displacements (Supplementary Fig. 3a, b) was performed by six of the same subjects plus one non-naïve subject, ZH (another author).

In the variants of Human Experiments 1 and 2 in which we collected full psychometric curves and perceptual thresholds (e.g. Figs. 2, 6 and Supplementary Figs. 4, 7), five subjects (24-29 years old; one female) participated. Three of these subjects were the same as those who performed Human Experiments 1 and 2 above, confirming that both variants of the experiments (either with a fixed flash contrast or with full threshold calculations) allowed similar conclusions.

Human Experiment 3 tested suppression selectivity for low spatial frequencies (Fig. 7). Six subjects (three females, 23-25 years old) participated, and only subject MB was non-naïve. Three subjects had also participated in Human Experiments 1 and 2 and their control versions above. A control version of Human Experiment 3 was also performed with black surrounds (Supplementary Fig. 10). This control experiment was performed by the same subjects that participated in Human Experiment 3.

Across all experiments, we ensured that the same subjects performed real and "simulated" saccade versions of a given paradigm so that we could make meaningful comparisons between these two eye movement conditions.

Coarse and fine textures

We created coarse and fine textures (Supplementary Fig. 1a) by convolving a random binary (i.e. white or black) pixel image with a two-dimensional Gaussian blurring filter[68] with the kernel

$$G(x,y) = e^{\frac{-(x^2+y^2)}{2\sigma^2}}$$

The parameter σ of the kernel influenced the amount of blurring. This resulted in textures having effectively low-pass spectral content (Supplementary Fig. 1b) with a cutoff frequency (f_c) depending on σ. As we describe below, we picked cutoff frequencies for coarse and fine textures that resulted in dark and bright image blobs approximating the receptive field sizes of RGCs (for coarse textures) and retinal bipolar cells (for fine textures). In other words, for a given species, coarse textures matched the resolution of RGCs, and fine textures matched the resolution of one processing stage earlier, the retinal bipolar cells.

For the *ex vivo* experiments with mouse and pig retinae, we assumed receptive field diameters for RGCs of at least 150 μm (Supplementary Fig. 1c; the parameter σ of the Gaussian blurring filter would be half that value), and diameters for bipolar cells of 25 μm (see Zhang et al[69]). For human psychophysics experiments, we estimated, from the literature[38], the sizes of human parasol RGC receptive fields at eccentricities >6 deg from the fovea (our flash eccentricities were 7 deg) to be around 200 μm. This translated into a cutoff frequency of ~0.68 cycles per deg (cpd) (Supplementary Fig. 1b). Bipolar cell receptive field sizes at this eccentricity were estimated to be 10 μm (corresponding to a cutoff frequency of ~13.7 cpd), based on sizes of human midget RGC receptive fields in the fovea[38]. When calculating the textures, the actual value of the parameter σ (in pixel-dimensions) always incorporated the specific experimental magnification factor between the stimulation screen and the retinal projection of the image. Calculating power spectra for coarse and fine textures confirmed that cutoff frequencies for a given species were consistent with our aimed designs described above (Supplementary Fig. 1b).

For both retinal and perceptual experiments, we normalized pixel intensities in the textures to have uniform variations in luminance around a given mean. In the retinal experiments, we used pixel intensities (from our 8-bit resolution scale) ranging from 0 to 60 around a mean of 30, or ranging from 30 to 90 around a mean of 60 (see R*etina electrophysiology experimental procedures* below for when each paradigm was used). For the human experiments, textures had a mean luminance of 22.15 cd m^{-2} with undulations in luminance in the texture within the range of 7.5-35.5 cd m^{-2}.

Because each texture, particularly when coarse, could have patterns of dark and bright blobs that human subjects can remember or interpret as potential shapes/objects/figures, we varied the displayed texture images from trial to trial. This was also necessary to avoid afterimages. We generated sets of 20 coarse and 20 fine textures, which we randomly interleaved across trials. Moreover, the textures themselves were designed to be larger than the viewable display area, allowing us to jitter the displayed sub-rectangle of each texture (within the viewable area of the display) from trial to trial (we jittered the displayed sub-rectangle within a range of 0.6 x 0.6 deg in steps of 0.024 deg). This way, even fine patterns at foveal fixation locations could not be memorized by the subjects across trials.

Retina electrophysiology experimental procedures

To simulate saccades in our *ex vivo* retina electrophysiology experiments, we displaced the texture across the retina in 6 display frames (100 ms at 60 Hz refresh rate). For easier readability, we sometimes refer to these saccade-like texture displacements as “saccades”. The textures were displaced in each frame by a constant distance along a linear trajectory. While each “saccade” lasted 100 ms, displacement direction was varied randomly for each “saccade” (uniformly distributed across all possible directions), and “saccade” amplitude could range from 310 μm to 930 μm (corresponding to a velocity range of 3,100-9,300 μm s^{-1} on the retinal surface). In visual degrees, this corresponds to a velocity range of 100-300 deg s^{-1} and displacement range of 10-30 deg in mice, well in the range of observed mouse saccade amplitudes[70]. In fact, similar to primates, mice also have oculomotor behavior, even under cortical control[71]. For example, they make, on average, 7.5 saccade-like rapid eye movements per minute when their head is fixed[70] (humans

make several saccades per second). We used the same retinal displacement range of 310 μm to 930 μm for pig retinae. To the best of our knowledge, pig oculomotor behavior has not been documented in the literature. However, with their larger eyeball sizes, our translations of the retinal image would correspond to slower saccades (e.g. small saccades in humans and monkeys), which are also associated with saccadic suppression. Moreover, we showed (Fig. 4) that retinal "saccadic suppression" is not critically dependent on the details of movement kinematics.

Each "trial" consisted of 39 successive sequences that each combined a "saccade" with a probe flash, as follows: there was first a "pre-saccade" fixation of 2 seconds, then a 100 ms "saccade", followed by "post-saccade" fixation. At a certain time from "saccade" onset (delay d, range: -177 ms to 2,100 ms), we presented a probe flash. In most cases, the probe flash had a duration of 1 frame (~16 ms). We used 2 frames (~33 ms) in a subset of experiments (mouse: 161 of 688 cells analyzed for "saccadic suppression"; pig: 112 of 228 cells). Results were pooled across these paradigms as they were indistinguishable. For sequences containing no probe flash, the next "saccade" happened 4 seconds after the previous one. The probe flash was a full-screen positive ("bright") or negative ("dark") stimulus transient. In different experiments, only a subset of possible delays was used within a given set of trials, depending on total recording time for a given retina (see below).

Bright or dark probe flashes could happen in two different ways across our experiments. The results were indistinguishable between the two ways, so we pooled results across them. Briefly, in one manipulation, the probe flash was a homogeneous bright (pixel intensity of 60 in our 8-bit projectors) or dark (pixel intensity of 0) full-screen rectangle replacing the background texture (in these

experiments, the textures themselves had intensities ranging from 0 to 60 pixel intensity; see *Coarse and fine textures* above). This way, the flash contrast from the underlying background luminance was variable (e.g. a bright flash on a bright portion of a texture had lower contrast from the underlying texture than the same flash over a dark portion of the texture). In the second manipulation, the bright and dark flashes were simply luminance increments or decrements (by pixel values of 30 on our 8-bit projectors) over the existing textures (like in our human perceptual experiments). This way, local contrast relationships in the background textures were maintained. In these experiments, the textures themselves had a range of 30-90 pixel intensities and a mean pixel value of 60 (on our 8-bit projectors). 332 of 688 cells that we analyzed for “saccadic suppression” experienced such probe flashes, whereas the rest (356 cells) experienced the homogenous probe flash. For pig retina recordings, we always used the homogenous framework. However, in the subset of pig experiments where the 2-frame probe flash was employed (112 of 228 RGCs), we used a high-contrast probe flash such that a bright flash would be achieved by first going completely dark in the first frame followed by the bright flash in the next frame and vice versa for a dark flash. Again, all data were pooled across these different paradigms because their outcomes were indistinguishable.

The number of trials required during a physiology experiment depended on the number of conditions that we ran on a specific day. For example, testing 7 different flash delays required 15 trials (7 with bright probe flashes, 7 with dark probe flashes, and 1 without probes). In a given experiment, we always interleaved all conditions; that is, in any one of the 15 necessary trials, each of the 39 “saccades” could be followed by a bright or a dark probe at any of the 7 delays, or no probe at all. Moreover, we repeated the total number of conditions (e.g. the interleaved 15 trials) 4

times per session, and we averaged responses across repetitions. Since one trial typically lasted for 2 minutes, the example of 15 trials repeated 4 times lasted for approximately 2 hours. This was usually combined with additional conditions (e.g. other background textures), such that typical recordings lasted 10-12 hours. If the combination of conditions would have required even longer recordings in a given session, we typically reduced the number of conditions (e.g. we presented flashes at fewer delays).

We sometimes replaced the 100 ms “saccade” with an instantaneous texture jump, to test the sensitivity of retinal “saccadic suppression” (Fig. 3) to the kinematic properties of saccade-like texture displacements (Fig. 4b). Here, the texture simply jumped, in one display frame, from the pre- to the post-displacement position. All other procedures were like described above. 31 RGCs were recorded with this paradigm.

In the control experiments of Fig. 4d, we used no textures at all. The screen was always a homogenous gray field, and the visual event of a "saccade" was replaced by an instantaneous step to a different gray value. The gray backgrounds had intensities between 30 and 90 (on our 8-bit projector). This instantaneous change in intensity caused either a positive contrast step (+0.03 to +0.50 Michelson contrast) or a negative contrast step (-0.03 to -0.50 Michelson contrast). A “trial” consisted of either 57 or 157 successive sequences that each combined a contrast step with a probe flash, as follows: there was first a “pre-step” fixation of 2 seconds (analogous to “pre-saccade” fixation in texture displacements), then an instantaneous switch to “post-step” fixation. At a certain time from the contrast step (delay: 17, 33, 50, 100, 250, 500, 1000 or 2,000 ms), we presented a 2-frame (~33 ms) probe flash. For

sequences containing no probe flash, the next contrast step happened 4 seconds after the previous one. The probe flash was either a uniform negative step of -0.33 Michelson contrast (“dark”) or a uniform positive step of +0.33 Michelson contrast (“bright”).

Finally, we used other stimuli unrelated to the main experiments to help us characterize RGC types and other receptive field properties (e.g. response polarity, latency, transiency, and spatial receptive fields). These stimuli had the same mean intensities and intensity ranges as the textures used in each experiment. Below, we describe these stimuli for the condition in which the texture intensities ranged from 0 to 60 pixel intensity (represented as grayscale RGB values in the units of our 8-bit projects). In experiments in which the textures ranged in intensity from 30 to 90, all intensities reported below were shifted upward by 30. (1) Full-field contrast steps. ON steps: stepping from 0 to 30 (+1 Michelson contrast) and from 30 to 60 (+0.33) for 2 s. OFF steps: stepping from 60 to 30 (-0.33) and from 30 to 0 (-1) for 2 s. (2) Full-field Gaussian flicker, 1 minute. Screen brightness was updated every frame and was drawn from a Gaussian distribution with mean 30 and standard deviation 9. This stimulus was used to calculate the linear receptive field filters of ganglion cells through reverse correlation (spike-triggered averaging of the stimulus history). (3) Binary checkerboard flicker, 10-15 minutes. The screen was divided into a checkerboard pattern; each checker either covered an area of 55 x 55 μm, 60 x 60 μm, or 65 x 65 μm depending on the recording rig. The intensity of each checker was updated independently from the other checkers and randomly switched between 10 and 50 or 0 and 120. This stimulus also allowed us to calculate the linear filters of cells’ receptive fields.

Human psychophysics experimental procedures

In Human Experiment 1, we presented a coarse or fine background texture (Fig. 1) for 800-1,700 ms in every trial. Over the texture, a white fixation marker (square of 7.3 x 7.3 arcmin) surrounded by a uniform gray circle of 30 min arc radius was presented at one screen location in order to guide gaze fixation onto the marker. The fixation marker was always at 4.8 deg eccentricity from display center, but its specific location was varied from trial to trial (up-right, up-left, down-right, or down-left relative to display center; 45 deg direction from horizontal). After the end of the initial interval, the fixation marker jumped to display center, instructing subjects to generate a saccade.

At a random time from the saccade instruction (47, 94, 153, 200, 247, or 507 ms), a luminance pedestal (probe flash) was applied for one display frame (~12 ms) at one of four locations relative to display center (7 deg above, below, to the right of, or to the left of center). Note that because the display was rasterized (that is, drawn by the computer graphics board from the top left corner in rows of pixels), the actual exact flash time and duration depended on the location of the flash on the display (but in a manner like other psychophysical experiments studying the same phenomenon, and also in a manner that is unlikely to affect our results). The luminance pedestal consisted of a square of 147.8 x 147.8 min arc in which we added or subtracted a value of 4.8 cd m^{-2} to the texture pattern. Therefore, local contrast within the luminance pedestal was the same as that without the pedestal. Since all of our analyses revealed identical results whether the pedestal was a luminance increment or decrement, we combined these conditions in all analyses. At the end of the trial,

subjects had to report their perceived flash location by pressing one of four buttons, corresponding to the four possible flash locations, on a hand-held response box.

Because saccadic reaction times were 156.9 +/- 3.3 ms s.e.m. across subjects, our choice of flash times above meant that we could analyze trials in which flashes appeared before or after saccade onset, allowing us to obtain full time courses (e.g. Fig. 1). Also, because of the display geometry, the retinal region that experienced a flash before, during, or after a saccade was always a region that was visually-stimulated by the texture before flash onset (rather than by the monitor edge or the black surround of the laboratory). Therefore, we maintained pre- and post-flash visual stimulation by texture background, as in the retinal experiments. We also ensured that flash locations were not coincident with saccade goal locations both retinotopically and also in display coordinates. We confirmed in separate analyses that similar effects of suppression (e.g. Fig. 1) occurred for each flash location separately.

We collected 576 trials per session in this experiment. Six subjects participated in 6 sessions each, and the remaining two participated in 3 or 4 sessions.

Human Experiment 2 (Fig. 5) was identical, except that the initial fixation marker was presented at display center and remained there for the entire duration of a trial. Instead of instructing a saccade 800-1,700 ms after fixation marker onset, we translated the entire background texture rapidly to simulate a saccade-like image displacement. Texture displacement consisted of a 6-frame translation at a speed of 176 deg s^{-1}. Note that, because of our display refresh rate and geometry, this meant a slightly larger displacement (of 12.4 deg) when compared to the saccade sizes in

Human Experiment 1. However, we chose this translation because it resulted in a sufficiently fast average speed of the displacement (average speed in the real saccades of Human Experiment 1 was 160 deg s^{-1}). This choice is not problematic because our retinal experiments revealed that visual mechanisms related to saccadic suppression were not sensitive to parameters of individual motion patterns (Fig. 4b).

In this experiment, the texture displacement happened in a diagonal direction to simulate the directions of saccadic displacements of Human Experiment 1 (and also to dissociate the direction of motion flow from the locations of the flashes, again as in Human Experiment 1). For example, the texture could move globally down-right, as might be expected (in terms of image motion) if subjects made upward-leftward saccades in Human Experiment 1. Also, flash times were chosen relative to the onset of texture displacement from among the following values: -35, -24, 24, 47, 84, 108, 141, 200, 259, 494 ms.

All subjects participated in 10 sessions each in this experiment.

We also performed a control experiment, in which there was neither a real saccade (Human Experiment 1) nor a texture displacement (Human Experiment 2), but otherwise identical to these 2 experiments. Subjects simply fixated display center, and we presented (after 1,200 to 2,400 ms from trial onset) a luminance pedestal exactly as in Human Experiments 1 and 2. To obtain full psychometric curves, we varied the luminance increment from among 6 values (Supplementary Fig. 3a, b). Subjects performed two sessions each of this experiment (600 trials per session).

To explore perceptual thresholds in a more quantitative manner for Human Experiments 1 and 2, we also performed additional real or simulated saccade experiments collecting full psychometric curves (Figs. 2, 6 and Supplementary Figs. 4, 7). The logic of both additional experiments (real or simulated) was the same as that of Human Experiments 1 and 2, except that we varied the luminance of the probe flash from trial to trial (like in the above control experiment of flash visibility; Supplementary Fig. 3a, b). Because this endeavor (allowing us to measure full psychometric curves) was very data intensive, we reduced the time samples relative to saccade onset or texture displacement onset at which we probed perceptual performance. For the experiment with real saccades, we used an automatic procedure to detect saccade onset in real time based on eye velocity, as described by Chen and Hafed[41]. We then presented the probe flash at 42, 65, 88, or 148 ms after saccade detection. These times were chosen because they covered intervals of maximum perceptual saccadic suppression as well as recovery, allowing us to get a time course of perceptual threshold elevation associated with saccadic suppression. In subsequent data analyses, we confirmed that these flash times were as planned (within the expected variability due to the asynchronous nature of saccade times relative to display update times; Fig. 2). For the experiment with simulated saccades, we presented the probe flash at -24, -12, 48, or 96 ms relative to the onset time of the texture displacement. In this case, we introduced a new negative time sample to the set (-12 ms) because the original Human Experiment 2 did not probe this particular time (e.g. Fig. 5). It was therefore important to clarify that the time course of perceptual suppression for simulated saccades was continuous and well-behaved, exactly like that for real saccades.

In order to also estimate perceptual thresholds online in these additional experiments, and therefore optimize the numbers of trials needed, we applied an adaptive QUEST procedure[42] on each randomly interleaved condition. Specifically, the first 40 trials of each randomly interleaved condition (e.g. flash time -24 ms and coarse texture, or flash -12 ms time and fine texture, and so on) were part of the QUEST procedure. The remaining trials in the session interleaved 4 additional flash luminances per condition, which were chosen to lie around the threshold luminance of each condition as detected by the QUEST procedure. These additional flashes had luminances that were +/- 1 or +/- 2 times a pre-defined luminance increment for a given condition, depending on the detected threshold and earlier pilot data. Specifically, if the detected threshold (according to QUEST) was very low (e.g. no suppression effect), the pre-defined luminance increment was 1 step of luminance (dictated by the luminance resolution of our display; Supplementary Fig. 3a). That is, the 4 additional flashes were at +/-1 and +/-2 display-determined luminance steps from the detected threshold. If the detected threshold (according to QUEST) was high (e.g. strong suppression), we made the pre-defined luminance increment 2 or 5 display-determined luminance steps (that is, +/- 2 and +/-4 display-determined luminance steps or +/- 5 and +/-10 display-determined luminance steps, respectively).This allowed filling the psychometric curves during subsequent data analyses, including measurements from the full dynamic range of perceptual performance. The reasoning behind this approach is as follows: depending on the amount of perceptual saccadic suppression to be expected per condition (e.g. peak suppression during saccades or texture displacements, or very weak suppression during recovery), it is expected that the psychometric curves would be shifted by different amounts from baseline depending on the particular condition (e.g. flash time or coarse versus fine texture). Finally, also note that we only used bright flashes in

these particular experiments instead of both bright and dark flashes. In total, we collected 240 trials per condition per subject.

In Human Experiment 3 (Fig. 7), the flashes of Human Experiments 1 and 2 were replaced by vertical Gabor gratings having one of five different spatial frequencies (0.41, 0.85, 1.71, 3.42, 4.56, or 6.8 cpd). Spatial phase was randomized from trial to trial, and the σ parameter of the Gaussian envelope was 0.49 deg. Also, a virtual monitor of 20 deg diameter was present at display center at the time of Gabor grating flashes. The virtual monitor had a uniform gray luminance equal to the average of the textures used in Human Experiments 1 and 2. Surrounding the virtual monitor, a coarse or fine texture could be visible.

In one block of trials, subjects generated saccades towards display center using the same procedures as in Human Experiment 1. Grating flash times were similar to Human Experiment 1, and the subjects performed 6 sessions each (576 trials per session).

In another block of trials, subjects maintained fixation at display center. In one third of the trials, the virtual monitor and surrounding texture did not move. These trials provided us with “baseline” visual performance (i.e. without saccades or virtual monitor displacements). It was necessary to have these trials because perceptual visibility of different spatial frequencies is not equal due to the well-known human contrast sensitivity function[72]. Therefore, we needed to establish “baseline” grating visibility first and then compare the effects of saccades or saccade-like virtual monitor displacements on such visibility. In the remaining two thirds of the trials, the virtual monitor and surrounding texture initially appeared displaced from display center at a

location near one corner of the display and along one of the diagonal directions. After 800-1,700 ms, the virtual monitor and surrounding texture were translated rapidly towards display center to simulate visual flow associated with the diagonal saccades of the real-saccade version of the paradigm (the translation parameters were similar to Human Experiment 2). Grating flashes happened 84 ms or 108 ms after virtual monitor and texture displacement. Note that we reduced the number of flash times here because of the larger number of conditions (5 different spatial frequencies of the Gabor gratings) that needed to be collected. However, our data were consistent with all other experiments in terms of recovery time courses of suppression (e.g. Figs. 1, 5, 7; Supplementary Figs. 8-10).

Because the initial displaced position of the virtual monitor (and texture) provided a cue to subjects that grating onset was expected soon, and because such a cue was not present in the one third of trials without image motion, we equalized subject expectations across these conditions by dimming the fixation point to black from the time of image motion onset until 200 ms after flash onset (equal timing was ensured in the one third of trials without image motions, such that the same expectation of grating onset was established by fixation marker dimming). The fixation marker then disappeared, and subjects had to report flash location.

Subjects performed 6 sessions each of this condition, with 576 trials per session (2 subjects performed 7 and 5 sessions each instead of 6).

Finally, we repeated the same experiment but with a black surround around the virtual monitor instead of a coarse or fine texture. Note that a black surround is theoretically equivalent to an infinitely coarse surround. We therefore expected

results conceptually similar to those with a coarse surround. Also, in this control experiment, we randomly interleaved all trial types together in the same session (fixation with virtual monitor displacement, real saccade, and fixation with neither virtual monitor displacement nor saccade). This allowed us to further confirm that our results from Human Experiment 3 were not influenced by the separate blocking of real saccade trials and virtual monitor displacement trials.

Retina electrophysiology data analysis and statistics

Low-density MEA recordings were high-pass filtered at a 500 Hz cutoff frequency using a tenth-order Butterworth filter. We extracted spike waveforms and times using thresholding, and we semi-manually sorted spikes using custom software. For high-density MEA recordings, we performed spike sorting by an offline automatic algorithm[73] and assessed the sorted units using UnitBrowser[74]. We judged the quality of all units using inter-spike intervals and spike shape variation. Low quality units, such as ones with high inter-spike intervals, missing spikes, or contamination, were discarded. All firing rate analyses were based on spike times of individual units.

We first characterized the properties of RGCs. We calculated linear filters in response to full-field Gaussian flicker and binary checkerboard flicker by summing the 500-ms stimulus history before each spike. The linear filters allowed determining cell polarity. Specifically, the amplitude of the first peak of the filter was determined. If the peak was positively deflected, the cell was categorized as an ON cell; if negatively deflected, the cell was an OFF cell. ON cells were later always analyzed with respect to their responses to bright probe flashes in the main experiment, and OFF cells were analyzed with dark probe flashes. We determined the spatial receptive fields of RGCs by calculating the linear filters for each region (checker)

defined by the binary checkerboard flickering stimulus. The modulation strength of each linear filter, measured as the s.d. along the 500 ms temporal kernel, is an estimate for how strongly that region drives ganglion cell responses. We fitted the resulting 2D-map of s.d. values with a two dimensional Gaussian and took the 2-σ ellipse (long axis) as the receptive field diameter. For all other figures and analyses, we converted spike times to estimates of firing rate by convolving these times with a Gaussian of σ = 10 ms standard deviation and amplitude 0.25 $\sigma^{-1}e^{1/2}$.

For each RGC, we used responses to full-field contrast steps to calculate an ON-OFF index, a transiency index, and a response latency index. These indices were used to characterize the properties of RGCs (Supplementary Fig. 6) that we included in our analyses. The ON-OFF index was calculated by dividing the difference between ON and OFF step peak response by their sum. The resulting index values ranged between -1 (OFF) and +1 (ON) and were then scaled to span between 0 (OFF) and +1 (ON). The transiency index was defined as the ratio of the response area within the first 400 ms and the total response area spanning 2,000 ms. The resulting index had a value of 1 for pure transient cells. Response latency was calculated as the time from stimulus onset to 90% of peak response. This value was normalized to the maximum response latency in our dataset to create the response latency index.

To quantify retinal “saccadic suppression”, we first determined a “baseline response”, defined as the response to a probe flash approximately 2 s after texture displacement onset (delay between 1,967 to 2,100 ms, depending on the specific flash times used in a specific experiment). This baseline response was compared to responses of the

same cell to the same flash when it occurred at an earlier time (i.e. closer in time to the “saccade”). Usually, the saccade-like texture displacements themselves caused significant neural responses even without flashes (“saccade-response”, e.g. Fig. 3b), and the responses to the flashes were superimposed on these “saccade-responses” (Fig. 3c). We therefore first isolated the component of the responses caused by the flashes by subtracting the “saccade-responses” from the composite responses.

To get a robust estimate of the response to “saccades” alone (i.e. without any flashes), we averaged spike rate from before “saccade” onset up until the next “saccade” onset for conditions in which no flash was presented, or until just before the flash onset for conditions in which a “post-saccade” flash was presented. This was done for each of the 39 successive “saccades” in a given trial.

We then computed a neural modulation index, ranging from -1 to +1. A value of -1 represents complete suppression of flash-induced responses, whereas +1 indicates “complete enhancement” of flash-induced responses (that is, there was only a response to a flash after saccades, but not to a flash in isolation). A modulation index of 0 meant no change in flash-induced response relative to the “baseline” response. The modulation index of an RGC for a given flash delay d after “saccade” onset was calculated as $(r_d – r_b)/(r_d + r_b)$ where r_d is the peak firing rate for the flash-component of the response (see above for how we isolated this from the composite “saccade”+flash response) and r_b is the peak firing rate for the baseline flash response (i.e. the same flash but occurring ~2 s away from any “saccade”; see above). In all cases, peak firing rate was estimated after averaging responses from all repetitions of a given condition (delay d or baseline) for a given RGC. For ON cells, the modulation index was based only on responses to bright flashes, and for

OFF cells, it was based on responses to dark flashes. For some analyses, we also calculated modulation indices of RGCs for each of the 39 individual “saccades” using the same procedure.

In some cells and trials, individual “saccades” from the sequence of 39 were discarded. This happened when the baseline response peak was less than 60% of the median baseline response peak across the 39 “saccades” of a given trial. We did this to ensure that our modulation indices were not marred by a numerator and denominator approaching zero (e.g. if both flash and baseline responses were weak). We did, however, re-include sequences in which the peak response to the flash after the “saccade” was above the median baseline response peak (across the 39 “saccades”). This was done in order to re-include sequences (if discarded by the first step) for which the baseline flash response was weak but a flash after “saccades” nonetheless gave a robust response. For example, this could happen if a cell did not respond to a flash in isolation but the “saccade” enhanced the response to a flash following it. Our main results (e.g. Fig. 3) were highly robust to such scenarios.

Finally, to perform statistics, we applied tests at either the individual cell level or at the level of the population. At the individual cell level, we determined whether a given RGC’s modulation index for a probe flash presented at a given delay was significantly different from 0 (i.e. “Is the response of this cell modulated by the ‘saccade’?”). For this, we performed a one-tailed sign test of the null hypothesis that the 39 individual modulation indices came from a distribution with zero median against the alternative hypothesis that the median was below (for negative modulation index) or above (for positive modulation index) zero. The modulation index was considered significant (i.e. the flash response was modulated by the

“saccade”) at $p<0.05$ if the test had a power ($1-\beta$) of at least 0.8. At the population level, we determined whether the retinal output as a whole was modulated by “saccades”. For this, we performed a two-tailed Wilcoxon signed rank test of the null hypothesis that the median of the distribution of modulation indices did not differ from 0. Lastly, we tested whether the modulation index of the population was significantly different across textures. For this, we performed a two-tailed Wilcoxon signed rank test of the null hypothesis that the median of the distribution of modulation indices did not differ across textures. Since our modulation index was based on responses to the brief probe flashes, it could only be computed for cells that did respond to these flash stimuli (mouse: N = 688 of 1,423 recorded cells; pig: N = 228 of 394). Only these cells, showing a measurable baseline flash response, were included in our analyses for retinal “saccadic suppression” (Fig. 3e, Supplementary Fig. 5).

To quantify retinal “saccadic suppression” in our control experiments with structure-free uniform backgrounds and contrast steps in place of textures and texture displacements (Fig. 4d), we used the same analyses and statistical procedures to those described above for the texture displacement paradigm. The only difference was that instead of 39 successive “saccades” in a trial, we now had either 57 or 157 successive contrast steps (depending on experiment setting). 22 of 57 or 66 of 157 steps had a Michelson contrast in the range of +/- 0.03 to 0.15 and these steps were used to quantify suppression for low contrast steps. 24 of 57 or 58 of 157 steps had a Michelson contrast in the range of +/- 0.20 to 0.40 and were used to quantify suppression for high contrast steps. As shown in Fig. 4d, low and high contrast steps resulted in the modulation of ganglion cell responses to the probe flashes that was reminiscent of the modulation observed after displacement of fine and coarse textures, respectively. Similar to the texture displacement paradigm, the modulation

index was based on responses to brief probe flashes, and it could therefore only be computed for cells that did respond to these flash stimuli (N = 376 of 650 recorded RGCs in mouse). The modulation index for ON RGCs was calculated from responses to bright probe flashes, and that for OFF RGCs was calculated from responses to dark flashes.

Human psychophysics data analysis and statistics

We analyzed eye movements in all trials. We detected saccades using established methods[41,75], and we manually inspected all trials to correct for mis-detections. In experiments requiring a saccade (e.g. Fig. 1), we excluded from analysis any trials with premature (before saccade instruction) or late (>500 ms reaction time) saccades. We also rejected all trials in which saccades landed >0.5 deg from the saccade target. In experiments requiring fixation, we excluded from analysis any trials in which a saccade or microsaccade happened anywhere in the interval from 200 ms before to 50 ms after any flash or grating onset.

For experiments with saccades (e.g. Fig. 1), we obtained time courses of perception by calculating, for each trial, the time of flash or grating onset from saccade onset. We then binned these times into 50 ms bins that were moved in 5 ms bin-steps relative to saccade onset. Within each bin, we calculated the proportion of correct trials, and we obtained full time courses of this perceptual measure. We obtained time course curves for each subject individually, and we then averaged the curves for the individual subjects in summary figures. All of our analyses were robust at the individual subject level as well (e.g. Supplementary Fig. 2).

For experiments with simulated saccades (i.e. saccade-like texture displacements), there were discrete flash or grating times relative to “simulated saccade” onset, so no temporal binning was needed. At each flash or grating time, we simply calculated the proportion of correct trials.

When we fitted performance to psychometric curves (e.g. Supplementary Fig. 3a, b), we used *the psignifit 4 toolbox*[76], and we used an underlying beta-binomial model.

We also used the same toolbox to analyze the variants of Human Experiments 1 and 2 in which we collected full psychometric curves (Figs. 2, 6). For these experiments, we defined the threshold as the flash luminance level that resulted in 62.5% correct perceptual performance. We then plotted the value of such threshold as a function of flash time relative to real or simulated saccade time.

For some analyses of Human Experiment 3 and its control version, we calculated a “suppression ratio” as a visualization aid (e.g. Fig. 7). This was obtained as follows. For a given spatial frequency grating, we calculated the fraction of correct trials within a given time window (from either simulated or real saccade onset) divided by the fraction of correct trials for the same spatial frequency when there was neither a saccade nor a virtual monitor and texture displacement (i.e. baseline perception of a given spatial frequency). This ratio therefore revealed the effect of suppression independently from the underlying visibility of any given spatial frequency[15]. However, note that we also report raw proportions of correct trials in all conditions.

All error bars that we show denote s.e.m. across individual subjects, except where we report individual subject analyses and control analyses. For individual subject

performance, error bars denote s.e.m. across trials; for control analyses, error bars denote 95% confidence intervals (e.g. Supplementary Fig. 3a, b) or s.d. (e.g. Supplementary Fig. 3d, f). All error bar definitions are specified in the corresponding figures and/or legends.

To statistically validate if the time courses for perceptual localization performance for saccades across the different background textures (coarse versus fine) differed significantly from each other (e.g. Fig. 1), we used a random permutation test with correction for time clusters of adjoining significant p-values[39,40]. First, for each time bin, we calculated a test statistic comparing performance for coarse versus fine background textures. This test statistic was the difference between the proportion of correct responses for the different textures. Then, we performed a random permutation with 1,000 repetitions for each time bin; that is, we collected all trials of both conditions, within a given time bin, into a single large set, and we randomly assigned measurements as coming from either coarse or fine textures, while at the same time maintaining the relative numbers of observations per time bin for each texture condition. From this resampled data, we calculated the test statistic again, and we repeated this procedure 1,000 times. Second, we checked, for each time bin, whether our original test statistic was bigger than 95% of the resampled test statistics (i.e. significant), and we counted the number of adjoining time bins that were significant at this level (i.e. clusters of time bins in which there was a difference between coarse and fine textures). We then repeated this for all 1,000 resampled test statistics. The p-value for our original clusters was then calculated as the number of resampled clusters that were bigger or the same size as the original clusters, divided by the total number of repetitions (1,000). This procedure was described in detail elsewhere[40]. We followed a conservative approach, paying no attention to which bins

in the resampled data formed a cluster of time bins. As discussed elsewhere[40], our statistical analysis constituted a highly conservative approach to establishing significance of differences between time courses for coarse and fine textures. In Human Experiment 3, we used the same approach to compare time courses of suppression ratio for coarse and fine surround contexts with real saccades.

For Human Experiment 2, we had discrete flash times relative to texture displacement onset. Here, the comparison between coarse and fine textures was tested with a Bonferroni-corrected χ^2 test at corresponding flash times. To compare between real and simulated saccades in Human Experiments 1 and 2, we also ran a Bonferroni-corrected χ^2 test. We only considered time bins in the real saccade data that corresponded to the discrete flash times in the simulated saccade data. A Bonferroni correction was necessary because we tested the same data sets on multiple time bins with the same hypothesis (that there is a difference in time courses).

In Human Experiment 3, we also compared suppression ratios for real and simulated saccades for a given texture surround. We again used a Bonferroni-corrected χ^2 test. This was justified because within a given surround, baseline data were the same for real and simulated saccades. Therefore, the relationship between the proportion of correct localizations and suppression ratio was identical. In contrast, testing suppression ratios between fine and coarse surrounds in the same experiment with a χ^2 test was not applicable because baseline values differed. Therefore, we used instead a random permutation test with 5,000 repetitions. To compare the different

spatial frequency Gabor gratings in one bin or time stamp, we used the Kruskal-Wallis test.

Data availability

All data presented in this paper are stored and archived on secure institute computers and are available upon reasonable request.

Acknowledgements

Andreas Hierlemann provided the HiDens CMOS MEA system and helped establish our high-density MEA recordings. Roland Diggelmann helped in setting up the pipeline (including providing code) for automatic spike sorting of high-density MEA recordings. This work was supported by funds of the Deutsche Forschungsgemeinschaft (DFG) to the Werner Reichardt Center for Integrative Neuroscience (EXC 307) and to T.A.M. (MU3792/3-1). T.A.M. received support from the Tistou and Charlotte Kerstan Foundation. T.A.M. and Z.M.H. were also supported by an intra-mural funding program (Projekt 2013-05) of the Werner Reichardt Centre for Integrative Neuroscience. F.F. was supported by a Swiss National Science Foundation Ambizione grant (PZ00P3_167989).

Author contributions

S.I., M.B., T.A.M., Z.M.H. designed the overall study; S.I., M.B., T.A.M., Z.M.H. designed experiments; S.I. performed *ex vivo* retina experiments; M.B., Z.M.H. performed human psychophysics experiments; S.I., M.B., F.F., T.A.M., Z.M.H. analyzed data; S.I., M.B., F.F., T.A.M., Z.M.H. wrote manuscript.

Competing interests

The authors declare no competing interests.

References

1. O'Regan, J. K. & Noë, A. A sensorimotor account of vision and visual consciousness. *Behav. Brain Sci.* **24**, 939–973 (2001).

2. Wurtz, R. H. Neuronal mechanisms of visual stability. *Vision Res.* **48**, 2070–89 (2008).

3. Wurtz, R. H., Joiner, W. M. & Berman, R. A. Neuronal mechanisms for visual stability: Progress and problems. *Philos. Trans. R. Soc. B Biol. Sci.* **366**, 492–503 (2011).

4. Thiele, A., Henning, P., Kubischik, M. & Hoffmann, K. P. Neural mechanisms of saccadic suppression. *Science (80-.).* **295**, 2460–2462 (2002).

5. Zuber, B. L. & Stark, L. Saccadic suppression: Elevation of visual threshold associated with saccadic eye movements. *Exp. Neurol.* **16**, 65–79 (1966).

6. Beeler, G. W. Visual threshold changes resulting from spontaneous saccadic eye movements. *Vision Res.* **7**, 769–775 (1967).

7. Matin, E. Saccadic suppression: a review and an analysis. *Psychol. Bull.* **81**, 899–917 (1974).

8. Riggs, L. A. & Manning, K. A. Saccadic suppression under conditions of whiteout. *Invest. Ophthalmol. Vis. Sci.* **23**, 138–143 (1982).

9. Volkmann, F. C. Human visual suppression. *Vision Res.* **26**, 1401–1416 (1986).

10. Burr, D. C., Morrone, M. C. & Ross, J. Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature* **371**, 511–513 (1994).

11. Ross, J., Burr, D. C. & Morrone, M. C. Suppression of the magnocellular pathway during saccades. *Behav. Brain Res.* **80**, 1–8 (1996).

12. Bremmer, F., Kubischik, M., Hoffmann, K.-P. & Krekelberg, B. Neural dynamics of saccadic suppression. *J. Neurosci.* **29**, 12374–12383 (2009).

13. Hafed, Z. M. & Krauzlis, R. J. Microsaccadic suppression of visual bursts in the primate superior colliculus. *J. Neurosci.* **30**, 9542–9547 (2010).

14. Krekelberg, B. Saccadic suppression. *Curr. Biol.* **20**, R228–R229 (2010).

15. Chen, C.-Y. & Hafed, Z. M. A neural locus for spatial-frequency specific saccadic suppression in visual-motor neurons of the primate superior colliculus. *J. Neurophysiol.* **117**, 1657–1673 (2017).

16. Duffy, F. H. & Lombroso, C. T. Electrophysiological evidence for visual suppression prior to the onset of a voluntary saccadic eye movement. *Nature* **218**, 1074–1075 (1968).

17. Diamond, M. R., Ross, J. & Morrone, M. C. Extraretinal control of saccadic suppression. *J. Neurosci.* **20**, 3449–3455 (2000).

18. Ross, J., Morrone, M. C., Goldberg, M. E. & Burr, D. C. Changes in visual perception at the time of saccades. *Trends Neurosci.* **24**, 113–121 (2001).

19. Gremmler, S. & Lappe, M. Saccadic suppression during voluntary versus reactive saccades. *J. Vis.* **17**, 1–10 (2017).

20. Mackay, D. M. Elevation of visual threshold by displacement of retinal image. *Nature* **225**, 90–92 (1970).

21. Mitrani, L., Mateeff, S. & Yakimoff, N. Is saccadic suppression really saccadic? *Vision Res.* **11**, 1157–1161 (1971).

22. Matin, E., Clymer, A. B. & Matin, L. Metacontrast and saccadic suppression. *Science* **178**, 179–182 (1972).

23. Mitrani, L., Yakimoff, N. & Mateeff, S. Saccadic suppression in the presence of structured background. *Vision Res.* **13**, 517–521 (1973).

24. Mateeff, S., Yakimoff, N. & Mitrani, L. Some characteristics of the visual masking by moving contours. *Vision Res.* **16**, 489–492 (1976).

25. Brooks, B. A., Impelman, D. M. K. & Lum, J. T. Backward and forward masking associated with saccadic eye movement. *Percept. Psychophys.* **30**, 62–70 (1981).

26. Macknik, S. L. & Livingstone, M. S. Neuronal correlates of visibility and invisibility in the primate visual system. *Nat. Neurosci.* **1**, 144–149 (1998).

27. Castet, E., Jeanjean, S. & Masson, G. S. ‘Saccadic suppression’ – no need for an active extra-retinal mechanism. *Trends Neurosci.* **24**, 316–317 (2001).

28. Castet, E. Perception of intra-saccadic motion. in *Dynamics of visual motion processing* 141–160 (Springer US, 2010). doi:10.1007/978-1-4419-0781-3

29. García-Pérez, M. A. & Peli, E. Visual contrast processing is largely unaltered during saccades. *Front. Psychol.* **2**, 1–15 (2011).

30. Ilg, U. J. & Hoffmann, K. P. Motion perception during saccades. *Vision Res.* **33**, 211–220 (1993).

31. Campbell, F. W. & Wurtz, R. H. Saccadic omission: Why we do not see a grey-out during a saccadic eye movement. *Vision Res.* **18**, 1297–1303 (1978).

32. Krueger, J. & Fischer, B. Strong periphery effect in cat retinal ganglion cells. Excitatory responses in ON- and OFF-center neurones to single grid displacements. *Exp. Brain Res.* **18**, 316–318 (1973).

33. Noda, H. & Adey, W. R. Retinal ganglion cells of the cat transfer information on saccadic eye movement and quick target motion. *Brain Res.* **70**, 340–345 (1974).

34. Barlow, H. B., Derrington, A. M., Harris, L. R. & Lennie, P. The effects of remote retinal stimulation on the responses of cat retinal ganglion cells. *J. Physiol.* **269**, 177–194 (1977).

35. Enroth-Cugell, C. & Jakiela, H. G. Suppression of cat retinal ganglion cell responses by moving patterns. *J. Physiol.* **302**, 49–72 (1980).

36. Roska, B. & Werblin, F. Rapid global shifts in natural scenes block spiking in specific ganglion cell types. *Nat. Neurosci.* **6**, 600–608 (2003).

37. Passaglia, C. L., Freeman, D. K. & Troy, J. B. Effects of remote stimulation on the modulated activity of cat retinal ganglion cells. *J. Neurosci.* **29**, 2467–2476 (2009).

38. Dacey, D. M. & Petersen, M. R. Dendritic field size and morphology of midget and parasol ganglion cells of the human retina. *Proc. Natl. Acad. Sci.* **89**, 9666–9670 (1992).

39. Maris, E. & Oostenveld, R. Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* **164**, 177–190 (2007).

40. Bellet, J., Chen, C.-Y. & Hafed, Z. M. Sequential hemifield gating of α- and β-behavioral performance oscillations after microsaccades. *J. Neurophysiol.* **118**, 2789–2805 (2017).

41. Chen, C.-Y. & Hafed, Z. M. Postmicrosaccadic Enhancement of Slow Eye Movements. *J. Neurosci.* **33**, 5375–5386 (2013).

42. Watson, A. B. & Pelli, D. G. Quest: A Bayesian adaptive psychometric method. *Percept. Psychophys.* **33**, 113–120 (1983).

43. Robinson, D. L. & Wurtz, R. H. Use of an extraretinal signal by monkey superior colliculus neurons to distinguish real from self-induced stimulus movement. *J. Neurophysiol.* **39**, 852–870 (1976).

44. Mayo, J. P. & Sommer, M. A. Neuronal Adaptation Caused by Sequential Visual Stimulation in the Frontal Eye Field. *J. Neurophysiol.* **100**, 1923–1935 (2008).

45. Krock, R. M. & Moore, T. Visual sensitivity of frontal eye field neurons during the preparation of saccadic eye movements. *J. Neurophysiol.* **116**, 2882–2891 (2016).

46. Chen, C.-Y., Ignashchenkova, A., Thier, P. & Hafed, Z. M. Neuronal response gain enhancement prior to microsaccades. *Curr. Biol.* **25**, 2065–2074 (2015).

47. Breitmeyer, B. G. Visual masking: past accomplishments, present status, future developments. *Adv. Cogn. Psychol.* **3**, 9–20 (2007).

48. Castet, E. & Masson, G. S. Motion perception during saccadic eye movements. *Nat. Neurosci.* **3**, 177–83 (2000).

49. Ramcharan, E. J., Gnadt, J. W. & Sherman, S. M. The effects of saccadic eye

movements on the activity of geniculate relay neurons in the monkey. *Vis. Neurosci.* **18**, 253–258 (2001).

50. Reppas, J. B., Usrey, W. M. & Reid, R. C. Saccadic eye movements modulate visual responses in the lateral geniculate nucleus. *Neuron* **35**, 961–974 (2002).

51. Kleiser, R., Seitz, R. J. & Krekelberg, B. Neural correlates of saccadic suppression in humans. *Curr. Biol.* **14**, 386–390 (2004).

52. Royal, D. W., Sáry, G., Schall, J. D. & Casagrande, V. A. Correlates of motor planning and postsaccadic fixation in the macaque monkey lateral geniculate nucleus. *Exp. Brain Res.* **168**, 62–75 (2006).

53. Hass, C. A. & Horwitz, G. D. Effects of microsaccades on contrast detection and V1 responses in macaques. *J. Vis.* **11**, 3–3 (2011).

54. Phongphanphanee, P. *et al.* Distinct local circuit properties of the superficial and intermediate layers of the rodent superior colliculus. *Eur. J. Neurosci.* **40**, 2329–2343 (2014).

55. Rajkai, C. *et al.* Transient cortical excitation at the onset of visual fixation. *Cereb. Cortex* **18**, 200–209 (2008).

56. Ibbotson, M. R., Crowder, N. A., Cloherty, S. L., Price, N. S. C. & Mustari, M. J. Saccadic modulation of neural responses: possible roles in saccadic suppression, enhancement, and time compression. *J. Neurosci.* **28**, 10952–60 (2008).

57. Cloherty, S. L., Mustari, M. J., Rosa, M. G. P. & Ibbotson, M. R. Effects of saccades on visual processing in primate MSTd. *Vision Res.* **50**, 2683–2691 (2010).

58. Duhamel, J. R., Colby, C. L. & Goldberg, M. E. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* **255**, 90–2 (1992).

59. Sommer, M. A. & Wurtz, R. H. Influence of the thalamus on spatial visual processing in frontal cortex. *Nature* **444**, 374–377 (2006).

60. Münch, T. A. *et al.* Approach sensitivity in the retina processed by a multifunctional neural circuit. *Nat. Neurosci.* **12**, 1308–1316 (2009).

61. Farrow, K. *et al.* Ambient Illumination Toggles a Neuronal Circuit Switch in the Retina and Visual Perception at Cone Threshold. *Neuron* **78**, 325–338 (2013).

62. Tikidji-Hamburyan, A. *et al.* Retinal output changes qualitatively with every change in ambient illuminance. *Nat. Neurosci.* **18**, 66–74 (2015).

63. Reinhard, K. *et al.* Step-By-Step Instructions for Retina Recordings with Perforated Multi Electrode Arrays. *PLoS One* **9**, e106148 (2014).

64. Frey, U., Egert, U., Heer, F., Hafizovic, S. & Hierlemann, A. Microelectronic system for high-resolution mapping of extracellular electric fields applied to

brain slices. *Biosens. Bioelectron.* **24**, 2191–2198 (2009).

65. Müller, J. *et al.* High-resolution CMOS MEA platform to study neurons at subcellular, cellular, and network levels. *Lab Chip* **15**, 2767–2780 (2015).

66. Hafed, Z. M. Alteration of Visual Perception prior to Microsaccades. *Neuron* **77**, 775–786 (2013).

67. Grujic, N., Brehm, N., Gloge, C., Zhuo, W. & Hafed, Z. M. Perisaccadic perceptual mislocalization is different for upward saccades. *J. Neurophysiol.* **120**, 3198–3216 (2018).

68. Schwartz, G. W. *et al.* The spatial structure of a nonlinear receptive field. *Nat. Neurosci.* **15**, 1572–1580 (2012).

69. Zhang, Y., Kim, I.-J., Sanes, J. R. & Meister, M. The most numerous ganglion cell type of the mouse retina is a selective feature detector. *Proc. Natl. Acad. Sci.* **109**, E2391–E2398 (2012).

70. Sakatani, T. & Isa, T. Quantitative analysis of spontaneous saccade-like rapid eye movements in C57BL/6 mice. *Neurosci. Res.* **58**, 324–331 (2007).

71. Itokazu, T. *et al.* Streamlined sensory motor communication through cortical reciprocal connectivity in a visually guided eye movement task. *Nat. Commun.* **9**, 338 (2018).

72. Peli, E., Arend, L. E., Young, G. M. & Goldstein, R. B. Contrast sensitivity to patch stimuli: effects of spatial bandwidth and temporal presentation. *Spat. Vis.* **7**, 1–14 (1993).

73. Diggelmann, R., Fiscella, M., Hierlemann, A. & Franke, F. Automatic spike sorting for high-density microelectrode arrays. *J. Neurophysiol.* **120**, 3155–3171 (2018).

74. Idrees, S., Franke, F., Diggelmann, R., Hierlemann, A. & Münch, T. A. UnitBrowser - A Tool to Evaluate and Post-Process Units Sorted by Automatic Spike Sorting Algorithms. *Front. Neurosci.* **10**, (2016).

75. Bellet, M. E., Bellet, J., Nienborg, H., Hafed, Z. M. & Berens, P. Human-level saccade detection performance using deep neural networks. *J. Neurophysiol.* **121**, 646–661 (2019).

76. Schütt, H. H., Harmeling, S., Macke, J. H. & Wichmann, F. A. Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Res.* **122**, 105–123 (2016).

Figures

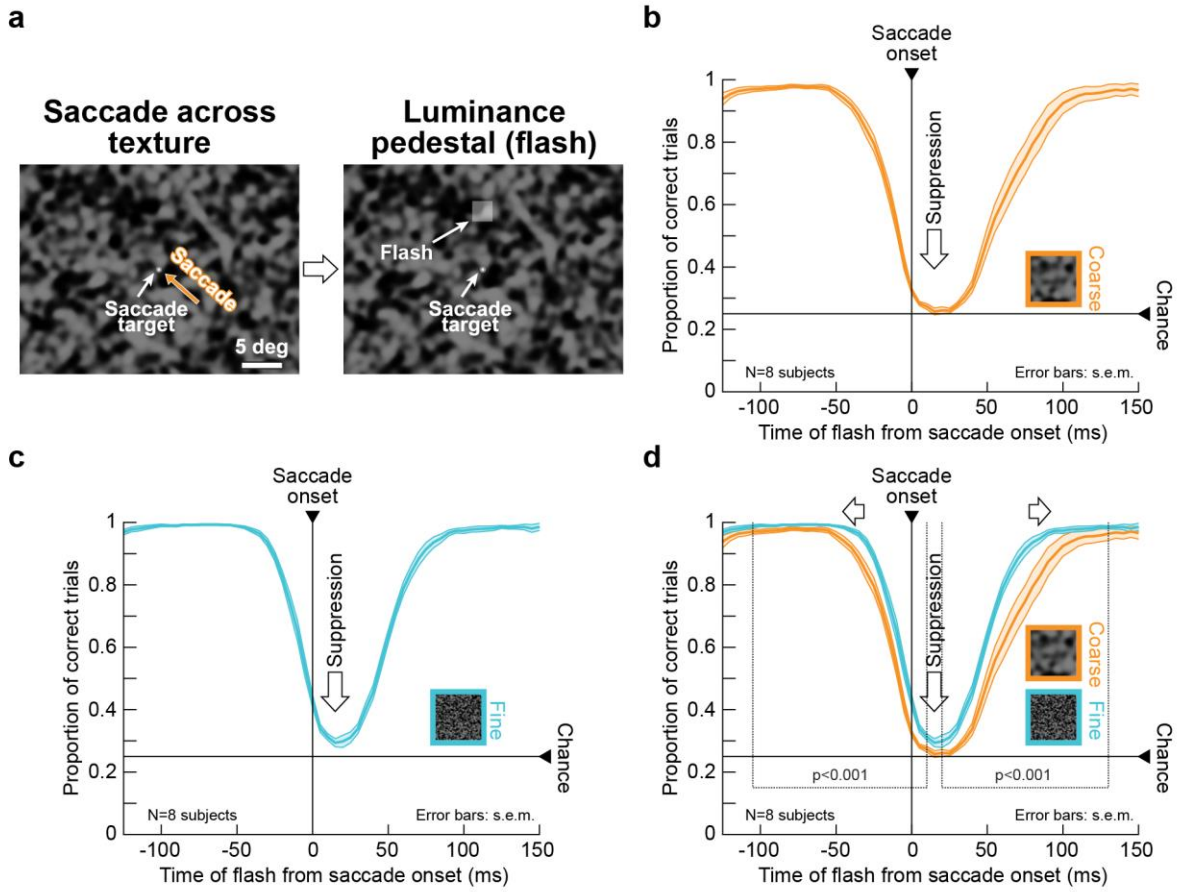


Figure 1 Image dependence of perceptual saccadic suppression. (a) Human subjects generated saccades from one of four diagonal locations towards display center (here: from the lower right). A luminance pedestal was flashed peri-saccadically at one of four locations around display center (right, left, up, or down; here: up). The example shows the coarse background texture (insets in **c**, **d** show fine textures for comparison; also see Supplementary Fig. 1 and Methods). **(b, c)** Subjects failed to localize peri-saccadic flashes with both coarse **(b)** and fine **(c)** textures (we binned perceptual reports as a function of flash time from saccade onset using 50-ms bins moved in steps of 5 ms). **(d)** Perceptual suppression started earlier and lasted longer with a coarse background (also see Fig. 2). The highlighted time points denote significantly different ($p<0.001$) time clusters between the coarse and fine conditions (Methods). Curves show averages (+/- s.e.m. bounds) of individual subjects' suppression curves. Supplementary Figs. 2, 3 show individual subject results, as well as controls for flash visibility (in the absence of saccades) and saccade motor variability.

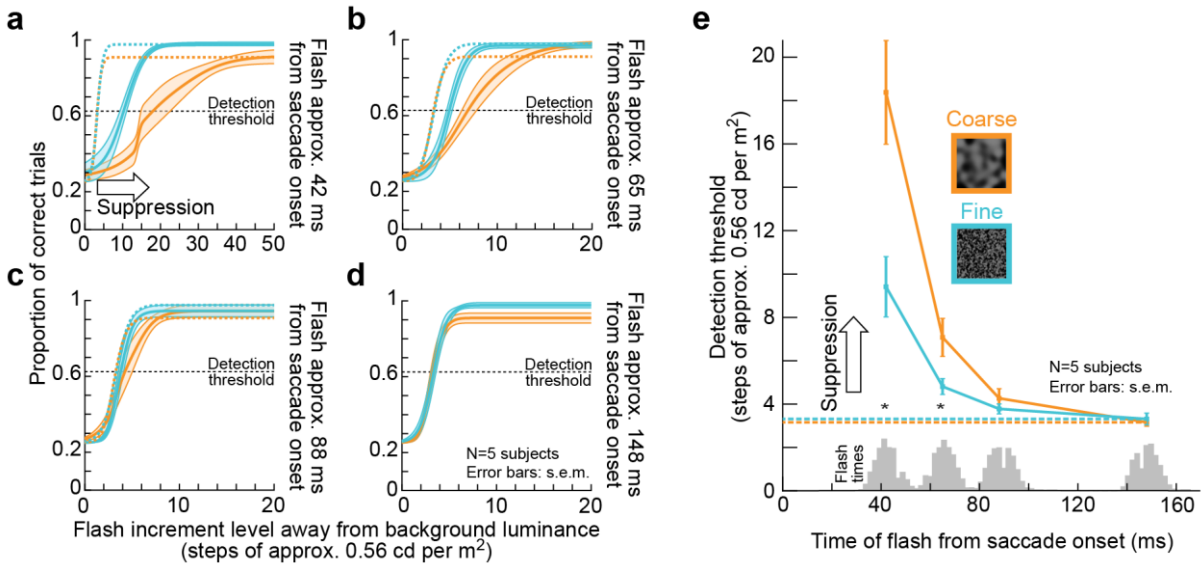


Figure 2 Image-dependent elevation of perceptual thresholds across saccades. (a-d) We repeated the experiment of Fig. 1 but collecting full psychometric curves of flash visibility. Solid curves: mean +/- s.e.m of the individual psychometric curves of five subjects (see Supplementary Fig. 4 for individual subject results). Dashed curves: psychometric curves near recovery from suppression long after saccades (same data as in **d**). Orange and light-blue indicate data for coarse and fine textures, respectively. (**a**) For flashes approximately 42 ms from saccade onset (Methods), strong perceptual saccadic suppression occurred (compare solid with dashed curves), and the psychometric curve for coarse textures was shifted to higher detection thresholds than that for fine textures, indicating stronger perceptual saccadic suppression. **(b)** At approximately 65 ms after saccade onset, substantial recovery was visible (note the different x-axis scale from **a**), but there was still stronger suppression for coarse than fine textures. **(c, d)** Recovery of visibility continued at even later times after saccade onset (88 ms, **c**, and 168 ms, **d**), consistent with Fig. 1. **(e)** Perceptual detection thresholds from **a**-**d** as a function of flash times from saccade onset. Since flash times were determined using online saccade detection (Methods), there was some variability of actual displayed flash times; the gray histograms on the x-axis show the actual distributions of flash times for each group of data from **a**-**d**. The results confirm the interpretation of Fig. 1: perceptual saccadic suppression was stronger and lasted longer for coarse than for fine background textures. Asterisks denote significant differences between coarse and fine textures (two-sample t-test; $p<0.05$). The dashed horizontal lines show the detection thresholds at the longest flash times (**d**); note that these thresholds are also similar to those in the visibility control experiments of Supplementary Fig. 3a, b.

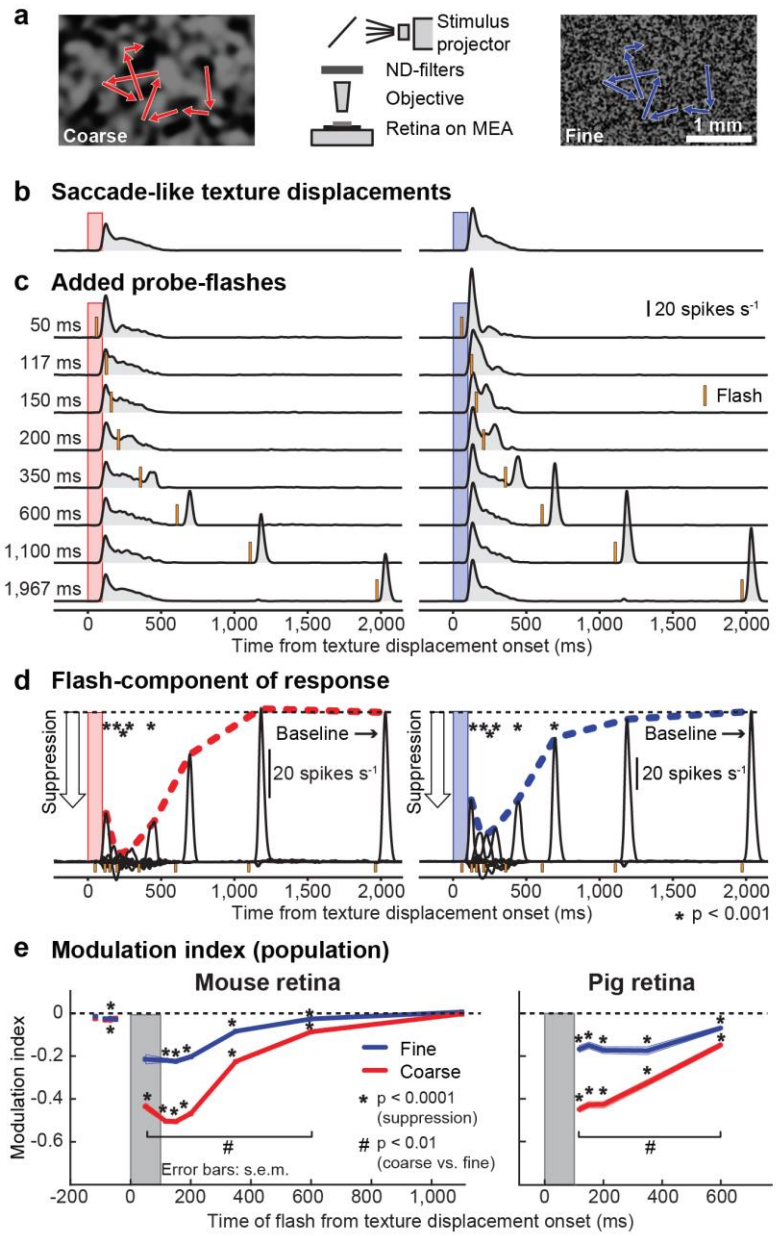


Figure 3 "Saccadic suppression" in retina. (a) We recorded RGC activity from *ex vivo* retinae placed on multielectrode arrays (MEA). A coarse (left) or fine (right) texture was repeatedly translated in a saccade-like manner (red or blue scan paths), and we presented brief visual flashes at different times relative to "saccades" (similar to Fig. 1). **(b, c)** Average activity of an example RGC to 39 texture displacements alone **(b)** or followed by probe flashes at different time delays **(c)**. Red and blue bars show the timings of the texture displacements; orange bars indicate probe flashes. Flash-induced responses were strongly suppressed immediately following saccade-like texture displacements. **(d)** Isolated flash responses obtained by subtracting responses in **b** from those in **c**. Dashed colored lines highlight the time courses of retinal "saccadic suppression" relative to baseline flash-induced responses. **(e)** Modulation index highlighting retinal "saccadic suppression" (Methods; negative values indicate suppressed flash-induced neural responses). Both mouse and pig retinae showed strong suppression during and after texture displacements, which also depended on texture statistics (similar to perception; Figs. 1, 2). Error bars denote s.e.m., and asterisks/hashes indicate statistical significance (Methods). The numbers of recorded cells at each flash time in **e** were as follows. Mouse RGCs: N=179 (-177 ms, -84 ms, -50 ms), 161 (-67 ms), 136 (50 ms), 527 (117 ms), 520 (150 ms), 502 (200 ms, 600 ms), 688 (350 ms), 345 (1,100 ms); pig RGCs: N=228 for each time point. Figure 4 shows additional properties of retinal "saccadic suppression", and Supplementary Figs. 5, 6 show the population data underlying panel **e** and different RGC types. Scale bars are defined in their respective panels.

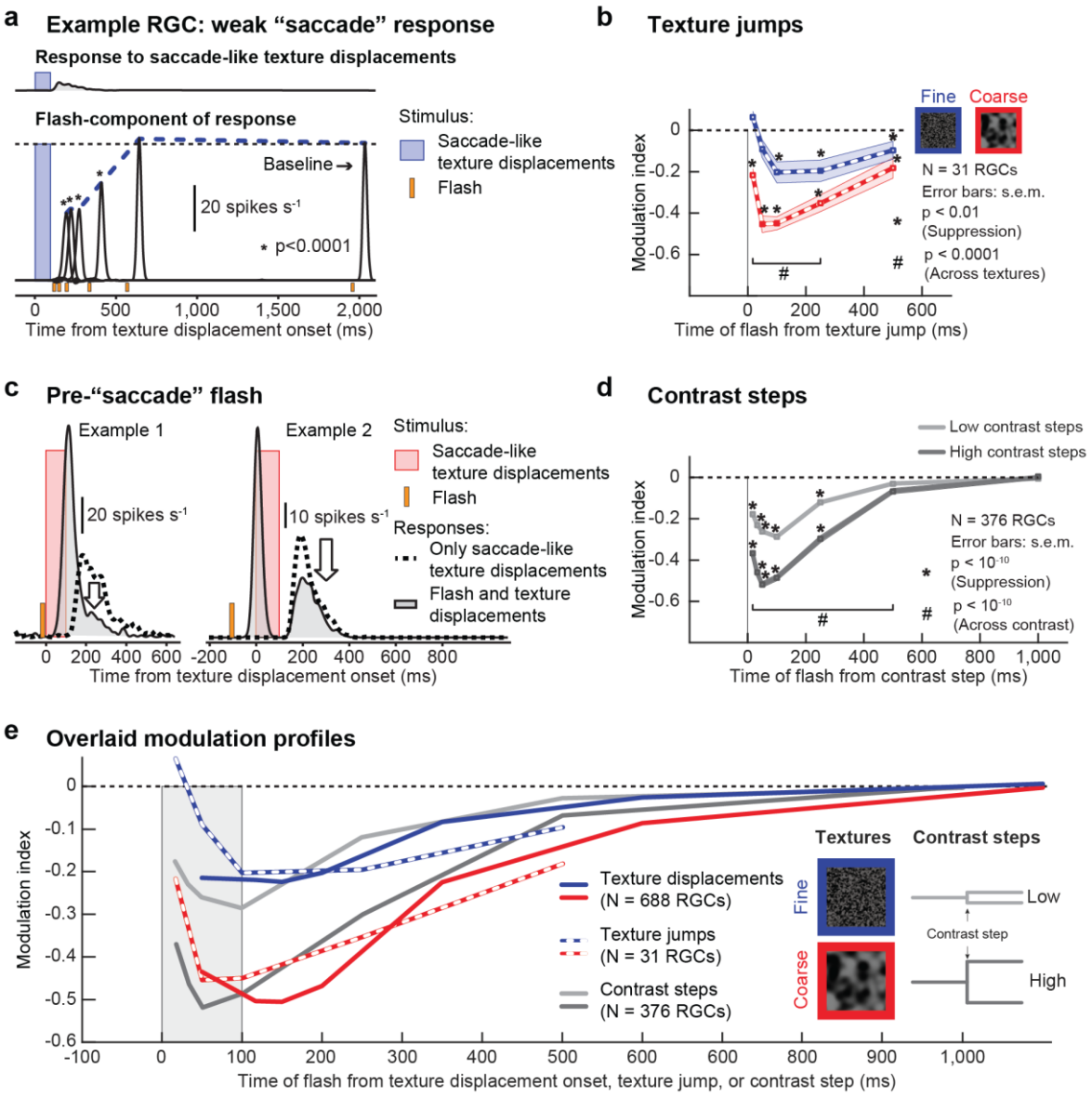


Figure 4 Stimulus-stimulus interactions in retinal "saccadic suppression". (a) Example RGC responding only weakly to texture displacements (top), but nevertheless exhibiting strong suppression of flash-induced neural responses (bottom; curves are plotted at the same scale). Suppression was much stronger than the response amplitude to the texture displacements alone. **(b)** Population modulation index (mean +/- s.e.m.) for a paradigm in which the textures jumped from their start to end positions instantaneously. Strong suppression (∗ p<0.01, two-tailed Wilcoxon signed-rank test) and significant differences between coarse (red) and fine textures (blue; # p<0.0001, Wilcoxon signed-rank test) were preserved. **(c)** Two example RGCs showing that a flash presented before saccade-like texture displacements suppressed the response to the displacements, supporting the notion that stimulus-stimulus interactions in the forward direction (first stimulus suppresses the response to the second stimulus) are the main drive for retinal "saccadic suppression". **(d)** Population modulation index (mean +/- s.e.m.) for a paradigm similar to panel **b**, but with textures replaced by spatially uniform backgrounds of different intensity. This created visual transients in the form of instantaneous contrast steps. Suppression of flash-induced responses was preserved (∗ $p<10^{-10}$, two-tailed Wilcoxon signed-rank test), and differences between low-contrast (light gray) and high-contrast (dark gray) steps (# $p<10^{-10}$, two-tailed Wilcoxon signed-rank test) resembled the differences between fine and coarse texture jumps in **b**. **(e)** Overlaid modulation profiles from saccade-like texture displacements (Fig. 3e), texture jumps (**b**), and contrast steps (**d**). Coarse texture displacements, coarse texture jumps, and high contrast steps had similar modulatory effects on probe flash responses; and so did fine texture displacements, fine texture jumps, and low contrast steps.

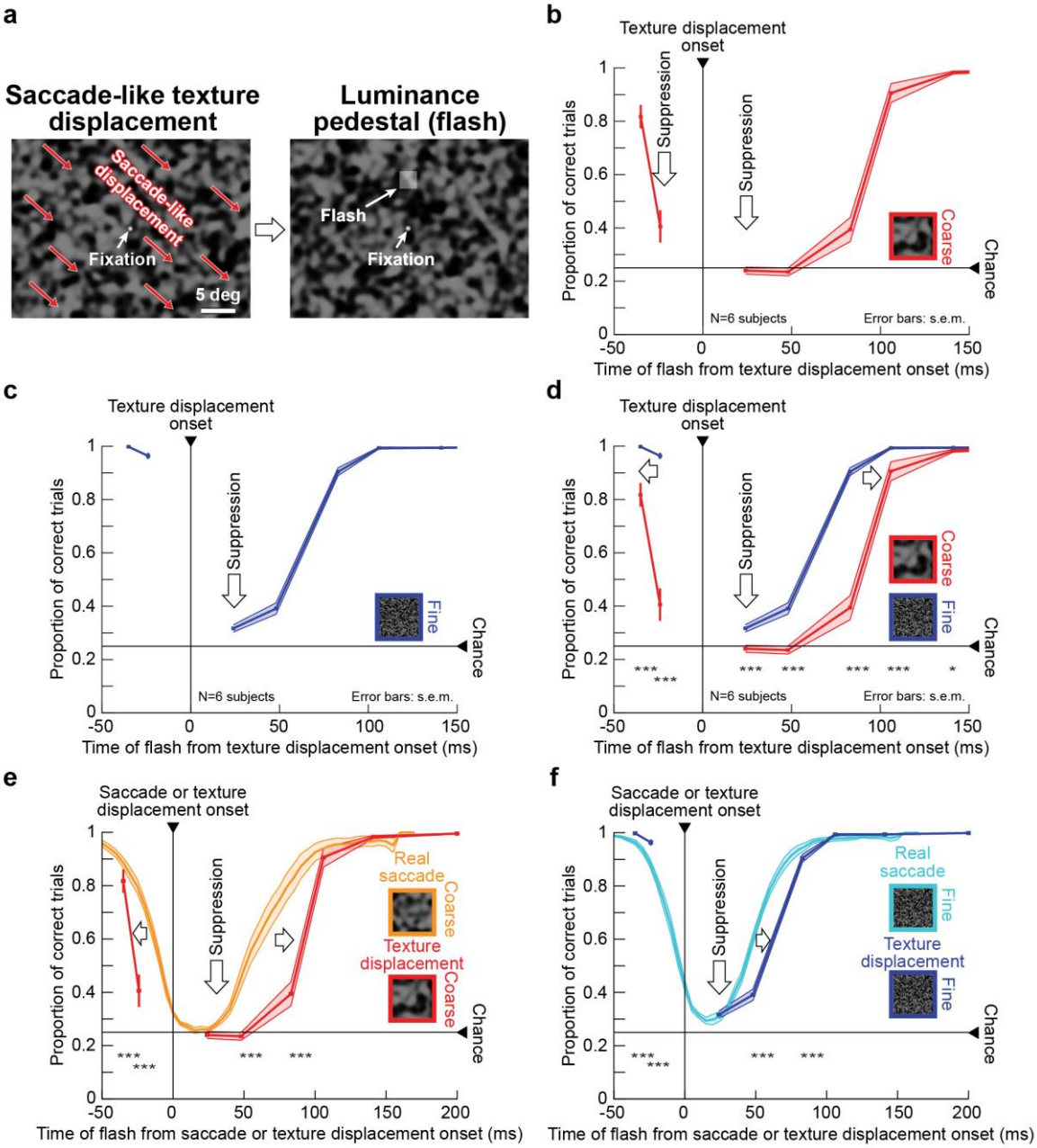


Figure 5 Image dependence of perceptual suppression without saccades. (a) Rapid texture displacements simulated saccade-like image displacements, similar to the retina experiments (Fig. 3). We used the same flashes and simulated saccade directions as in Fig. 1. The example shows a coarse texture (fine textures are shown in insets in **c**, **d**, and **f**). **(b, c)** Pre-, peri-, and post-displacement perceptual suppression occurred for both coarse **(b)** and fine **(c)** textures without real saccades. **(d)** As with real saccades (Fig. 1), suppression started earlier and lasted longer with coarse textures (also compare to similar retinal effects in Fig. 3e). Notably, pre-displacement suppression depended on texture statistics, just like with real saccades (Fig. 1). **(e, f)** Simulated saccades were associated with significantly longer suppression than real saccades for both fine and coarse textures. For coarse textures (**e**, which were most effective in causing suppression overall), flashes presented before the "saccade" event were suppressed earlier in the simulated saccade condition than in the real saccade condition (also see Fig. 6); thus, prolonged suppression with texture displacements was not restricted to post-displacement flashes only. Error bars denote s.e.m. across individual subjects' curves. Asterisks denote significant differences between coarse and fine textures (**d**) or between real and simulated saccades (**e**, **f**) at each indicated time point (χ^2 tests with Bonferroni corrections; ∗ p<0.005 in **d** and p<0.007 in **e**, **f**; ∗∗∗ p<0.0001 in **D** and p<0.00014 in **e**, **f**). Supplementary Fig. 2 shows individual subject results.

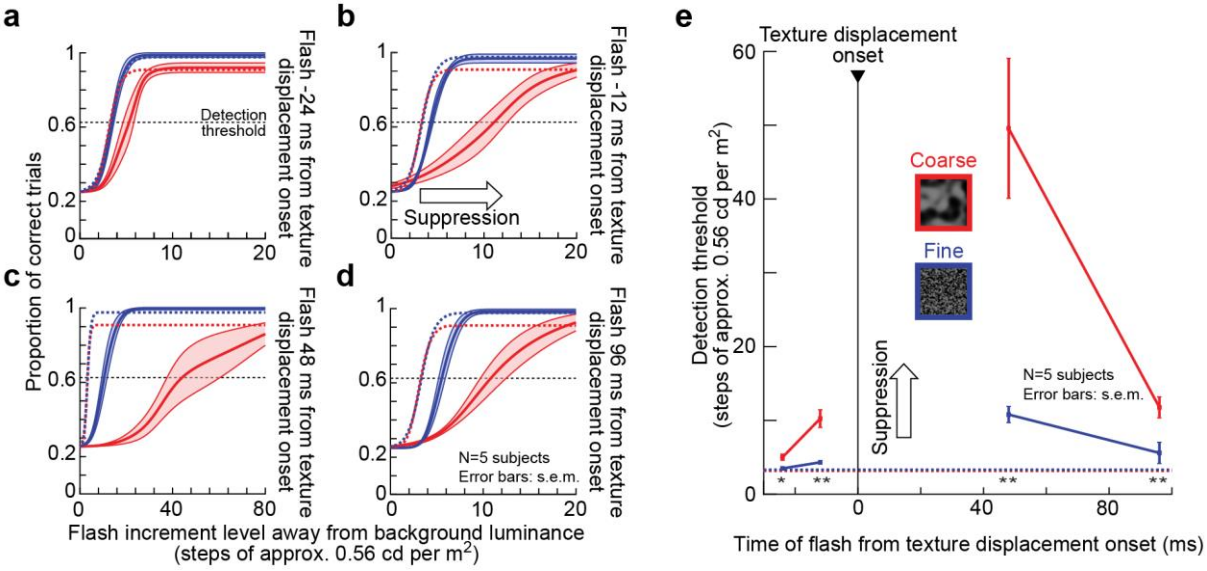


Figure 6 Image-dependent elevation of perceptual thresholds without saccades. (a-d) Similar to Fig. 2, we collected full psychometric curves of flash visibility around the time of simulated saccades (similar paradigm to Fig. 5). Solid curves: mean +/- s.e.m of individual psychometric curves of five subjects (see Supplementary Fig. 7 for individual subject results). Dashed curves: baseline data from the same subjects without simulated saccades and long after any real saccades (same data data as in Fig. 2d; also similar to Supplementary Fig. 3a, b with additional subjects). Red and blue indicate data for coarse and fine textures, respectively. (**a**) For a flash 24 ms before texture displacement onset, the red curve was shifted rightward towards higher flash contrasts (that is, reduced sensitivity) relative to baseline. This effect was much weaker with fine textures. **(b)** For a flash closer in time to the texture displacement but still before its onset (12 ms before displacement onset), both coarse and fine textures were associated with significant perceptual suppression relative to baseline, consistent with Fig. 5. Moreover, once again, suppression was stronger for coarse than fine textures (evidenced by the larger rightward shift in the psychometric curve). **(c)** Perceptual suppression was the strongest (note the different x-axis scale from the other panels) immediately after texture displacement onset. **(d)** 96 ms after texture displacement onset, there was still significant perceptual suppression, again significantly stronger for coarse than fine textures. This result is consistent with Fig 5, and highlights the longer-lasting suppression around simulated saccades compared to real saccades (Figs. 1, 2). **(e)** Detection thresholds from **a**-**d** as a function of flash time from texture displacement onset. Pre- and post-displacement perceptual suppression occurred, and suppression was stronger with coarse textures. Asterisks denote significant differences between coarse and fine textures (two-sample t-test; * $p<0.05$; ** $p<0.01$). Horizontal dashed lines show the baseline detection thresholds from Fig. 2d, e. All other conventions are similar to Figs. 1, 2, 5.

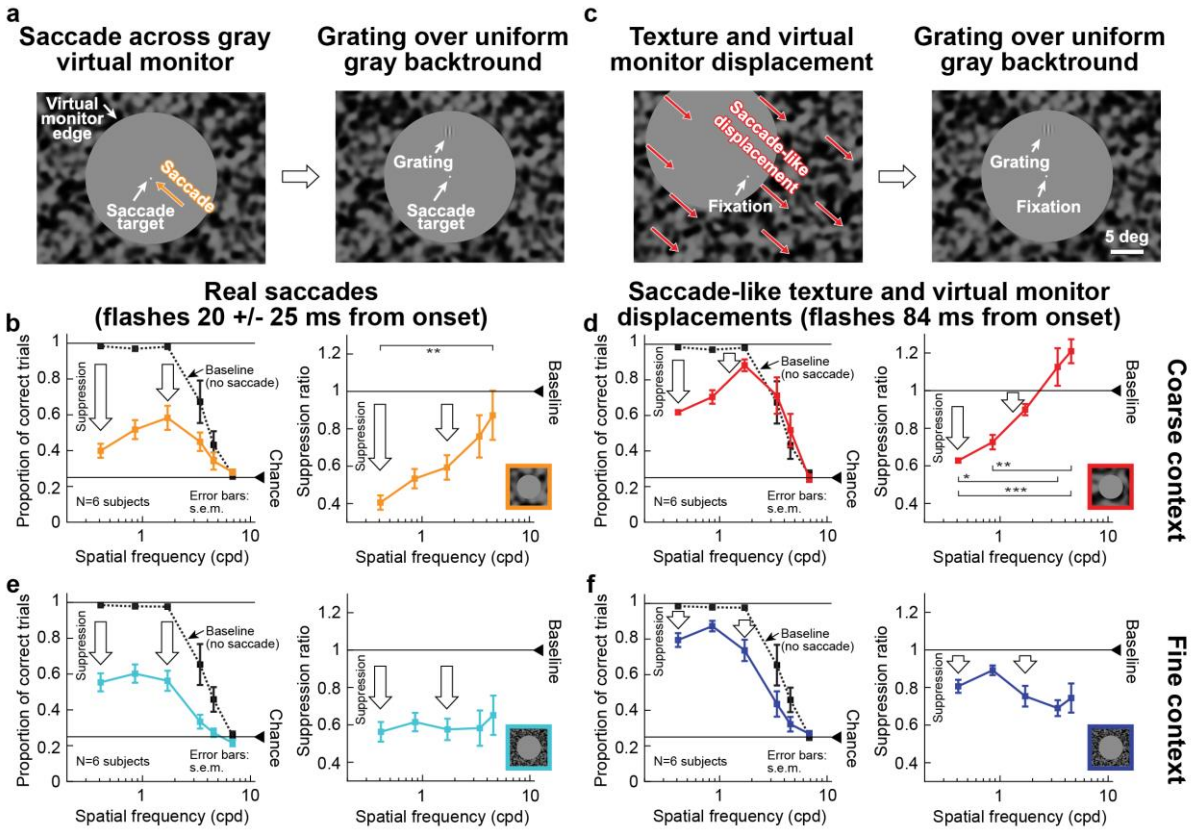
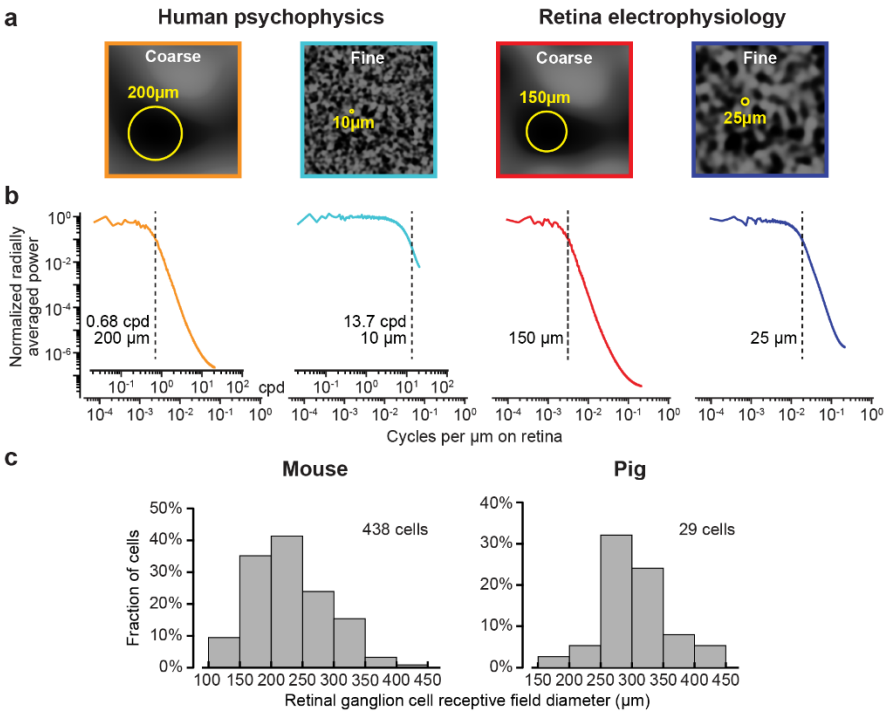
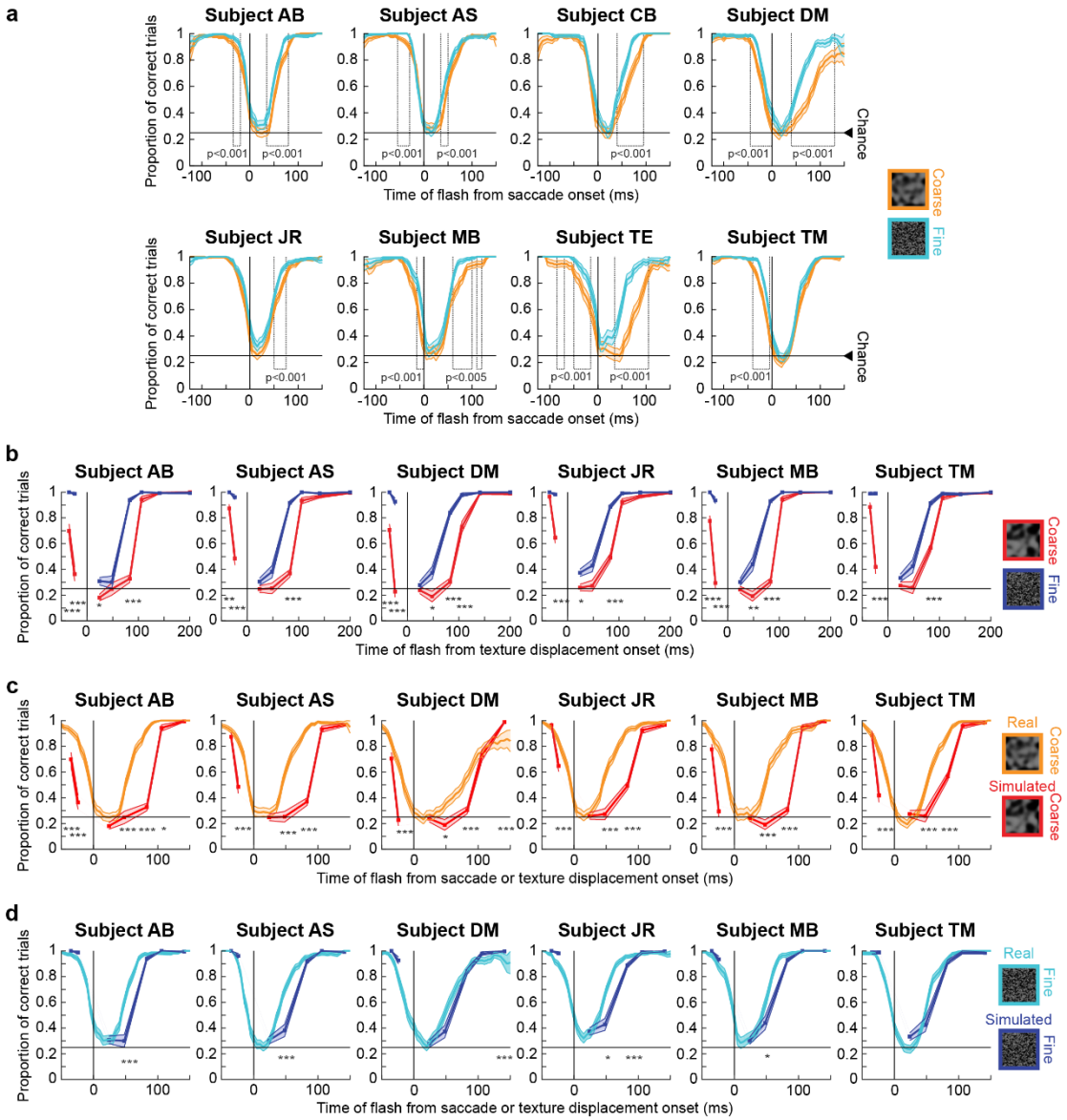


Figure 7 Selective peri-saccadic suppression of low spatial frequencies[10] is a visual phenomenon. (a) Left: Subjects made saccades towards display center. Right: gratings were flashed peri-saccadically over a uniform gray background (circular "virtual monitor" surrounded by a coarse texture; saccade directions and flash locations were similar to Figs. 1, 5). **(b)** Left: proportion of correct localizations of gratings with different spatial frequencies during fixation ("Baseline"; dashed curve) and for peri-saccadic flashes (solid curve). Low spatial frequencies were associated with the strongest suppression relative to baseline. Right: ratio of peri-saccadic to baseline performance (highest spatial frequency not shown because it was at chance performance even in baseline). Suppression depended on grating spatial frequency (χ^2=13.46, p=0.0092, df=4, Kruskal-Wallis test; ** p<0.01 for post-hoc pairwise comparisons between the lowest and highest spatial frequencies). **(c)** Left: we simulated saccade-induced image displacements by translating the virtual monitor and surrounding texture from one corner towards display center. Right: gratings appeared as in **a** (Methods). **(d)** The same selective suppression of low spatial frequencies as in **b** occurred. "Baseline" in this context means both no saccades and no virtual monitor and texture displacements. Suppression depended on spatial frequency (χ^2=25.33, p<0.0001, df=4, Kruskal-Wallis test; * p<0.05, ** p<0.01, *** p<0.001 for post-hoc pairwise comparisons between individual spatial frequencies). **(e, f)** With a fine surround texture, both real **(e)** and simulated **(f)** saccades were associated with suppression for all spatial frequencies; suppression selectivity[10] was eliminated (χ^2=0.8,p=0.938, df=4 for **e** and χ^2=7.74, p=0.102, df=4 for **f**, Kruskal-Wallis test). Error bars denote s.e.m. across individual subjects' curves. Supplementary Figs. 8-10 show full time courses as well as controls with black surrounds around the virtual monitor. Note that in **d**, **f**, we exploited the longer time course of visual suppression (Fig. 5, Supplementary Figs. 8, 9) to probe perception at a later time than in **b**, **e**. This also explains why suppression appeared quantitatively weaker in **d**, **f** than in **b**, **e**.

Supplementary information

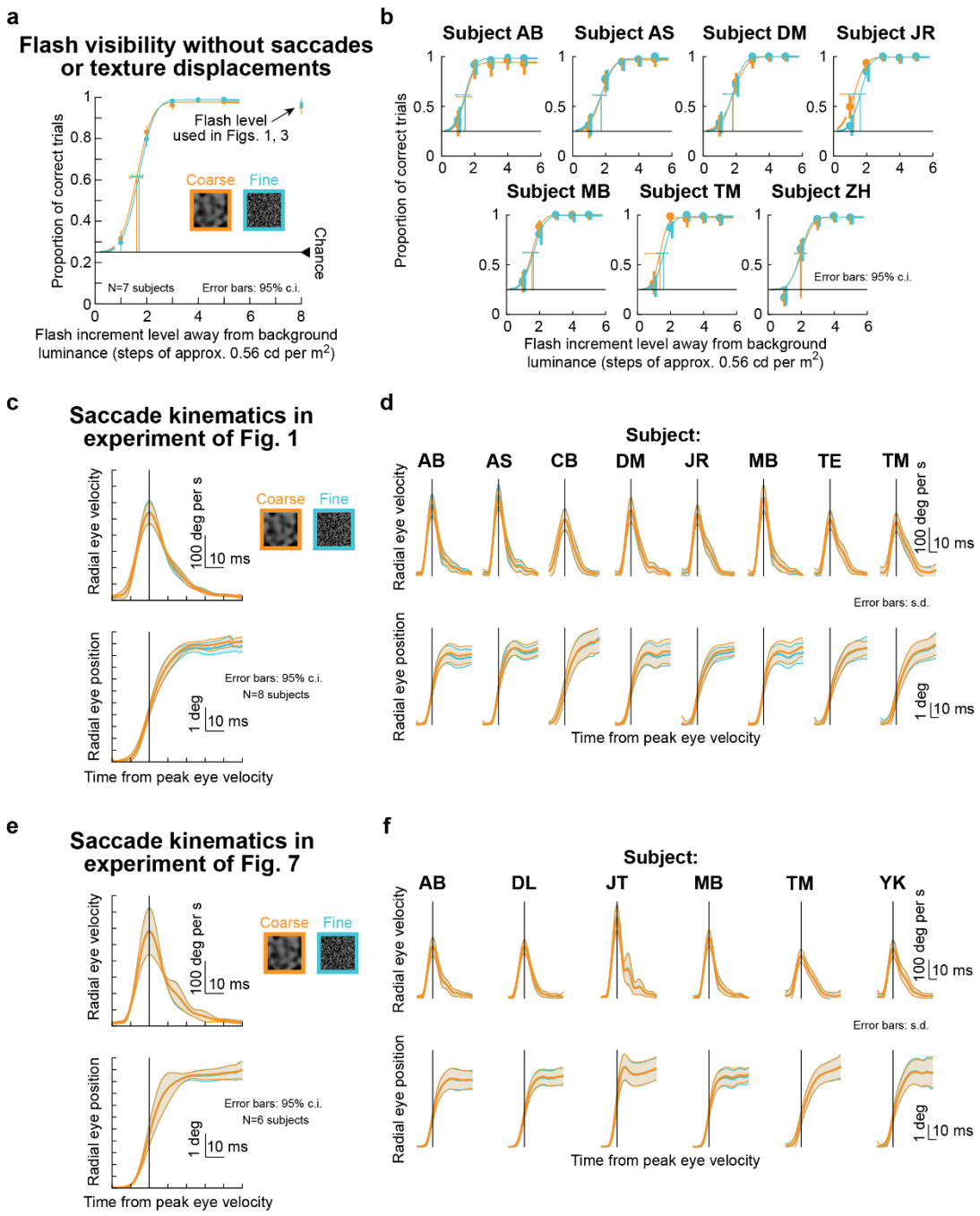


Supplementary Figure 1 Textured backgrounds tailored to receptive field sizes of retinal ganglion cells (coarse) or bipolar cells (fine) in the different species that we studied. (a) We created textures by convolving random binary pixel images with a Gaussian blurring filter. We varied the σ parameter of the Gaussian blurring filter (Methods) to define a so-called spatial scale for the resulting texture (indicated as yellow circles in the examples shown). For each species, we picked the spatial scale to result in dark or bright image blobs that approximated the sizes of either retinal ganglion cell (coarse) or bipolar cell (fine) receptive fields, and we then set σ to half the spatial scale value (Methods). **(b)** Radially-averaged power spectra for textures like in **a**, normalized to the maximum average power. Low-pass characteristics in all spatial scales were clear, as expected: less than 5% of the total average power was above the spatial frequency corresponding to the specific spatial scale of a given texture (vertical dashed lines). The inset x-axes in the first two spectra (used for human perceptual experiments) show units of cycles per degree (cpd) in addition to cycles per µm on the retina. **(c)** Histograms showing the distributions of receptive field diameters (Methods) in mouse (left) and pig (right) for a subset of retinal ganglion cells that we recorded. Since the distributions were generally similar, we used the same spatial scale parameter for the retinal recordings in both species. Human spatial scale parameters were estimated based on human receptive field diameters from the literature (Methods).



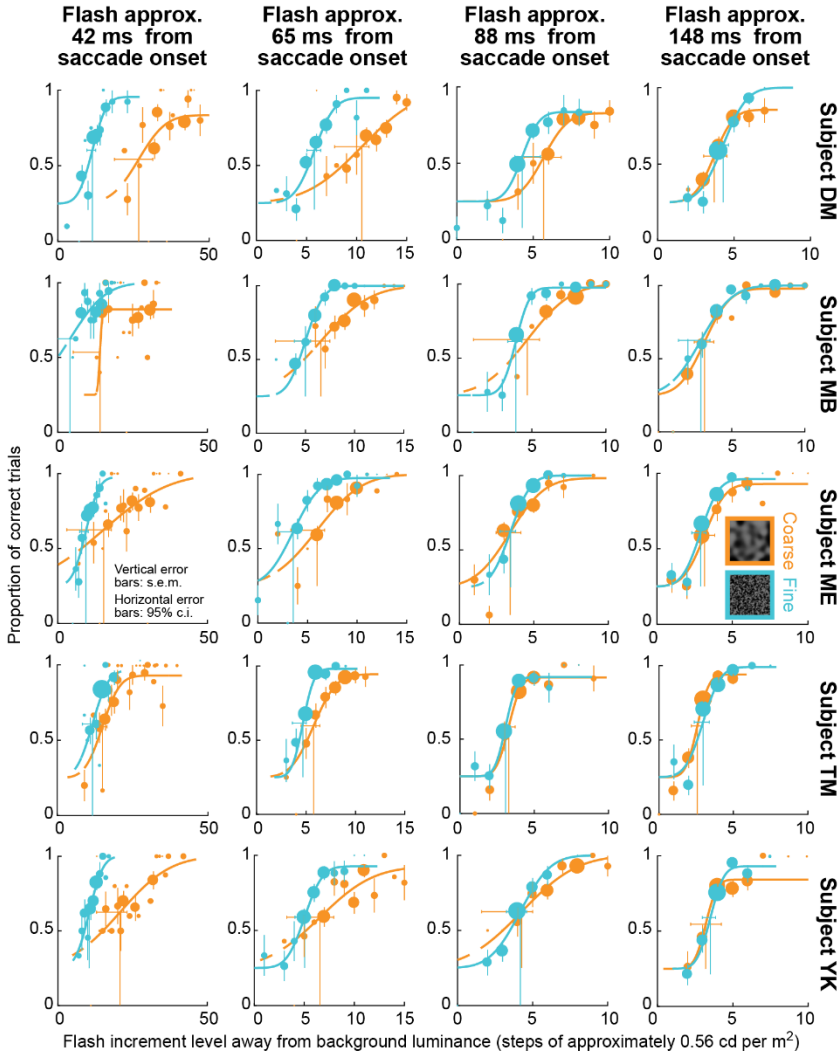
Supplementary Figure 2 Individual subject results from the perceptual experiments of Figs. 1, 5. (a) Identical analyses to Fig. 1d, shown separately for each individual subject. Error bars: s.e.m. across trials. All subjects experienced strong perceptual saccadic suppression, going from near-perfect localization performance to near-chance performance at peak suppression. Moreover, using strict statistical criteria (indicated in the figure and described in detail in Methods), all subjects had significant time clusters during which perception was different between saccadic suppression for saccades generated across coarse or fine textures. Also see Fig. 2 and Supplementary Fig. 4. **(b)** Same analyses as in Fig. 5d, but now showing individual subject results when saccades were replaced by saccade-like texture displacements during fixation. All subjects showed longer suppression after coarse texture displacements than after fine texture displacements; all subjects also showed earlier and stronger "pre-saccadic" suppression for coarse textures. Note that this "pre-saccadic" effect is purely visual, since the subjects never made saccades in this condition. Also, note that all subjects who participated in this experiment had also participated in the version with real saccades in **a**. Therefore, whether with or without saccades, perceptual suppression depended on image statistics. Also see Fig. 6 and

Supplementary Fig. 7. **(c, d)** Comparisons of perceptual suppression between real and simulated saccades across coarse **(c)** and fine **(d)** textures, as in Fig. 5e, f but now separating data from individual subjects. Note how even pre-saccadic suppression was prolonged in simulated relative to real saccades (i.e. started earlier in simulated saccades) in the coarse texture condition, which was most effective in causing suppression overall. Error bars: s.e.m. across trials. Asterisks in **b** denote a significant difference between coarse and fine conditions at the indicated flash time (χ^2 tests with Bonferroni corrections; $*$ $p<0.005$, $**$ $p<0.001$, $***$ $p<0.0001$). Asterisks in **c**, **d** denote significant differences ($*$ $p<0.007$, $**$ $p<0.0014$, $***$ $p<0.00014$) between real and simulated saccades, comparing perception of a flash at the indicated time delay after simulated saccades to the corresponding time bin (+/- 25 ms) from the real saccade condition.

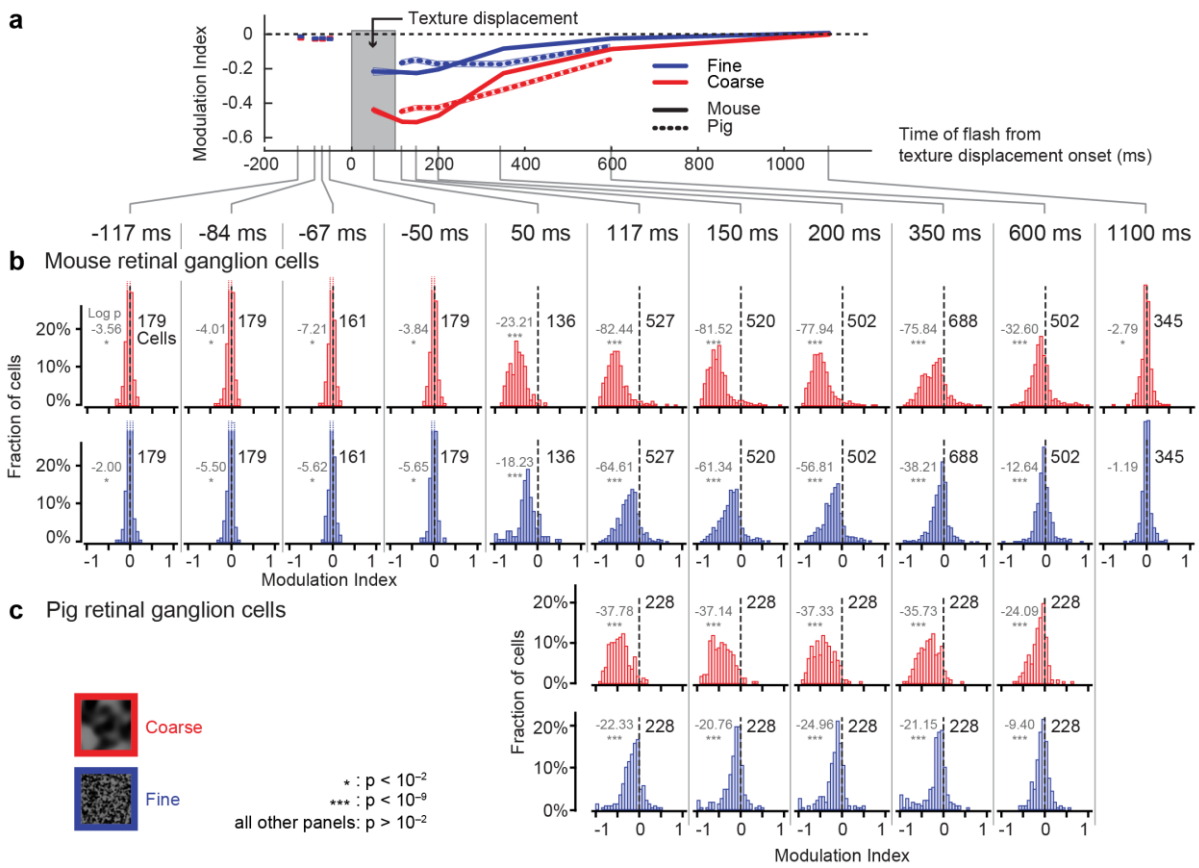


Supplementary Figure 3 Controls for flash visibility and motor variability in our perceptual experiments of Figs. 1, 5, 7. (a) For the same textures as in Figs. 1, 5, we asked subjects to maintain fixation. At a random time, a luminance pedestal appeared as in the main experiments (Figs. 1, 5), but this time, we varied its contrast across trials (Methods). We ensured that no microsaccades occurred near the flash onset time (Methods). Psychometric curves of localization performance indicate that, at the flash contrast used in Figs. 1, 5 (highlighted by the black arrow), subjects could easily detect flashes during simple fixation. Importantly, flash visibility was identical for coarse or fine textures at all contrasts. Therefore, flash visibility alone (or lack thereof) did not explain the main experiments' results (Figs. 1, 5). The strong perceptual suppression observed in Figs. 1, 5 was instead likely a function of interaction between visual transients associated with saccades or texture displacements and the flashes.

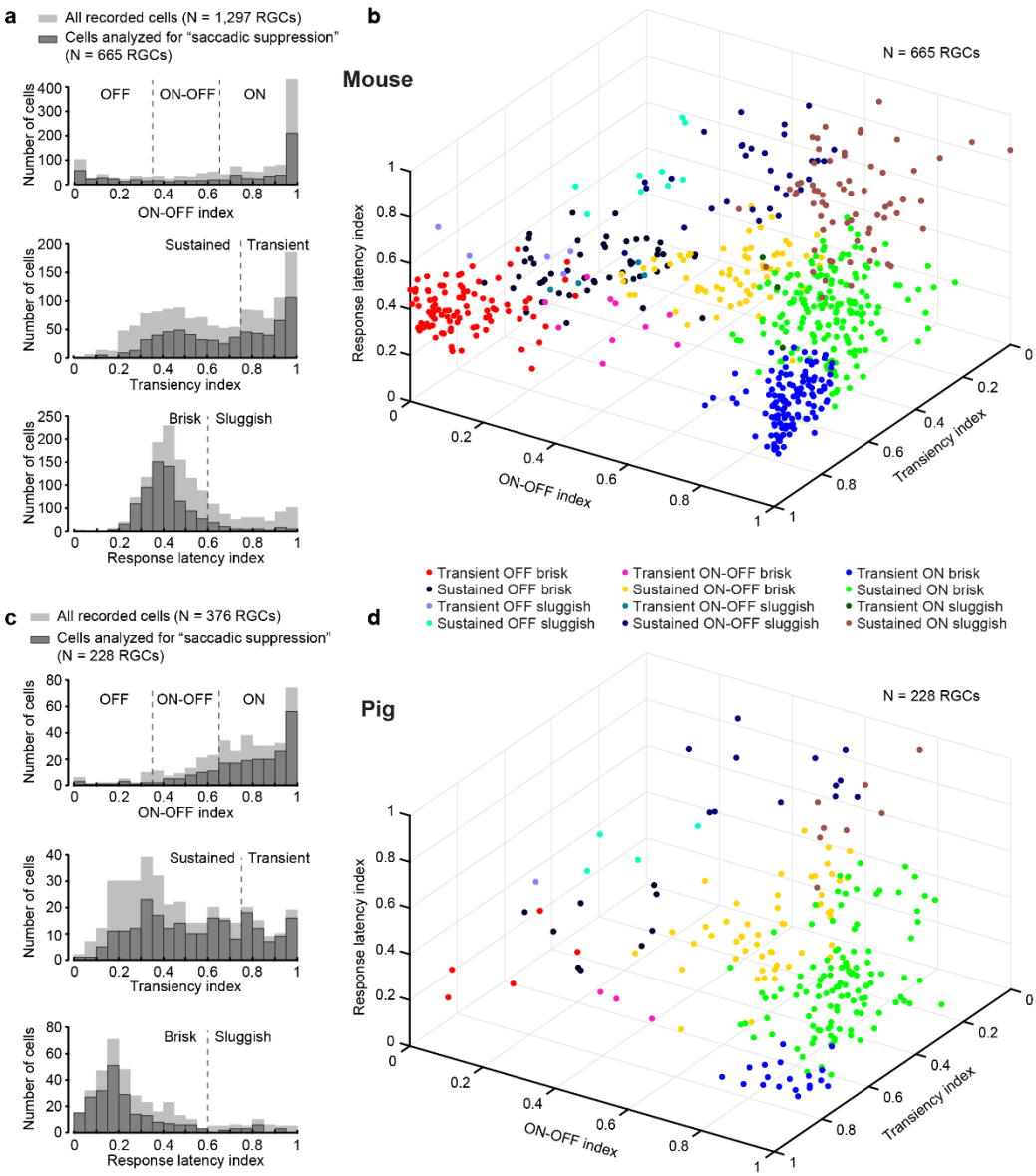
Also see Figs. 2, 6. **(b)** This idea is further supported by the fact that all individual subjects showed consistent results. All of these subjects had also participated in the experiments of Figs. 1, 5 (with the exception of subject ZH who only performed the control experiment). Psychometric curves were fit using the *psignifit 4 toolbox*[1], and error bars denote 95% confidence intervals. **(c)** We also checked for potential effects of motor variability on perceptual performance, in order to rule out the possibility that differences in performance between textures (Fig. 1) were due to differences in eye movement kinematics. For the experiments of Fig. 1, we plotted average radial eye velocity (top) and average radial eye position (bottom) across subjects (error bars denote 95% confidence intervals across the individual subjects' curves). There was no effect of background texture on movement kinematics. **(d)** This was also true for each subject. In this case, error bars denote s.d. across trials; saccade kinematics were not different when saccades were made across coarse or fine textures. **(e, f)** Same kinematic analyses, but now for the saccades of the experiment of Fig. 7. Scale bars are defined in their respective panels.



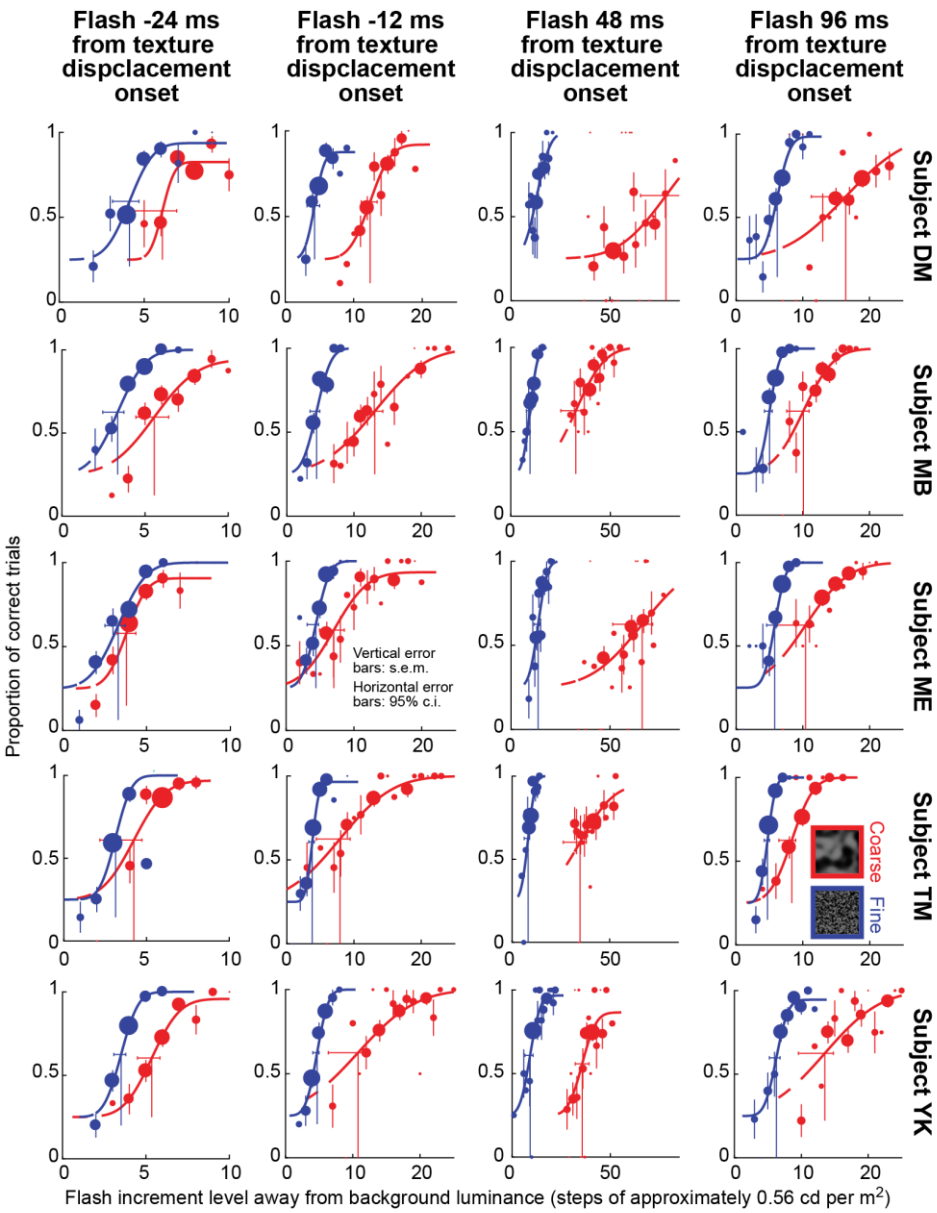
Supplementary Figure 4 Individual subject results from the perceptual experiment of Fig. 2. Each row shows psychometric curves like those shown in Fig. 2a-d, but for a single individual subject. Different rows show results from different subjects. The same conventions as in Fig. 2a-d apply. Here, we also scaled the size of each data point shown by the number of repetitions collected during the experiment. Note that we only show vertical error bars for data points with >10 repetitions, for clarity. Vertical error bars denote s.e.m. across repetitions of a given condition; horizontal error bars indicate 95% confidence intervals for the detection threshold of a given psychometric curve (i.e. the flash contrast resulting in threshold perceptual performance). Note that the x-axis ranges for the different columns (i.e. different flash times from saccade onset) are different from each other because of the varying amounts of perceptual saccadic suppression that occurred (Figs. 1-2). As can be seen, all subjects showed strong perceptual suppression near the time of saccade onset, with recovery occurring later in time, consistent with Fig. 1. Moreover, all subjects showed stronger perceptual suppression with coarse textures when compared to fine textures, again consistent with Fig. 1.



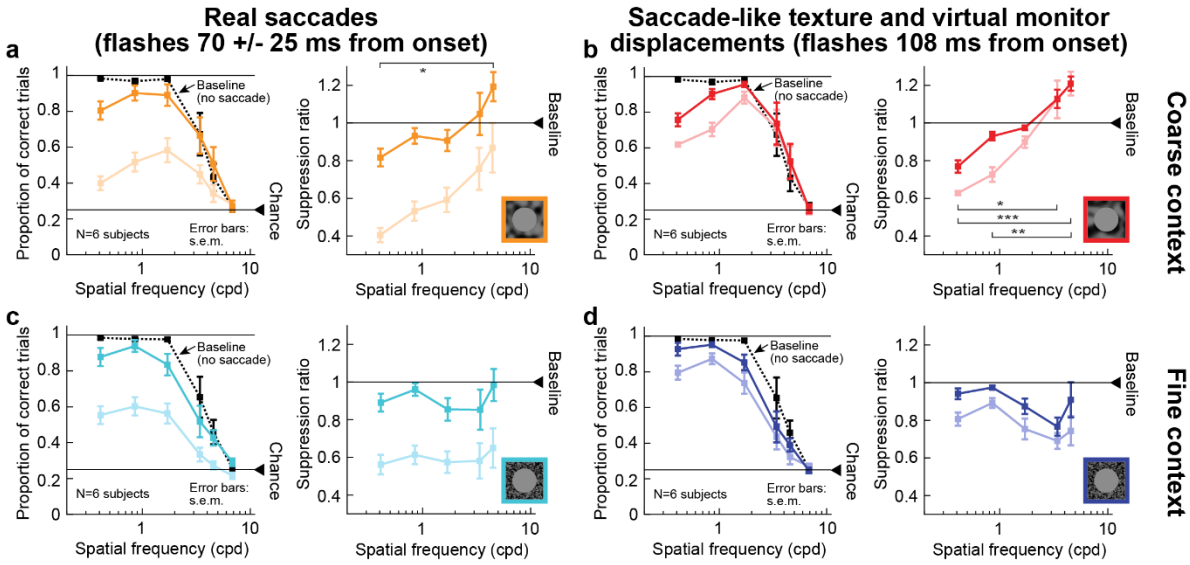
Supplementary Figure 5 Population data detailing the properties of retinal "saccadic suppression". (a) Replication of Fig. 3e, showing the time courses of retinal "saccadic suppression" in mouse and pig retinae. **(b, c)** Histograms of neuronal modulation indices for mouse **(b)** and pig **(c)** RGCs at different flash times relative to texture displacement onset. Red and blue denote coarse and fine textures, respectively. Black numbers in each panel indicate the numbers of RGCs analyzed for each condition; gray numbers show the logarithm (base 10) of the exact p-value (two-tailed Wilcoxon signed-rank test to determine if the population median was shifted away from 0). Asterisks additionally indicate the level of significance.



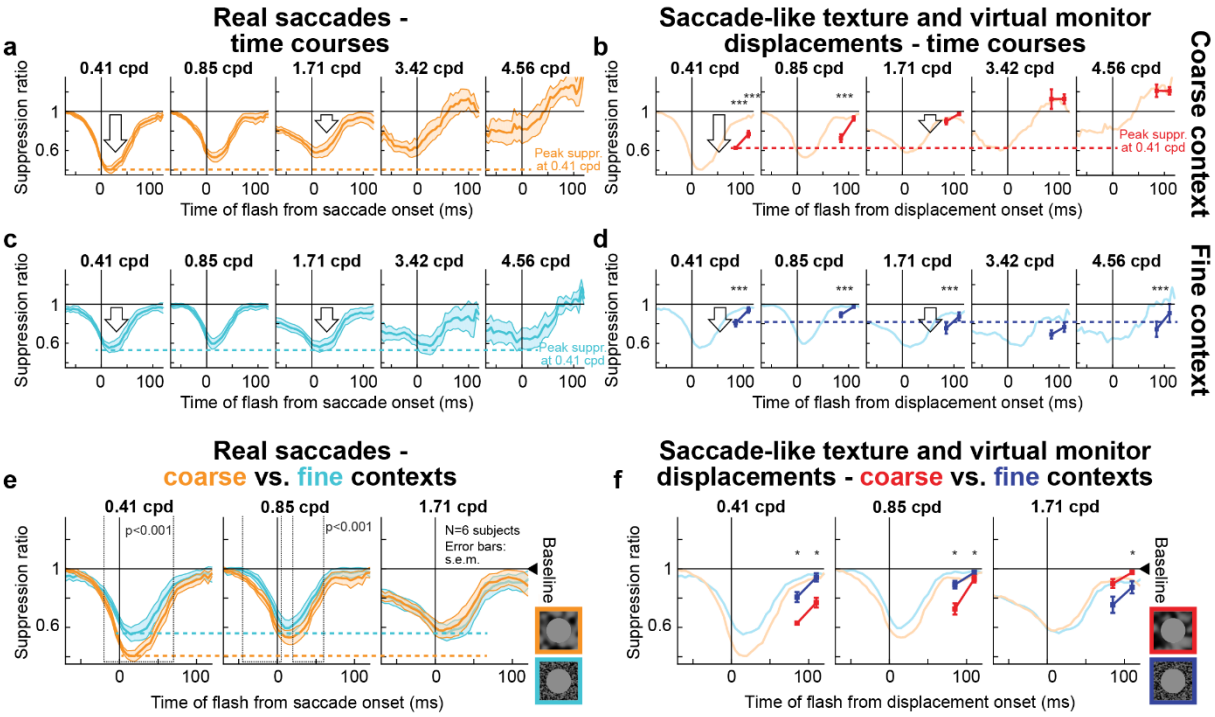
Supplementary Figure 6 Diverse properties of RGCs included in our analysis. We quantified the response properties of all recorded mouse **(a, b)** and pig **(c, d)** RGCs with respect to three neuronal response properties (see Methods): ON-OFF index, transiency index, and response latency index. Each histogram **(a**, **c)** was divided into 2 or 3 groups: RGCs could be OFF, ON-OFF, or ON (top histograms); transient or sustained (middle histograms); and brisk (short response latency) or sluggish (long response latency) (bottom histograms). Combined, this resulted in 12 response categories to which each recorded RGC belonged. The cells that could be analyzed for "saccadic suppression" and for which the response properties could be computed (dark gray histograms) spanned the entire range of response indices exhibited by all recorded cells for which these response properties were analyzed (light gray histograms). The three-dimensional scatter plots **(b**, **d)** show the projection of the RGC subsets considered in our analysis for "saccadic suppression" onto the 3 neuronal response indices. The 12 response categories, formed by the combination of histograms in **(a**, **c)**, can be seen in different colors.



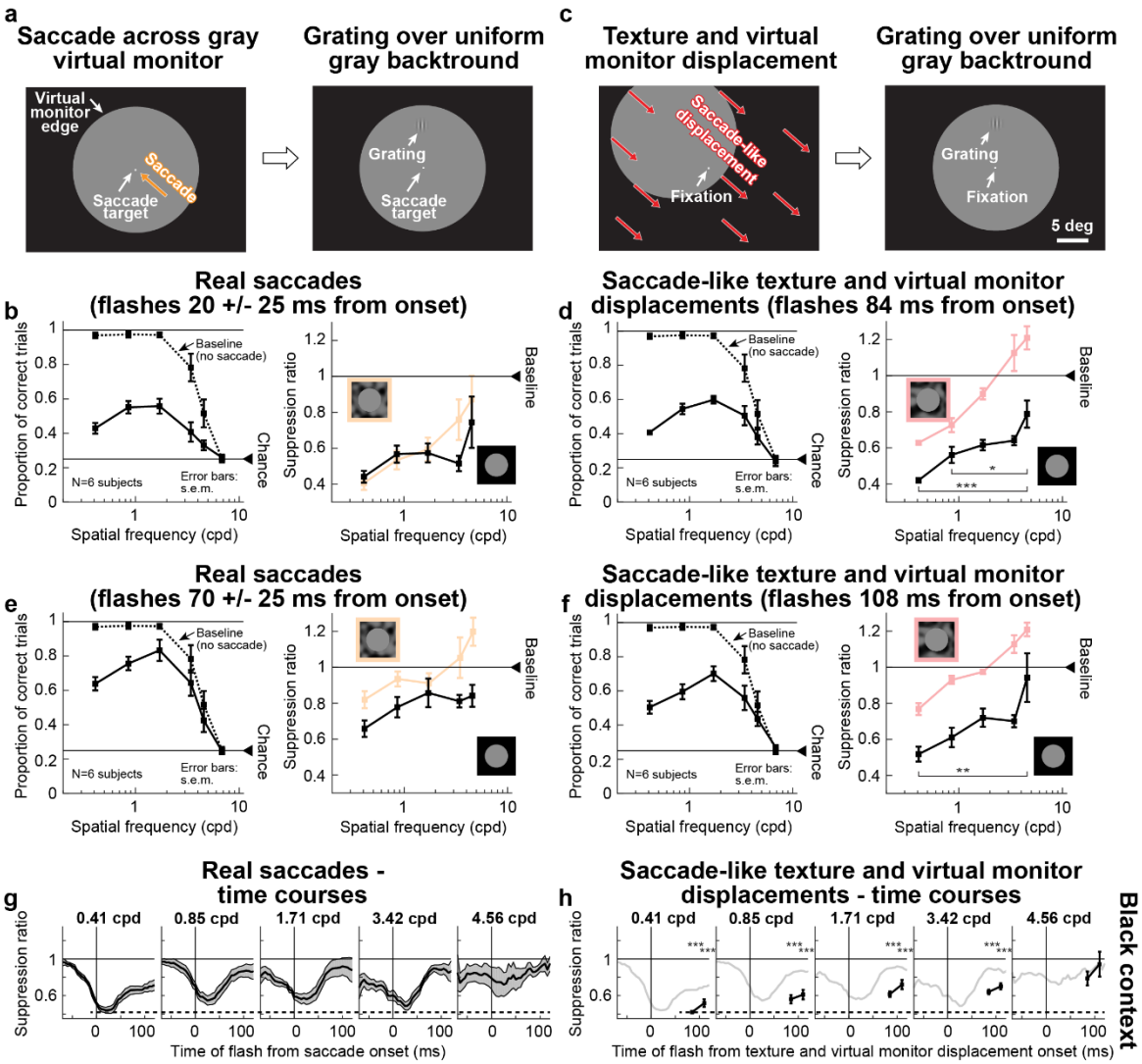
Supplementary Figure 7 Individual subject results from the perceptual experiment of Fig. 6. Same as Supplementary Fig. 4, but now for the experiment of Fig. 6. All subjects showed similar results: there was strong perceptual suppression before and after texture displacements in the absence of saccades, and the suppression effect was stronger when the displaced texture was coarse rather than fine. Note that in this experiment, we added an additional time sample prior to texture displacement onset, in comparison to Fig. 5, in order to demonstrate the robustness of this pre-displacement effect, and also to demonstrate the continuity of perceptual suppression in a time-locked fashion to texture displacement onset (Fig. 5).



Supplementary Figure 8 Recovery, with time, of perceptual suppression with real and simulated saccades in the experiment of Fig. 7. (a) Same analysis as in Fig. 7b, but at a later time point of grating flash onsets relative to saccade onset. Faint curves show the data from Fig. 7b for comparison. At around 70 ms after saccade onset, perceptual recovery from saccadic suppression emerged, but the selectivity of suppression across different spatial frequencies was still present (there was a main effect of spatial frequency on suppression ratio; χ^2=11.4, p=0.022, df=4, Kruskal-Wallis test). All other conventions are as in Fig. 7b (∗ p<0.05, post-hoc pairwise test between the lowest and highest spatial frequencies). **(b)** Same analysis as in Fig. 7d, but at a later time point of flash onset after virtual monitor and texture displacement. The same observations as in **a** were made: perceptual recovery occurred at the later time point, but selectivity of suppression was still obvious (χ^2=25.26, p<0.0001, df=4, Kruskal-Wallis test, same post hoc conventions as in Fig. 7d). The faint curves show the data from Fig. 7d for comparison. Note how this condition of displacements of the virtual monitor and texture surround resulted in longer lasting suppression than with real saccades (also see Supplementary Fig. 9). **(c, d)** Same analyses as in **a** and **b**, but with a fine texture surrounding the virtual monitor (Fig. 7e, f). Error bars in all panels denote s.e.m. All other conventions are as in Fig. 7.



Supplementary Fig. 9 Time courses of perceptual suppression with real and simulated saccades, as well as coarse and fine textures, in the experiment of Fig. 7. (a) Time courses of suppression from Fig. 7a, b with a coarse surround around the virtual monitor. We used similar binning procedures to Fig. 1. Peak suppression was strongest when 0.41 cpd gratings were flashed and progressively weakened for higher spatial frequency gratings (horizontal colored dashed line across panels). **(b)** With simulated saccade-like virtual monitor and texture displacements, we sampled two grating flash times relative to displacement onset. Recovery at the later time point for each grating spatial frequency was evident. Moreover, selectivity of suppression as a function of grating spatial frequency was evident (horizontal colored dashed line across panels demonstrating the peak suppression for the lowest spatial frequency). The faint curves show time courses from **a** for comparison. Note how simulated saccades caused longer-lasting suppression than real saccades, exactly as in the experiment of Fig. 5 (∗∗∗ $p<0.0001$, χ^2 tests with Bonferroni corrections comparing perceptual suppression with the simulated condition to a corresponding time bin in the real condition). **(c, d)** Similar analyses for fine texture surrounds around the virtual monitor. In this case, suppression was the same across all spatial frequencies (horizontal colored dashed lines across panels). **(e)** For real saccades, and for low spatial frequencies of gratings (i.e. when both coarse and fine surround contexts were associated with strong saccadic suppression), the coarse surround was associated with longer lasting suppression than the fine surround. This is consistent with the results of Fig. 1 when saccades were generated across full-screen textures. **(f)** This texture-dependence was also true with simulated saccades (∗ $p<0.05$, random permutation test comparing coarse and fine textures at a given grating flash time). Error bars in all panels denote s.e.m. All other conventions are as in Figs. 1, 5, 7.



Supplementary Fig. 10 Replicating the results of Fig. 7 but with black surrounds around a uniform gray display. (a) We repeated the same experiment as in Fig. 7a but this time using a black surround around the virtual monitor, as we might normally do in experiments on saccadic suppression[2,3]. Note that a black (or white) surround is theoretically equivalent to an infinitely coarse surround; hence, we expected observations more similar to Fig. 7a-d (i.e. selectivity of suppression for low spatial frequencies) than Fig. 7e, f. **(b)** Similar suppression selectivity for low spatial frequencies occurred with real saccades as in Fig. 7b (faint curves replicate that data for comparison). **(c)** Same experiment as in Fig. 7c, but with a black surround. **(d)** Selectivity of suppression for low spatial frequencies was even more evident with simulated saccades. Faint curves show results from Fig. 7d for comparison. **(e, f)** Similar analyses at a later time point, identical to Supplementary Fig. 8. There was recovery for both real (**e**) and simulated (**f**) saccades (faint colored curves show data from Supplementary Fig. 8a, b at the same time points for easier comparison). Note that with black surrounds, suppression strength was larger overall than with either coarse or fine texture surrounds (as if the black surround was indeed an extension of the coarseness of the texture). **(g, h)** Full time courses of suppression as in Supplementary Fig. 9. All error bars denote s.e.m., and all conventions are similar to Fig. 7 and Supplementary Figs. 8, 9.

Supplementary references

1 Schutt, H. H., Harmeling, S., Macke, J. H. & Wichmann, F. A. Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Res* **122**, 105-123 (2016).

2 Hafed, Z. M. & Krauzlis, R. J. Microsaccadic suppression of visual bursts in the primate superior colliculus. *J Neurosci* **30**, 9542-9547 (2010).

3 Chen, C. Y. & Hafed, Z. M. A neural locus for spatial-frequency specific saccadic suppression in visual-motor neurons of the primate superior colliculus. *J Neurophysiol* **117**, 1657-1673 (2017).