Retinal mechanisms of saccadic suppression and their role in processing sequential stimuli

Saad Idrees[1,2], Matthias-Philipp Baumann[1], Marili Korympidou[1,2,3], Timm Schubert[1,3], Alexandra Kling[4], Katrin Franke[3,5], Ziad M. Hafed[1], Felix Franke[6,7,8]*, Thomas A. Münch[1,3,9]*

[1]Werner Reichardt Centre for Integrative Neuroscience, University of Tübingen, 72076 Tübingen, Germany

[2]International Max Planck Research School, University of Tübingen, 72074 Tübingen, Germany

[3]Institute for Ophthalmic Research, University of Tübingen, 72076 Tübingen, Germany

[4]Department of Neurosurgery, Stanford School of Medicine, Stanford, CA 94305, USA

[5]Bernstein Center for Computational Neuroscience, University of Tübingen, 72076 Tübingen, Germany

[6]Bio Engineering Laboratory, ETH Zürich, 4058 Basel, Switzerland

[7]Institute of Molecular and Clinical Ophthalmology Basel, 4031 Basel, Switzerland

[8]Faculty of Science, University of Basel, 4056 Basel, Switzerland

[9]Lead Contact

*Correspondence: felfranke@googlemail.com (F.F.), thomas.muench@gmail.com (T.A.M.)

Summary

Visual perception remains stable across eye movements such as saccades, despite the concurrent strong disruptive visual flow. This stability is partially associated with a reduction in visual system sensitivity, a phenomenon known as saccadic suppression. Saccadic suppression already starts in the retina, where responses of ganglion cells to flashes are strongly suppressed after saccade-like image shifts. However, the retinal circuit mechanisms giving rise to such suppression remain unknown. Here, we describe these mechanisms using electrophysiology in mouse, pig, and macaque retina, 2-photon calcium imaging, computational modeling, and human psychophysics. We find that several mechanisms support retinal saccadic suppression, including a novel retinal processing motif, “dynamic reversal suppression”, which is triggered by sequential stimuli containing contrast reversals. This motif relies on nonlinear transformation of the inherently slow responses of cone photoreceptors. We describe two further components of suppression present in ON ganglion cells which originate in the cells’ receptive field surround, highlighting a novel disparity between ON and OFF ganglion cells. Our results increase our understanding of retinal processing of any time-varying stimulus.

Keywords: retina, ganglion cell, cone photoreceptor, saccadic suppression, computational model, sequential stimuli, rapid image shifts, perception, dynamic vision, visual processing

Introduction

Vision appears as a continuous and coherent process. This is a striking achievement of the visual system, considering that the visual flow across the retina is not continuous, but characterized by frequent and sudden changes, irregularities, and disruptions. The most prominent cause for such disruptions are eye movements. For example, as a result of saccades, the number of photons falling onto a given area of the retina can change by several orders of magnitude within tens of milliseconds. These fast eye movements can therefore cause sudden and frequent visual transients of local intensity across the entire retina. At the same time, saccades are critical for efficiently sampling the visual world (Schütz et al., 2011; Tatler et al., 2010; Yarbus, 1967), which is particularly true for species in which high visual resolution is limited to a small fraction of the overall visual space, such as the foveal region in primates.

Perceptually, saccadic disruptions are minimized by reducing the sensitivity of the visual system to new input around the time of saccades - a phenomenon known as saccadic suppression. While this phenomenon has been extensively characterized over the past few decades (Beeler, 1967; Bremmer et al., 2009; Idrees et al., 2020; Krekelberg, 2010; Matin, 1974; Volkmann, 1986; Wurtz, 2008; Zuber and Stark, 1966), its underlying mechanisms still remain unclear. Several electrophysiological studies have shown neural correlates of saccadic suppression throughout the visual system, namely a modulation of neural activity and/or sensitivity around the time of saccades (Bremmer et al., 2009; Chen and Hafed, 2017; Hafed and Krauzlis, 2010; Ibbotson et al., 2008; Kleiser et al., 2004; Krekelberg, 2010; Wurtz, 2008). These observations have often been interpreted to be caused by active suppressive signals originating from (pre-) motor areas, such as corollary discharge signals related to the saccadic eye movement command (Bremmer et al., 2009; Diamond et al., 2000; Duffy and Lombroso, 1968; Ross et al., 2001; Zuber and Stark, 1966). Most studies investigating the mechanisms of saccadic suppression have therefore focused on cortical or subcortical neuronal recordings and/or on behavioral measures of perceptual state, largely neglecting the consequence of visual processing in early visual pathways, for example in the retina.

The retina is an independent signal processing “front end” in the visual system, before visual information is sent along the optic nerve to higher brain areas. Consequently,

image processing triggered by visual transients, such as those that naturally occur during saccades, could potentially lead to altered retinal output. Retinal signal processing could therefore contribute to perceptual saccadic suppression. Some studies have investigated how the retina processes information in the context of spatio-temporal dynamics that occur during natural visual behavior (Appleby and Manookin, 2019; Baccus et al., 2008; Berry et al., 1999; Chen et al., 2013; Fried et al., 2002; Garvert and Gollisch, 2013; Geffen et al., 2007; Gollisch, 2013; Münch et al., 2009; Tikidji-Hamburyan et al., 2015; Wei, 2018; Zaghloul et al., 2007; Zhang et al., 2012), including saccades (Amthor et al., 2005; Barlow et al., 1977; Enroth-Cugell and Jakiela, 1980; Krishnamoorthy et al., 2017; Krueger and Fischer, 1973; Noda and Adey, 1974; Passaglia et al., 2009; Roska and Werblin, 2003). A retinal neural correlate of perceptual saccadic suppression has recently been shown by a previous study from our labs (Idrees et al., 2020). There, we showed that the retinal output is indeed altered by saccade-like image shifts. In most mouse and pig retinal ganglion cells (RGCs) that we recorded from, responses to brief probe flashes were suppressed when preceded by saccade-like image displacements across the retina. This retinal saccadic suppression had properties consistent with the perceptual suppression of probe flashes reported by human subjects using similar images, and following either real or simulated saccades. In fact, we already observed elementary properties of perceptual saccadic suppression, such as its dependency on background scene statistics, at the level of the output of the retina, providing strong evidence of a retinal mechanism directly contributing to perceptual saccadic suppression.

In this study, we describe such a mechanism. We experimentally mimicked the visual flow resulting from saccades and recorded the neural activity of the output neurons of the retina (RGCs) from ex vivo retinae of mice, pigs, and macaque monkeys. We found that retinal saccadic suppression was the result of multiple mechanisms. A major component originated from a novel visual processing motif within an RGC's receptive field center, which we call "dynamic reversal suppression". Dynamic reversal suppression results from a combination of inherent response properties of cone photoreceptors and non-linear processing of cone output by downstream pathways in the retina. Two further components of suppression originated from beyond the RGC's receptive field center, one of them driven by GABAergic inhibition. Interestingly, these

105 two additional components were observed primarily in ON RGCs, highlighting a novel
106 disparity between ON and OFF type RGCs. Perhaps one of the most intriguing
107 outcomes of this study, also consistent with observations of perception (Idrees et al.,
108 2020), is that the suppressive effects observed in RGCs are not exclusively triggered
109 by saccades, but occur for many scenarios involving sequential visual stimulation,
110 which are ever-present during natural vision. Therefore, while the results described
111 here are crucial for understanding the mechanisms of saccadic suppression, they also
112 elucidate more general mechanisms of retinal signal processing across any
113 time-varying visual input over short time scales (10 ms to 1 s).

114 # Results

115 ## Experimental Approach

116 We measured the modulation of retinal ganglion cell (RGC) output following
117 saccade-like changes of the visual input with a variety of different light stimulation
118 strategies (Fig. S1). The basic experimental paradigm was similar to that described in
119 (Idrees et al., 2020). Briefly, we first recorded spiking activity of RGCs in isolated ex
120 vivo mouse retinae using both high-density and low-density multi-electrode arrays
121 (MEAs). Each retina was exposed to a background texture having one of several
122 possible spatial scales that defined its spatial spectrum ("fine" to "coarse", Methods,
123 Fig. S2). We simulated saccade-like image displacements by rapidly translating the
124 texture globally across the retina (Methods; Fig. 1a). Most RGCs responded robustly
125 to such saccade-like texture displacements (see Fig. 1b for responses of example ON
126 and OFF RGCs). At different times relative to the saccade-like texture displacements
127 ("saccades" from now on), we presented a brief probe flash (Fig. 1c). We then
128 analyzed how the response to this probe flash (spike rate of the RGC) was influenced
129 by the preceding saccade, by comparing it to the response to the flash presented in
130 isolation (baseline). To quantify RGC response modulation, we calculated a
131 modulation index (Methods) as follows: we first isolated the flash-induced response
132 component by subtracting the saccade-only response (e.g., Fig. 1b) from the response
133 to the composite saccade-flash stimulus (e.g., Fig. 1c). Based on this flash-induced
134 response component (Fig. 1d), we calculated the modulation index as $(r_d - r_b)/(r_d + r_b)$.
135 Here, r_d is the peak response to the probe flash presented with a delay d relative to

136 saccade onset, and r_b is the baseline (peak response to the flash presented ~2 s after
137 the saccade). This modulation index is negative when flash-induced responses are
138 suppressed (Fig. 1d shows, on the horizontal dashed line, the modulation index for the
139 responses at each flash-time). In yet further recordings we applied various
140 manipulations to this base paradigm to probe for the mechanisms underlying
141 modulation of RGC responses following saccades. To generalize our findings across
142 other species, we also performed similar analyses of pig and macaque RGC data.

143 **Similarities and differences in retinal saccadic suppression across ON and OFF**
144 **type RGCs**

145 Suppression was robust across most RGCs that we recorded from, consistent with
146 what we reported previously (Idrees et al., 2020). Here, we more closely inspected
147 functionally different RGCs. Specifically, throughout this study, we divided RGCs into
148 ON and OFF types, i.e. into RGCs responding best to light increments or decrements,
149 respectively (Methods). We always quantified the modulation index defined above for
150 ON RGCs based on their responses to bright probe flashes and for OFF RGCs based
151 on their responses to dark probe flashes (Fig. 1c, d). Flash responses following a
152 saccade were suppressed in both ON and OFF RGCs. Fig. 1d shows example cells,
153 Fig. 1e the temporal profile of the mean population modulation index, and Fig. S3 the
154 underlying population data. Suppression was consistently stronger for coarser
155 background textures (Figs. 1e, S3), for both ON and OFF RGCs. This is consistent
156 with (Idrees et al., 2020), where we showed that this dependency on the texture
157 statistics can be explained by the distinct statistics of luminance and contrast changes
158 when coarse or fine textures move across the RGCs' receptive fields. However, a
159 striking difference existed in suppression recovery times: OFF RGCs recovered by
160 ~350 ms after saccade onset, whereas ON RGCs recovered only by ~1 s. Similar
161 results were obtained under scotopic conditions for coarse textures, while suppression
162 for fine textures was very weak (Fig. S4; all other mouse retina data was recorded at
163 mesopic conditions). In general, the presence of post-saccadic suppression of probe
164 flash responses in both ON and OFF type RGCs suggests a common mechanistic
165 theme across these cell types (Idrees et al., 2020). On the other hand, the different
166 recovery times indicate either additional suppressive mechanisms in ON RGCs or
167 additional recovery mechanisms in OFF RGCs.

168 **Spatial origin of retinal saccadic suppression**

169 ***Global component of suppression***

170 To probe the mechanisms underlying suppression and its differences across ON and
171 OFF type RGCs, we first examined the spatial origin of suppression. We hypothesized
172 that suppression of flash responses was caused by circuits detecting rapid global shifts
173 across the retina. Typically, these circuits include a lateral network of interneurons,
174 communicating with RGCs even from beyond their classical center-surround receptive
175 field (i.e., from their periphery, or far surround) (Lin and Masland, 2006; Roska and
176 Werblin, 2003). To test whether suppression was caused by such circuits, we modified
177 the spatial layout of the paradigm: we placed a square mask of 1000 x 1000 μm^2
178 (Fig. 2a, right) to restrict the saccades to the periphery of an RGC's receptive field.
179 Similar to the previous experiments, the probe flash was either a dark or bright flash
180 presented over the entire retina, including the masked region. Figure 2b shows the
181 mean population modulation indices of ON RGCs (top) and OFF RGCs (bottom) from
182 these experiments (Fig. S5 depicts the underlying population data and shows
183 responses of representative ON and OFF RGCs from these experiments). In OFF
184 RGCs, responses to full-field probe flashes were no longer suppressed when
185 saccades were restricted to the periphery. The responses of ON RGCs, on the other
186 hand, were still suppressed in this condition. The resulting suppression was however
187 weaker and short-lived (recovered by 350 ms) than with full-field saccades. These
188 observations (Figs. 2b, S5c) were robust across ON and OFF RGCs whose receptive
189 fields were completely contained within the mask (Fig. S5b).

190 We will refer to this component of suppression in ON RGCs, which originates from the
191 periphery, as the "global component" from now on. Such spatially far-reaching
192 inhibition is often mediated through GABAergic wide-field amacrine cells. We tested
193 this hypothesis by blocking $GABA_A$ receptors. Indeed, in the periphery saccade
194 condition, the modulation index for most ON RGCs was around 0 in the presence of
195 the $GABA_A$ receptor antagonist SR-95531 (Fig. 2b, S5d). These results suggest that
196 this short-lived global component of suppression is caused by inhibition via GABAergic
197 amacrine cells, perhaps similar to the polyaxonal amacrine cells described previously
198 (Ölveczky et al., 2003; Roska and Werblin, 2003). Thus, suppression is indeed
199 partially caused by circuits detecting global changes across the retina. However, those

200 circuits seem to act predominantly on ON RGCs, and even there, they do not account
201 for the total suppression observed with full-field saccades (without mask), which lasts
202 longer. Other, probably more local sources of suppression must exist that account for
203 the remaining suppression in ON RGCs and all of the suppression in OFF RGCs.

204 ***Local components of suppression***

205 To understand the more local components of suppression, we used different analyses
206 and manipulations of the main experimental paradigm. As we will see below, the local
207 components can be subdivided into a "central" and a "surround" component. First, we
208 eliminated the global component, by repeating our normal full-field saccade paradigm
209 in the presence of GABA receptor blockers. The suppression profile of both ON and
210 OFF RGCs was only weakly affected upon blocking $GABA_{A,C}$ receptors (5 μM
211 SR-95531 and 100 μM Picrotoxin; Fig. 2c, S6a). Since the GABA-block eliminates the
212 global component of suppression, the remaining more local components did not seem
213 to rely on GABAergic inhibition. Also, this suggests that the local components
214 dominate retinal saccadic suppression under full-field conditions. We then also
215 blocked glycine receptors (1 μM of Strychnine; Fig. 2d, S6b) to test if the local
216 components of suppression were caused by local inhibition via glycinergic pathways.
217 Here again, the suppression profiles of both ON and OFF RGCs were only weakly
218 affected upon blocking glycine receptors in combination with blocking $GABA_{A,C}$
219 receptors. Therefore, inhibitory synaptic interactions are not the major mechanism
220 behind the local components of suppression.

221 Next, we tested whether these local components originated from within the receptive
222 field center. For this, we modified the spatial layout of our paradigm to exclude
223 saccades from the very center of the receptive field. Simply reducing the size of our
224 mask would have severely decreased the number of simultaneously recorded cells
225 located inside the mask, and we therefore resorted to a different strategy: saccades
226 and flashes were presented in small square regions spread across the retina,
227 separated by gaps kept at mean luminance ("checkerboard mask", Fig. 2e). In one
228 condition (Fig. 2e, left), we presented saccades and flashes in all regions of the
229 checkerboard mask; in the other condition (Fig. 2e, right), saccades and flashes were
230 presented in alternate regions. With this arrangement, saccades were excluded from
231 at most ~300 x 300 μm^2 of a cell's receptive field center, even if that cell was perfectly

232 centered on a non-saccade region. Flashes were presented in the set of regions that
233 included the square region covering the receptive field center of the analyzed RGC
234 (Fig. S7a).

235 Probe flash responses following saccades were suppressed in both ON and OFF
236 RGCs when the saccade and flash were presented in all regions (Fig. 2f, thick lines;
237 Fig. S7c; see Fig. S7b for example cells), consistent with the suppression observed
238 after full-field saccades (Figs. 1, S3). When saccades were excluded from the
239 receptive field center, and were presented in alternate regions to the flash, the flash
240 responses were no longer suppressed in OFF RGCs (Fig. 2f, bottom, thin line), even
241 though these cells showed spiking responses to saccades themselves (Fig. S7b). In
242 fact, flash responses were even enhanced. This suggests that the local component of
243 suppression in OFF RGCs arises fully from within the receptive field center ("central
244 component"). This highly localized origin of suppression in OFF RGCs was further
245 confirmed by additional analysis of the large mask experiments (see Figs. S7d, e). In
246 ON RGCs, on the other hand, suppression persisted (Fig. 2f, top; Fig. S7c), even
247 though a loss in suppression was apparent for flashes presented immediately after the
248 saccade, at 117 and 150 ms (marked with an arrow in Fig. 2f). This suggests that in
249 ON RGCs, part of the early suppression originates from the "central component". The
250 leftover suppression during these early time points might be explained by the global
251 component of suppression, described above (Fig. 2b, S5c), which should also be
252 triggered under this experimental setting. However, since the global component also
253 recovers quickly (by 350 ms, Fig. 2b, S5c), the persisting suppression at the later time
254 points (350 ms and beyond) needs to originate from yet another source beyond the
255 receptive field center. We call this the "surround component", and it may originate from
256 the ON RGCs' immediate surround, which also experiences the saccade under this
257 experimental setting. Therefore, in ON RGCs, the local component of suppression can
258 be divided into a central and a surround spatial component.

259 ***Summary of retinal saccadic suppression spatial origins***

260 In summary, our data suggests that retinal saccadic suppression is mediated by at
261 least three components with distinct spatial origins and temporal properties (Fig. 3): a
262 central, surround, and global component. Suppression in OFF RGCs is mediated
263 exclusively by the central component, which originates from the cell's receptive field

264 center and is characterized by fast onset and fast recovery (by 350 ms after saccade
265 onset). In ON RGCs, we most directly observed the global component (Figs. 2b, S5a).
266 It extends into the periphery and its timing is similar to the central component in OFF
267 RGCs. Only this global component is affected by blocking GABA receptors (Figs. 2b,
268 S5b). During full-field saccades, removing this component by blocking GABA
269 receptors has little effect on the overall suppression (Figs. 2c, 2d, S7), suggesting a
270 more dominant role of the remaining components. The central component in ON RGCs
271 can only be observed by the loss in suppression for early flashes when saccades are
272 excluded from the receptive field center (marked with an arrow in Fig. 2f, top). Its full
273 duration and time course are obscured by the concurrently acting global and surround
274 components. However, given the identical pharmacological dependencies and spatial
275 origins, it is plausible that the central component is symmetric across ON and OFF
276 RGCs with a common underlying mechanism. Therefore, the longer suppression in
277 ON RGCs can neither be attributed to the central nor global components. It likely
278 originates from the immediate surround of the receptive field. This surround
279 component is long lasting (recovers by ~1s) and has a slow onset (Figs. 2f, S3).

280 **Suppression is triggered by interaction between consecutive stimuli of opposite**
281 **polarity**

282 We previously showed (Idrees et al., 2020) that retinal and perceptual saccadic
283 suppression not only occurs after texture displacements, but also after instantaneous
284 texture jumps and structure-free uniform luminance steps. These observations
285 suggested that saccadic suppression is the consequence of rather general
286 mechanisms in which the response to a second stimulus (here: probe flash) gets
287 suppressed by a previous visual transient (caused by saccades or luminance steps).
288 In the following, we apply additional analysis to the luminance-step paradigm
289 (Fig. S1b) dataset of (Idrees et al., 2020), to investigate how the polarity and strength
290 of the visual transients affect the suppression of ON and OFF RGCs.

291 Similar to all previous experiments, we analyzed the modulation index of ON and OFF
292 RGCs separately, using bright probe flashes to analyze ON RGCs and dark probe
293 flashes for OFF RGCs. Consistent with the suppression after texture displacements
294 (Figs. 1e, S3), responses to flashes after luminance steps were strongly suppressed

295 in both ON and OFF RGCs, and ON RGC suppression outlasted suppression in OFF
296 RGCs (Fig. 4a, S8a). The two seemingly different experimental paradigms may
297 therefore trigger similar mechanisms in the retina.

298 We hypothesized that the response to a luminance step might strongly activate RGCs,
299 so that the response to a subsequent probe flash would drive the cells into adaptation
300 or saturation, effectively resulting in suppressed flash responses. At least for the
301 central component of suppression, this could be a viable mechanism as suppression
302 is not caused by inhibitory synaptic interactions. If this was indeed the case, then
303 positive-contrast luminance steps would suppress responses to bright flashes in ON
304 RGCs, and negative-contrast luminance steps would suppress responses to dark
305 flashes in OFF RGCs. To test this, we separately analyzed the effects of positive- and
306 negative-contrast luminance steps on probe flash responses (Fig. 4b). Surprisingly,
307 the resulting effects were contrary to our adaptation/saturation hypothesis: the
308 responses of ON RGCs to bright probe flashes were only weakly suppressed after
309 positive-contrast luminance steps (Fig. 4b, left), but strongly suppressed following
310 negative-contrast luminance steps (Fig. 4b, right). Similarly, responses of OFF RGCs
311 to dark probe flashes were weakly suppressed by negative-contrast luminance steps
312 (Fig. 4b, right), but strongly suppressed by positive-contrast luminance steps (Fig. 4b,
313 left). Fig. S8 shows the underlying population data for these experiments, and
314 Figs. S9a, b show the spiking response of a representative ON and OFF RGC,
315 respectively. While ON RGCs did show a small component of suppression in support
316 of our adaptation/saturation hypothesis (Fig. 4b left panel, see Fig. S9a left column
317 and Fig. S10a for a detailed analysis), the dominant suppressive effect in both ON and
318 OFF RGCs was caused by luminance steps with the opposite contrast as the
319 subsequent flash.

320 Such crossover style of suppression would be consistent with mechanisms involving
321 crossover inhibition via amacrine cells (Werblin, 2010), where activation of OFF
322 pathways (here: by the negative-contrast luminance step) would inhibit responses in
323 the ON pathway (here: to the bright probe flash) and vice versa. However, consistent
324 with our earlier experiments (Fig. 2c, d), suppression in ON RGCs still persisted upon
325 blocking $GABA_{A,C}$ and glycine receptors (5 μM SR-95531, 100 μM Picrotoxin and 1 μM
326 Strychnine) (Fig. 4c). We could not calculate a modulation index for OFF RGCs under

these conditions because they did not respond to brief probe flashes in the presence of the pharmacological agents, and therefore the modulation index was mathematically undefined. However, in our texture displacement experiments, the same pharmacological agents (Fig. 2d) had no substantial effect on OFF RGC suppression. The crossover-style suppression observed in Fig. 4b was therefore unlikely to be caused by classical crossover inhibition pathways involving amacrine cells and $GABA_{A,C}$ or glycine receptors.

Central component of suppression results from cone response kinetics and nonlinearities in downstream retinal pathways

So far our experiments suggest that suppression in OFF RGCs (1) is mediated solely by the central component of suppression that originates in the receptive field center (Figs. 2e, S7b-e), (2) is predominantly triggered by the interaction between consecutive stimuli with opposite polarity (Figs. 4b, S8b), and (3) is not caused by inhibitory amacrine cells (Figs. 2c, 2d, S6). Similar conclusions can be drawn for the central component of suppression in ON RGCs. Together, these results restrict the possible cellular substrates for this central suppressive mechanism to the feed-forward pathway in the retina, namely photoreceptors - bipolar cells - RGCs. We wondered whether opposite-polarity stimulus-stimulus interactions could already modulate the responses of photoreceptors themselves. For this, we recorded the output of cones with an intensity-based glutamate-sensitive fluorescent reporter (iGluSnFR) (Marvin et al., 2013; Szatko et al., 2020), predominantly expressed in horizontal cells postsynaptic to cone terminals (Methods). We presented a shortened version of the luminance step paradigm in which a homogeneous background alternated between a brighter and darker gray value (Fig. S1c) to induce positive-contrast (+0.4 Michelson contrast) and negative-contrast (-0.4 Michelson contrast) luminance steps. Dark or bright probe flashes (100 ms duration, -0.33 or +0.33 Michelson contrast, respectively) followed the luminance steps at different delays (50, 250, and 2000 ms), with the flash at 2000 ms serving as baseline.

The luminance steps caused sustained changes in the cones' glutamate output (Fig. 5a). The transient responses to the probe flashes were superimposed on these glutamate modulations (Fig. 5b). This superposition was mostly linear and did not

358 indicate nonlinear effects such as adaptation or saturation. Therefore, when we
359 isolated the flash responses by subtracting the step responses, the resulting peak
360 flash responses were only weakly affected by the preceding luminance step (Fig. 5c).
361 Thus, at the level of the cone output (Fig. 5b), there was hardly any suppression when
362 only considering the peak of the probe flash responses (Fig. 5c). How does the
363 suppression observed at the level of RGC output arise from effectively linear cone
364 responses? The answer lies in other properties of the cone response, such as its
365 kinetics, which will be captured by downstream retinal pathways.

366 To demonstrate this, we used a previously published phenomenological model of
367 retinal processing. In the model (Drinnenberg et al., 2018), RGC responses are
368 described by passing a “light stimulus” to model photoreceptors, feeding their output
369 through a set of linear-nonlinear filters to describe the processing in different bipolar
370 pathways (Methods), and eventually converting these filter outputs into RGC spiking.
371 Here, as the first step of analysis, we fitted the parameters of the model cone to reflect
372 our measured data of cone output. The model faithfully explained the observed cone
373 responses (Fig. S11) and gave us the opportunity to calculate cone responses to
374 flashes presented at additional time points. This modeled cone output to step-flash
375 combinations was fed into the model bipolar cells, finally yielding model RGC
376 responses (Fig. 6). In the model, different RGC types can be described by varying the
377 bipolar cell filter properties. We first investigated transient RGC responses (Fig 6a-c)
378 and calculated a modulation index (Fig. 6d) comparable with the modulation index of
379 our real RGC data. As a control, we also passed the raw cone output data, instead of
380 the fitted cone model, to the bipolar cells and found qualitatively the same results
381 (Fig. S12).

382 The model’s ON and OFF RGCs (Fig. 6d) showed crossover-style suppression that
383 was consistent with the suppression of real RGCs under similar luminance step
384 experiments (Fig. 4b): the model’s ON RGC showed strong suppression to the bright
385 flashes presented immediately after the negative-contrast luminance step (Fig. 6d,
386 right), while flashes after the positive-contrast luminance step were only weakly
387 affected (Fig. 6d, left). Suppression recovered by 200 ms, consistent with the recovery
388 time for the central component of suppression in real RGCs. Similarly, the model’s
389 OFF RGC showed strong suppression of its response to dark flashes presented

390 immediately after the positive-contrast luminance step (Fig. 6d, left); suppression was
391 absent in OFF RGC when the dark flash was presented later or after a
392 negative-contrast luminance step (Fig. 6d). The markers on the curves in Fig. 6d
393 correspond to the time points when the flashes were presented to the cones in the
394 experiments depicted in Fig. 5. Model RGC responses to step-flash combinations and
395 flash-induced responses at these time points are shown in Fig. 6a-c.

396 What mechanism within the model led to the emergence of the suppressive effect in
397 RGC responses, despite the mostly linear response superposition at the cone
398 output? In the model, the bipolar cells have transient filter properties and are driven
399 predominantly by the instantaneous rate of change in the cone output, i.e. its
400 derivative, rather than by the absolute cone output. The response to a probe flash
401 presented immediately after an opposite-contrast luminance step (50 ms) occurred
402 during the initial phase (ramp) of the cone response to the luminance step (Fig. 5b,
403 columns 2 and 3). This causes a much smaller rate of change in cone output and
404 therefore drives the downstream bipolar cells only weakly, resulting in weak or even
405 completely suppressed model RGC responses (Fig. 6c, columns 2 and 3). On the
406 other hand, flashes presented during the steady state phase of the luminance step
407 response (250 ms and 2000 ms in Fig. 5b, columns 2 and 3), or flashes presented
408 immediately after (50 ms) a same-contrast luminance step (Fig. 5b, columns 1 and 4)
409 caused larger instantaneous changes in the cone output, and therefore resulted in
410 relatively stronger spiking (250 ms and 2000 ms in Fig. 6c; 50 ms in Fig. 6c, columns
411 1 and 4).

412 If suppression observed at RGC output did indeed result from the weaker rate of
413 change in the cone output, less transient RGCs should show weaker suppression
414 because they reflect the absolute cone output signal more faithfully. To test this, we
415 varied the linear-nonlinear filter of the model bipolar cell to make the retinal pathway
416 less transient (Methods). Indeed, the resulting crossover-style suppression was
417 weaker in model RGCs with less transient response properties, indicated by the
418 negative modulation index slope in Fig. 6e: see ON RGC suppression after
419 negative-contrast steps (blue lines in row 2) and OFF RGC suppression after
420 positive-contrast steps (red lines in row 1); columns represent different time points
421 sampled. Curious if this effect was also present in our real RGC data, we re-analyzed

422 the data of Fig. 4 and evaluated suppression as a function of RGC transiency
423 (Methods). The result was congruent with the model prediction: suppression was
424 weaker in less transient (more sustained) cells (Fig. 6f).

425 The analyses above suggest that the central component of retinal saccadic
426 suppression does not have a single site of origin. Instead, it appears to emerge from
427 the relatively slow kinetics of cone responses (such that flash responses ride on the
428 initial rising/falling phase of the cone's step response), combined with temporal
429 properties and nonlinearities of downstream retinal pathways. The downstream
430 pathways shape the exact strength and time course of RGC suppression. We term
431 this novel retinal processing motif "dynamic reversal suppression": "dynamic" because
432 of the necessarily tight temporal link between responses to the two consecutive stimuli;
433 and "reversal" because the response to an event gets suppressed when it reverses
434 the cone output triggered by the first stimulus.

435 **Generalization to other species**

436 Retinal saccadic suppression, at least its central component, was triggered by
437 stimulus-stimulus interactions (Figs. 4, 6, S8), governed by general retinal signal
438 processing, without the need for any specialized "saccadic suppression" circuit. It is
439 likely that such general processing is conserved across species. Indeed, we observed
440 quantitatively similar retinal saccadic suppression in pig ON and OFF RGCs
441 (Fig. S13), including the dependency on background texture statistics. Interestingly,
442 like in mouse, OFF RGCs in pig retina also recovered faster than ON RGCs,
443 suggesting that the surround component of suppression was also present in pig ON
444 RGCs. In an additional experiment, we also recorded the activity of RGCs from ex vivo
445 macaque retina while subjecting it to a shorter version of the luminance step paradigm
446 (Fig. S1b) with fewer conditions than in the original paradigm. Our results (Fig. 7)
447 provide an indication that macaque RGC responses to probe flashes, following
448 luminance steps, are suppressed in a way similar to mouse retina. However, more
449 data will be required to determine the population trend and for characterizing the
450 dictionary of response modulations in macaque retina.

Downstream visual areas may modulate retinal saccadic suppression

Given the similarities we previously described between retinal and perceptual saccadic suppression (Idrees et al., 2020), it was tempting to test whether the crossover style of suppression, observed in the retina (Fig. 4b), was also reflected in perception. We therefore conducted human psychophysics experiments where we asked human subjects (N = 5) to maintain saccade-free fixation, while we simply changed the luminance of the homogenous background to a brighter (0.3 to 0.56 Michelson contrast) or darker (-0.3 to -0.56 Michelson contrast) background (Fig. 8a; Methods). At random times relative to the luminance step, we presented a dark (-0.033 Michelson contrast) or bright probe flash (+0.033 Michelson contrast), at one of four locations in the subjects' field of view. At trial end, the subjects were asked to localize the probe flash.

Irrespective of the step → flash combination, subjects were strongly impaired in their ability to localize the probe flashes presented around the time of the luminance step (Fig. 8b). Most interesting in this context was the combination of negative-contrast luminance steps with dark probe flashes. In mouse retina, even though OFF RGCs did show weak suppression to this combination (Fig. S8b, inverted histograms in row 2), this effect was virtually absent at the population level (Fig. 4b, right panel). In human perception, however, this combination led to strong suppression (Fig. 8b, right panel). We cannot exclude that stronger retinal suppression to this specific combination might be present under different light or stimulus conditions. Another possibility is that it might be more pronounced in the retina of humans and other primate species (Fig. 7). Nonetheless, visual mechanisms of suppression certainly exist in higher visual brain areas (Idrees et al., 2020). Perceptual suppression after same-contrast stimulus combinations may arise from processing in these higher visual brain areas, which may modulate and complement retinal saccadic suppression to achieve robust effects at the perceptual level.

Discussion

For most RGCs that we recorded, responses to brief probe flashes were strongly suppressed when presented after saccade-like texture displacements across the retina. In fact, similar suppression occurred when texture displacements were replaced

482 by sudden uniform changes in background luminance, suggesting that suppression
483 was caused by rather generic mechanisms, triggered by visual transients across the
484 retina, rather than specialized suppression circuits that react to saccadic image
485 motion. We found that the suppression strength depended on four main factors:
486 (1) strength of the visual transients, governed by the statistics of the background
487 texture or the contrast of the luminance step; (2) elapsed time following the visual
488 transient; (3) RGC polarity (ON vs. OFF RGCs); and (4) RGC response properties
489 (RGCs transiency). Stronger visual changes, elicited either by coarser textures or
490 larger luminance step contrasts, caused stronger suppression, peaking approximately
491 50 ms after the stimulus offset (Figs. 1e, 4). The recovery times depended on RGC
492 polarity: OFF RGCs recovered by 250 - 350 ms whereas suppression in ON RGCs
493 lasted for up to 1 s (Figs. 1e, 4). The suppression was stronger in more transient RGCs
494 (Fig. 6e, f). We identified at least three components of retinal saccadic suppression,
495 with distinct spatial origins, defined as central, surround, and global components
496 (Fig. 3). These components were mediated by different underlying mechanisms.

497 The central component was the only source of suppression in OFF RGCs, and the
498 dominant source in ON RGCs for time points immediately after a full-field saccade or
499 luminance step. This component was short-lived (~250-350 ms) and originated from a
500 cell's receptive field center. It was triggered by opposite-polarity stimulus-stimulus
501 interactions, which naturally occur during saccades and other forms of sequential
502 visual stimulation. The relatively slow temporal dynamics of the cone photoreceptor
503 responses to two opposite-polarity consecutive stimuli were the basis for subsequent
504 non-linear processing, ultimately suppressing RGC response to the second stimulus.
505 Therefore, the central component of suppression cannot be pinpointed to one specific
506 site but seems to be an emergent property. We call this novel processing motif
507 "dynamic reversal suppression". Such a mechanism, where the cone response itself
508 remains linear, but nonetheless forms the basis for subsequent nonlinear response
509 modulation, is clearly different from adaptation (Clark et al., 2013) or desensitization
510 (Baylor and Hodgkin, 1974) mechanisms within the cones, which would evoke
511 nonlinear responses of the cones themselves.

512 The suppressive mechanism of the central component also suggests that perceptual
513 saccadic suppression is derived, at least in part, from the inherent response properties

514 of photoreceptors, the very first cell in the visual processing cascade. In fact, this early
515 implementation could also explain why we observed suppression in most RGCs we
516 recorded from (Figs. S3, S8, S13), covering a wide spectrum of response properties
517 and therefore presumably many RGC cell types (see Supplementary Fig. 6 of Idrees
518 et al., 2020). Still, our results suggest that the suppression initiated at the level of cone
519 photoreceptors is translated differently by the different parallel pathways in the retina,
520 leading to variability in response suppression at the ganglion cell level (such as the
521 stronger suppression in more transient RGCs, Fig. 6e, f). Further cell type
522 classification will be required to relate the degree of modulation with pathway
523 specificity. The type and degree of modulation might also differ across species, even
524 though we see qualitatively similar suppression in mouse, pig, and macaque RGCs.

525 ON RGCs, in addition to this central component, derive suppression from two more
526 components. First, the global component is a fast but short-lived (~250-350 ms)
527 component, caused by inhibition via GABAergic wide field amacrine cells, triggered by
528 global image changes and carried to the RGC from as far as the cell's periphery
529 (Figs. 2b, S5). This likely belongs to the same class of circuits as described by others
530 in the context of peripheral stimulation (Baccus et al., 2008; Brooks and Holden, 1973;
531 Enroth-Cugell and Jakiela, 1980; Roska and Werblin, 2003; Zhang et al., 2012). The
532 role of this global component in perceptual saccadic suppression might not be so
533 prominent as it is masked by more local components during full-field saccades
534 (Fig. 2c, 2d, S6). It is nevertheless another mechanism to process global motion, in
535 addition to the several others (Wei, 2018). It may, for example, play a role in perceptual
536 modulations during pursuit eye movements (Schütz et al., 2011). Here, the central
537 component of suppression will not be triggered in RGCs whose receptive field centers
538 are locked to the tracked object; but these RGCs will still be suppressed by the global
539 component.

540 The second additional component is the surround component. It seems to act with a
541 delay of ~200 ms and can last for up to 1000 ms. The spatial origins of this component
542 are not very well understood, but our data indicates that it presumably arises from the
543 immediate surround of a cell's receptive field. Additionally, similar to the central
544 component, it does not rely on $GABA_{A,C}$ or glycine receptors (Fig. 2c, 2d, S6), but the
545 exact mechanisms remain to be explored. Possible mechanisms could involve

546 negative feedback of horizontal cells onto the cones (Drinnenberg et al., 2018;
547 Kemmler et al., 2014). This slower component most likely contributes to the longer
548 recovery times observed in ON cells. Interestingly, visual masking in cat LGN also
549 lasts longer in ON versus OFF cells (Schiller, 1968), which may be a consequence of
550 the effects we describe here in the retina. While this surround component plays an
551 important role in shaping RGC and downstream neural responses following a visual
552 transient, its contribution to perceptual saccadic suppression can also be disputed.
553 This is because, during real saccades, eye-movement related signals (e.g., corollary
554 discharge) shorten the duration of suppression caused by visual mechanisms (Idrees
555 et al., 2020), such that the long-lasting surround component may not critically shape
556 perception. Yet another additional component of suppression, based on saturation-like
557 mechanisms (Fig. 4b left, S10a), was found only in ON RGCs. It is possible that this
558 component originates at the level of bipolar cells as response saturation has been
559 observed predominantly in ON bipolar cells but not in OFF bipolar cells (Schreyer and
560 Gollisch, 2020). In summary, while Fig. 3 summarizes the three spatial components of
561 suppression and their temporal properties, these components in turn can have further
562 sub-components. An example is the saturation-type of suppression in ON RGCs which
563 acts in addition to the dynamic contrast reversal. Both can be assigned to the central
564 component.

565 Irrespective of which components of retinal saccadic suppression contribute towards
566 perceptual saccadic suppression, our results show that retinal responses to stimuli
567 following visual transients are modulated concurrently by several mechanisms (Fig. 3).
568 Additional mechanisms might exist under different stimulus conditions. From the
569 perspective of retinal visual feature processing, this would be consistent with how
570 multiple mechanisms concurrently process other visual features in the retina, such as
571 motion (Wei, 2018). Consequently, our study is not only relevant from the perspective
572 of saccadic suppression, but also contributes new insights into retinal signal
573 processing during the dynamics of natural vision where local and global visual
574 transients are frequent. Our visual system might exploit this retinal signal processing
575 to enhance visual performance during high-acuity tasks. For example, while threading
576 a needle, microsaccades will jitter the needle and its background across the retina.
577 RGCs exposed to the background will experience stronger visual transients than

578 RGCs exposed to the needle. As a result, RGCs ‘seeing’ the background might show
579 suppressed responses following microsaccades, perceptually blurring out the
580 background. This relative enhancement of the fixated object by background
581 suppression could be valuable for visual tasks such as object recognition.

582 Suppression during different paradigms of sequential visual stimulation has regularly
583 been observed in downstream visual areas, e.g. when studying saccadic suppression
584 (Idrees et al., 2020), visual masking (Breitmeyer, 2007) or neural adaptation (Mayo
585 and Sommer, 2008). The similarities between the observations there and the
586 suppression we observe in the retina suggest that the retina may be a common neural
587 substrate of these seemingly different types of suppression, unifying their underlying
588 mechanisms. Indeed, it is remarkable that an elementary property of retinal saccadic
589 suppression, i.e. its dependence on the scene statistics (Fig. 1e), is clearly preserved
590 all the way to perception (Idrees et al., 2020). On the other hand, retinal saccadic
591 suppression alone cannot explain all perceptual suppressive effects. For example,
592 suppression before a saccade (or a luminance step) is observed perceptually (even in
593 the absence of eye movements, Fig. 8), but not in the retinal output (Idrees et al.,
594 2020), and therefore most likely originates from visual processing beyond the retina.
595 Similarly, not all properties of retinal saccadic suppression were preserved in
596 perception. The crossover style of suppression observed in the retinal output of mice
597 (Fig. 4b) was not observed perceptually in humans (Fig. 8). The more far-reaching
598 perceptual suppression may be the result of additional visual processing beyond the
599 retina (Shapley and Enroth-Cugell, 1984). Alternatively, our results, despite their
600 robustness, may not capture the full array of retinal processing. For one, our stimulus
601 conditions in mouse retina experiments may not have been adequate to capture all
602 retinal suppressive effects. Further, the retinal output of humans and other primate
603 species (Fig. 7) might differ from mouse retina in this respect.

604 Retinal saccadic suppression, or rather sequential-stimuli suppression, is only one
605 way the retina alters its output during dynamic vision. Other forms certainly co-exist,
606 such as brief changes in RGC polarity following peripheral shifts (Geffen et al., 2007),
607 or sensitization of some RGC types following a change in background luminance
608 (Appleby and Manookin, 2019). These and several other studies (Amthor et al., 2005;
609 Barlow et al., 1977; Enroth-Cugell and Jakiela, 1980; Krishnamoorthy et al., 2017;

Noda and Adey, 1974; Passaglia et al., 2009; Roska and Werblin, 2003), together with ours, demonstrate the complex image processing capabilities in the retina to facilitate downstream visual processing for the ultimate service of perception during natural vision. Looking forward, the detailed characterization of retinal output provided here paves the way to investigate the interactions between retinal and extra-retinal (visual and non-visual) mechanisms of saccadic suppression, to further our understanding of how the visual system maintains stability in the face of constant disruptions.

Acknowledgements

We thank Elisabeth Gustafsson for providing technical support; Andreas Hierlemann for providing the HiDens CMOS MEA system and helping establish our high-density MEA recordings; Roland Diggelmann for helping in setting up the pipeline (including providing code) for automatic spike sorting of high-density MEA recordings; Martin Schenk at the Department of Experimental Surgery, Tübingen, for providing us domestic pig eyes. Our 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. (MU3792/1-1 and MU3792/3-1). T.A.M. received support from the Tistou and Charlotte Kerstan Foundation. S.I., T.A.M. and Z.M.H. were also supported by an intra-mural funding program (Project 2013-05) of the Werner Reichardt Center for Integrative Neuroscience. F.F. was supported by a Swiss National Science Foundation Ambizione grant (PZ00P3_167989) and Swiss National Science Foundation Eccellenza grant (PCEFP3_187001). M.P.B. and Z.M.H. were further funded by the SFB 1233 on “Robust Vision” (DFG) - project number 276693517. K.F. was supported by the Bundesministerium für Bildung und Forschung (BMBF, 01GQ1002) and the Max Planck Gesellschaft (MPG, M.FE.A.KYBE0004).

Author contributions

TAM conceptualized the study; SI and TAM designed the overall study; FF developed and supervised computational modeling; SI, MPB, AK, KF, ZMH, FF and TAM designed experiments; SI performed ex vivo mouse and pig retina experiments; AK performed ex vivo macaque retina experiment; MPB performed human psychophysics experiments; KF performed viral injections; MK performed cone imaging experiments;

SI and FF implemented computational model; All authors analyzed and interpreted the data and wrote the manuscript.

Declaration of interests

The authors declare no competing interests.

Figures

Figure 1

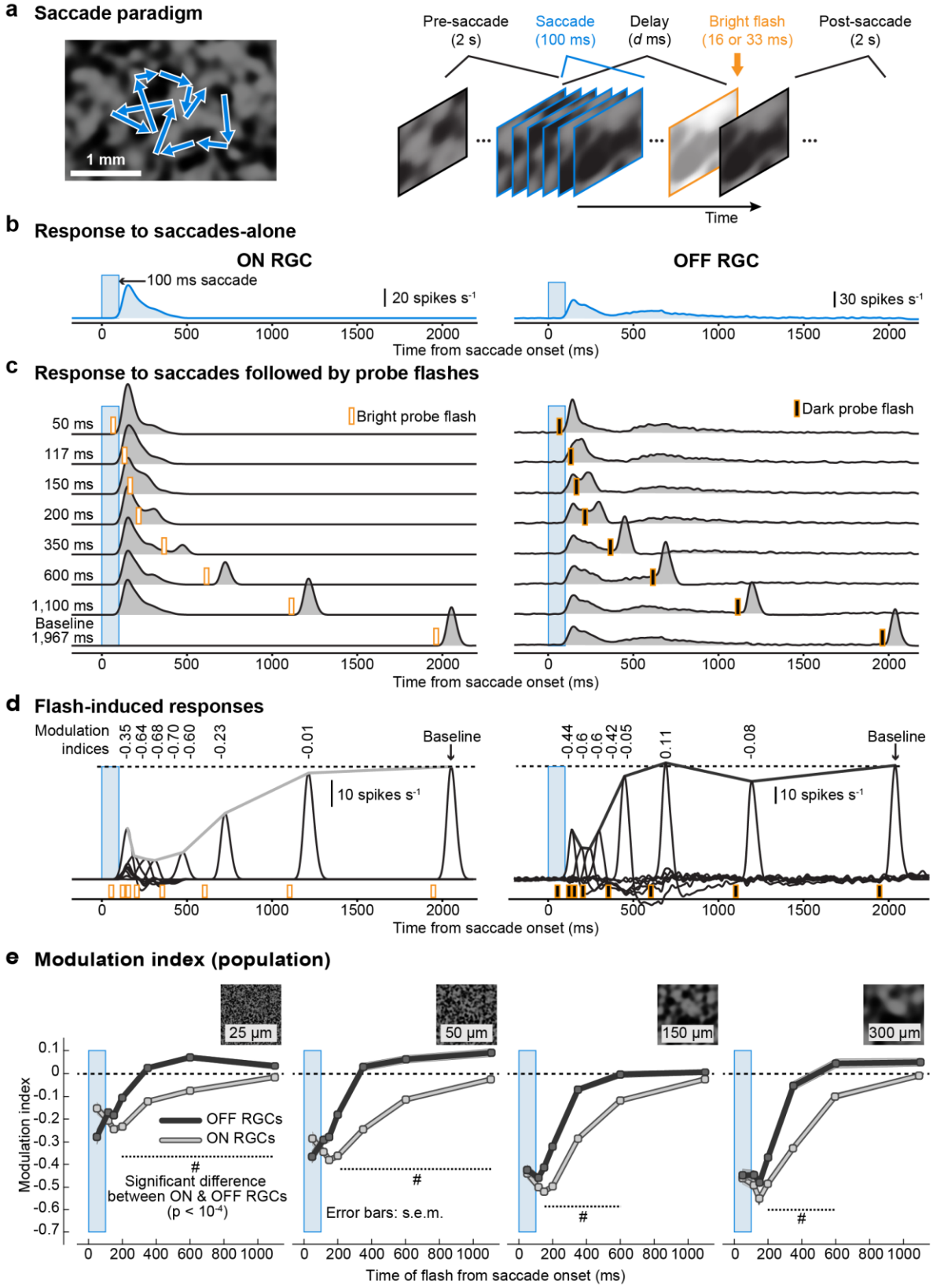


Figure 1 Similarities and differences in retinal saccadic suppression across ON and OFF RGCs.

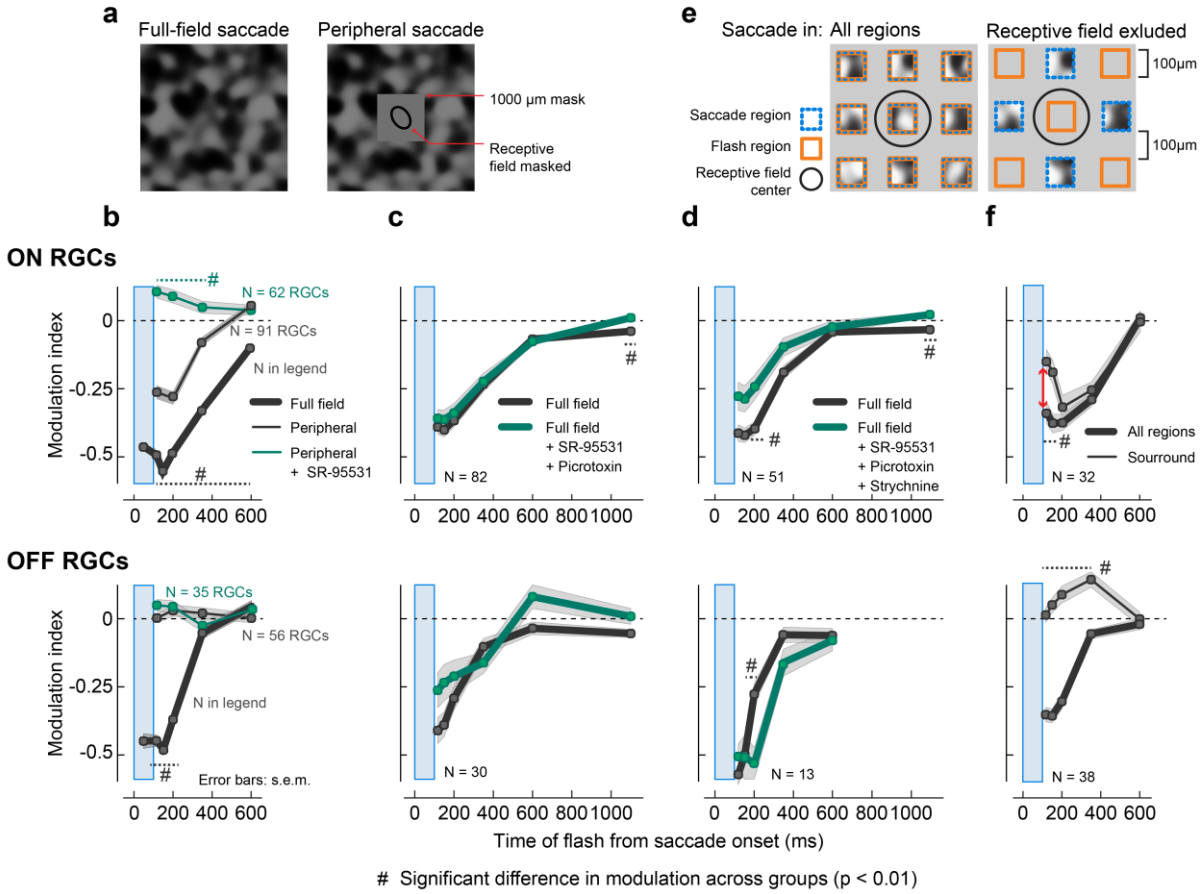
a. RGC action potentials were recorded from ex vivo retinae placed on multielectrode arrays. Saccades were mimicked by displacing a texture projected onto the retina (blue arrows in left panel indicate texture displacement paths). The texture remained static for 2 s and was then displaced over 100 ms (blue outlines) followed by a brief probe flash (here bright probe flash is depicted, orange outline). Each trial consisted of 39 such successive saccade-flash sequences (Fig. S1a).

b,c. Average activity (firing rate) of an example ON RGC (left column) and OFF RGC (right column) to 39 saccade sequences not followed by a probe flash (**b**), and to 39 saccade sequences followed by probe flashes at different delays after saccade onset (**c**). Blue window: timing of saccades; orange markers: timing of probe flashes.

d. Isolated flash-induced responses (firing rate) of the same RGCs obtained by subtracting responses to saccades-alone (**b**) from responses to saccades followed by probe flashes (**c**). Lines connecting the response peaks highlight the time courses of retinal saccadic suppression relative to baseline flash-induced responses. Numbers above each response peak represent the modulation index which quantifies how much the probe flash response is modulated by the preceding saccade (Methods, negative modulation indices correspond to suppressed flash-induced responses).

e. Population modulation index (mean ± s.e.m.) of ON (light gray) and OFF (dark gray) RGCs, for different background textures with different spatial scales (left to right: fine to coarse). The number of ON and OFF RGCs in the population varied between 68 and 574 for different flash times and textures (see Fig. S3 for exact numbers and relevant statistics). Hash symbols: significant modulation difference between ON and OFF RGCs ($p < 10^{-4}$, two-tailed Wilcoxon rank-sum test).

673 **Figure 2**



674

675 **Figure 2 Spatial origins of retinal saccadic suppression.**

676 **a.** Spatial layout of the visual stimulation paradigm used in experiments to probe the
677 global component of suppression. Saccades were presented either full-field (left; same
678 as in Fig. 1) or in the periphery (right), where a 1000 x1000 μm^2 mask (intensity: mean
679 luminance of texture) covered at least 2-σ of the 2D Gaussian fit to the RGC receptive
680 fields (Fig. S5b).

681 **b.** Population modulation index (mean ± s.e.m.) of ON (top) and OFF (bottom) RGCs
682 for full-field saccades condition (thick gray lines, same as Fig. 1e rightmost panel;
683 N = 68 to 574 RGCs (see Fig. S3 for exact numbers)); periphery saccades condition
684 (thin gray lines; N = 91 ON RGCs, N = 56 OFF RGCs); and periphery saccades
685 condition in the presence of $GABA_A$ receptor blocker (5µM SR-95531; green lines;
686 N = 62 ON RGCs, N = 35 OFF RGCs). Blue window shows the timing of the saccade.
687 In these experiments, we used a coarse background texture (300 µm spatial scale).
688 Timing of probe flashes: 50 and 150 (only for full-field saccade), 117, 200, 350, 600
689 and 2100 ms (baseline) after saccade onset.

690 **c,d.** Population modulation index (mean ± s.e.m.) of ON (top) and OFF (bottom) RGCs
691 for full-field saccades without any pharmacological agents (gray lines; N = 82 ON
692 RGCs, N = 30 OFF RGCs) and with $GABA_{A,C}$ receptor blockers 5 µM SR-95531 +
693 100 µM Picrotoxin (**c**; green lines), and for a subset of RGCs where we additionally

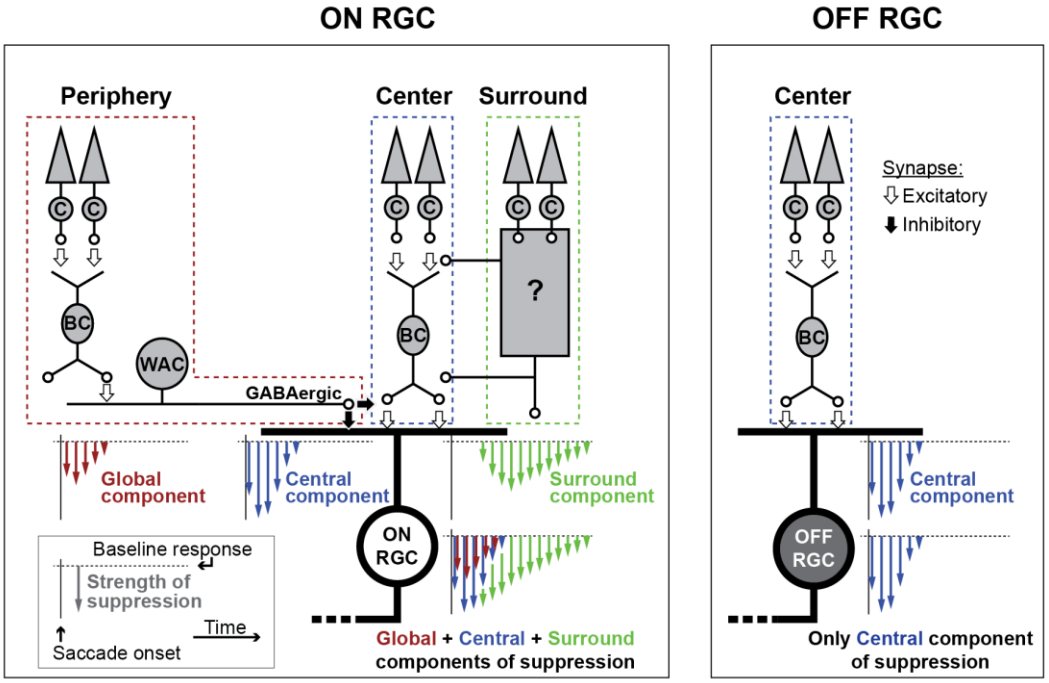
blocked glycine receptors using 1 μM Strychnine (**d**; green lines; N = 51 ON RGCs, N = 13 OFF RGCs). In these experiments, we used a coarse background texture (150 μm spatial scale). Probe flashes were presented at 117 ms, 150, 200, 350, 600, 1100 and 2100 (baseline) after saccade onset.

e. Spatial layout of the visual stimulation paradigm used in experiments to probe local components of suppression. Saccades and flashes were presented in 100 x 100 μm^2 square regions, separated by 100 μm gaps with mean overall luminance. **Left:** Saccades and flashes were presented in all regions. **Right:** Saccades and flashes were presented in alternate regions; only cells with receptive fields (RFs) in the non-saccade regions (orange) were analyzed (black ellipse: 1-σ of the 2D Gaussian fit to an example RGC receptive field). Consequently, saccades were excluded from at most ~300 x 300 μm^2 of a cell's RF center. In these experiments, we used a coarse background texture (150 μm spatial scale).

f. Population modulation index (mean ± s.e.m.) of ON (top; N = 32) and OFF (bottom; N = 38) RGCs for saccades and flashes in all regions (thick lines) or saccades excluded from RGC RF center (thin lines). Red arrow indicates significant loss in suppression in ON RGCs for early flashes at 117 and 150 ms upon excluding saccades from RF center ($p = 0.0016$ and $p = 0.002$ respectively; two-tailed Wilcoxon rank-sum test).

In all panels, hash symbols indicate statistically significant difference between groups ($p < 0.01$, two-tailed Wilcoxon rank-sum test).

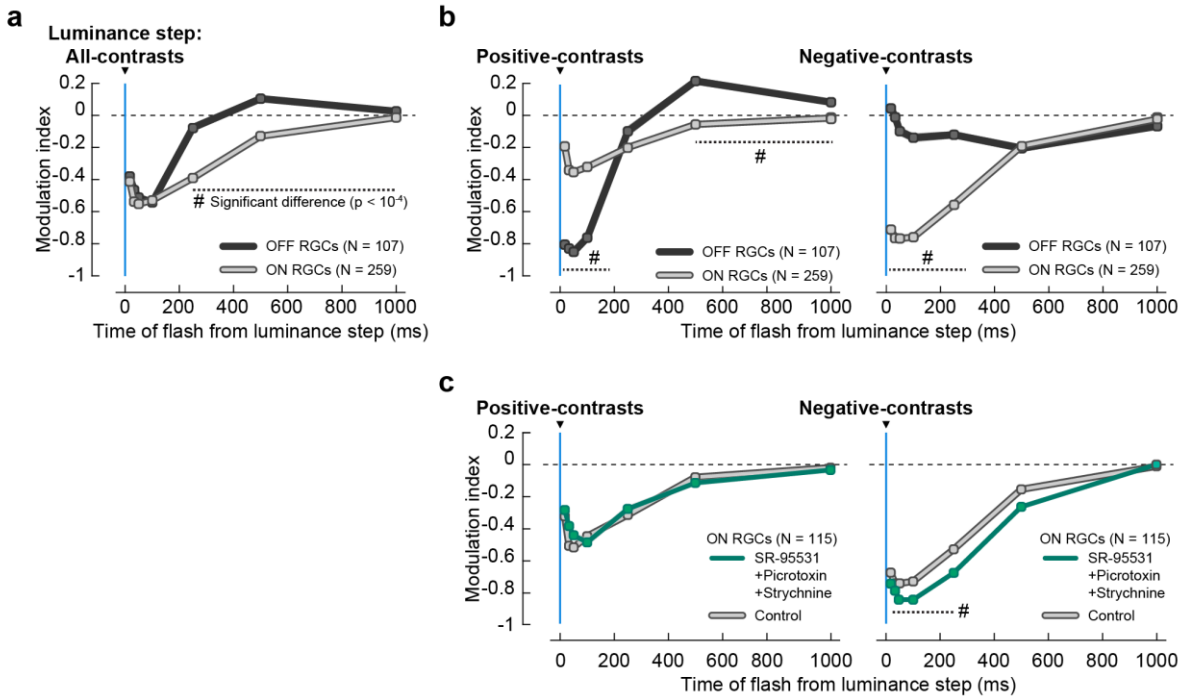
716 **Figure 3**



717

718 **Figure 3 Schematic summarizing the spatial origins of retinal saccadic**
719 **suppression.**
720 Saccadic suppression in OFF RGCs (right) is mediated primarily by the central
721 component of suppression (blue) that originates from the cells' receptive field center.
722 ON RGCs (left) get suppressed from two additional components: First, the fast but
723 short-lived global component (red), mediated by GABAergic inhibition, that originates
724 from as far as the cells' periphery. This global component has a similar temporal profile
725 as the central component. It is however weaker than the central component and acts
726 in parallel to it, indicated by the red arrows parallel to blue arrows in the total
727 suppression schematic. Second, the delayed but long-lasting surround component
728 (green), which might originate from the cell's immediate surround. The central
729 component and surround component do not depend on classical GABAergic or
730 glycinergic inhibitory pathways. The differences in the suppression recovery time in
731 ON and OFF RGCs was mainly due to this surround component acting on ON RGCs.
732 Inset shows the legend for arrow schematics. Length of the arrows represent
733 suppression strength; spread of the arrows show the temporal profile of suppression.
734

735 **Figure 4**



736

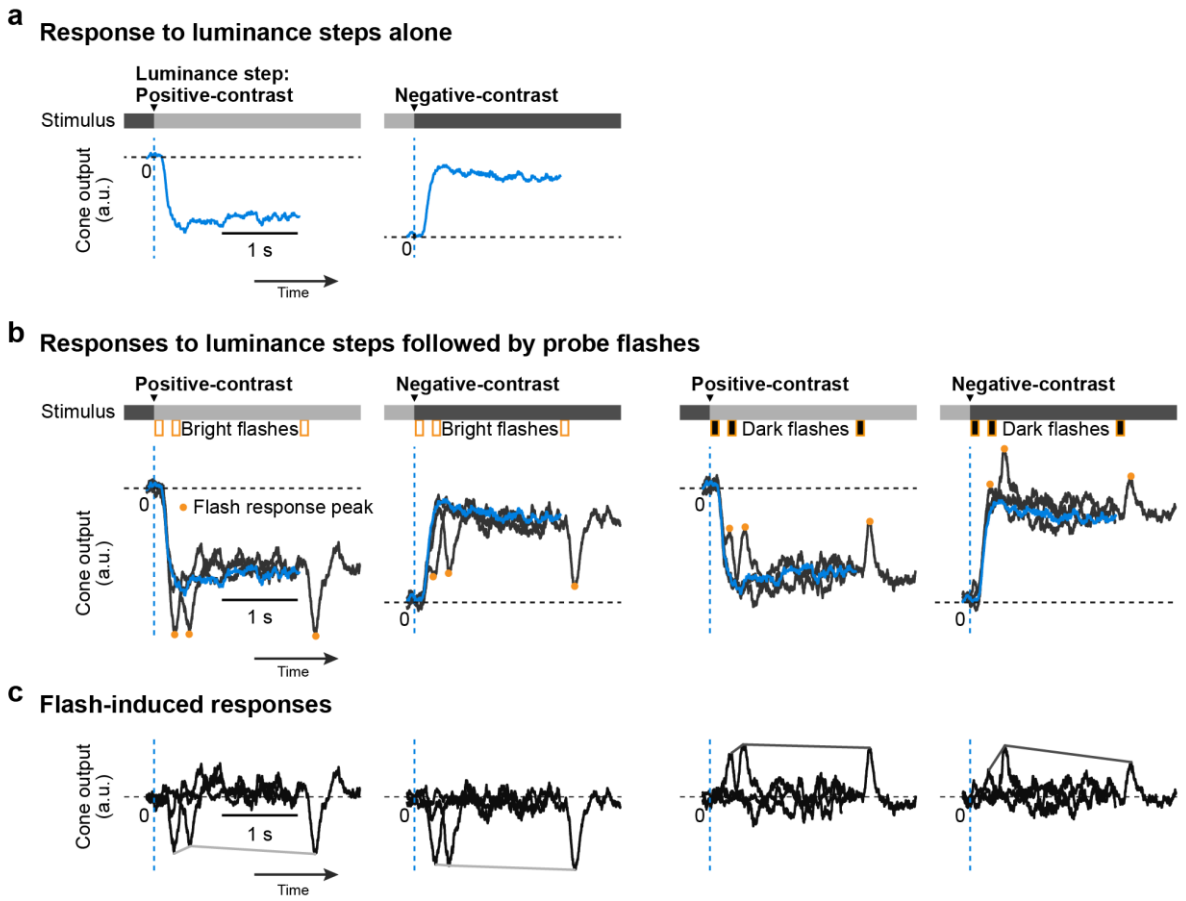
737 **Figure 4 Suppression following luminance steps.**

738 **a.** Population modulation index (mean ± s.e.m.) of ON (light gray, N = 259) and OFF
739 (dark gray, N = 107) RGCs for probe flashes following luminance steps (blue line).
740 Modulation index for each RGC was based on its average response to 56 or 156
741 luminance step sequences (Fig. S1b) spanning a contrast range of -0.5 to +0.5
742 Michelson contrast (Methods). Probe flashes were presented at 17 ms, 33, 50, 100,
743 250, 500, 1000 and 2000 (baseline) after luminance steps. Probe flash responses
744 were suppressed in both ON and OFF RGCs, with similar time course and recovery
745 as in the saccade paradigm with textures (Fig. 1e). Error bars are not visible due to
746 small s.e.m.

747 **b.** Same as in **a**, except that the modulation index for each RGC was separately based
748 on average responses to probe flashes after positive-contrast luminance steps (left
749 panel; +0.03 to +0.5 Michelson contrast), and after negative-contrast luminance steps
750 (right panel; -0.03 to -0.5 Michelson contrast). Underlying population data is shown in
751 Fig. S8.

752 **c.** Same as in **b**, for a subset of ON RGCs (N = 115) in control conditions (light gray
753 lines) and with $GABA_{A,C}$ and glycine receptors blocked (green lines; cocktail of 5μM
754 SR-95531, 100 μM Picrotoxin and 1μM Strychnine). Hash symbol: significant
755 difference between modulation of ON and OFF RGCs in **a** or between ON RGCs
756 without and with pharmacological blockers in **c** ($p < 10^{-4}$, two-tailed Wilcoxon rank-sum
757 test).

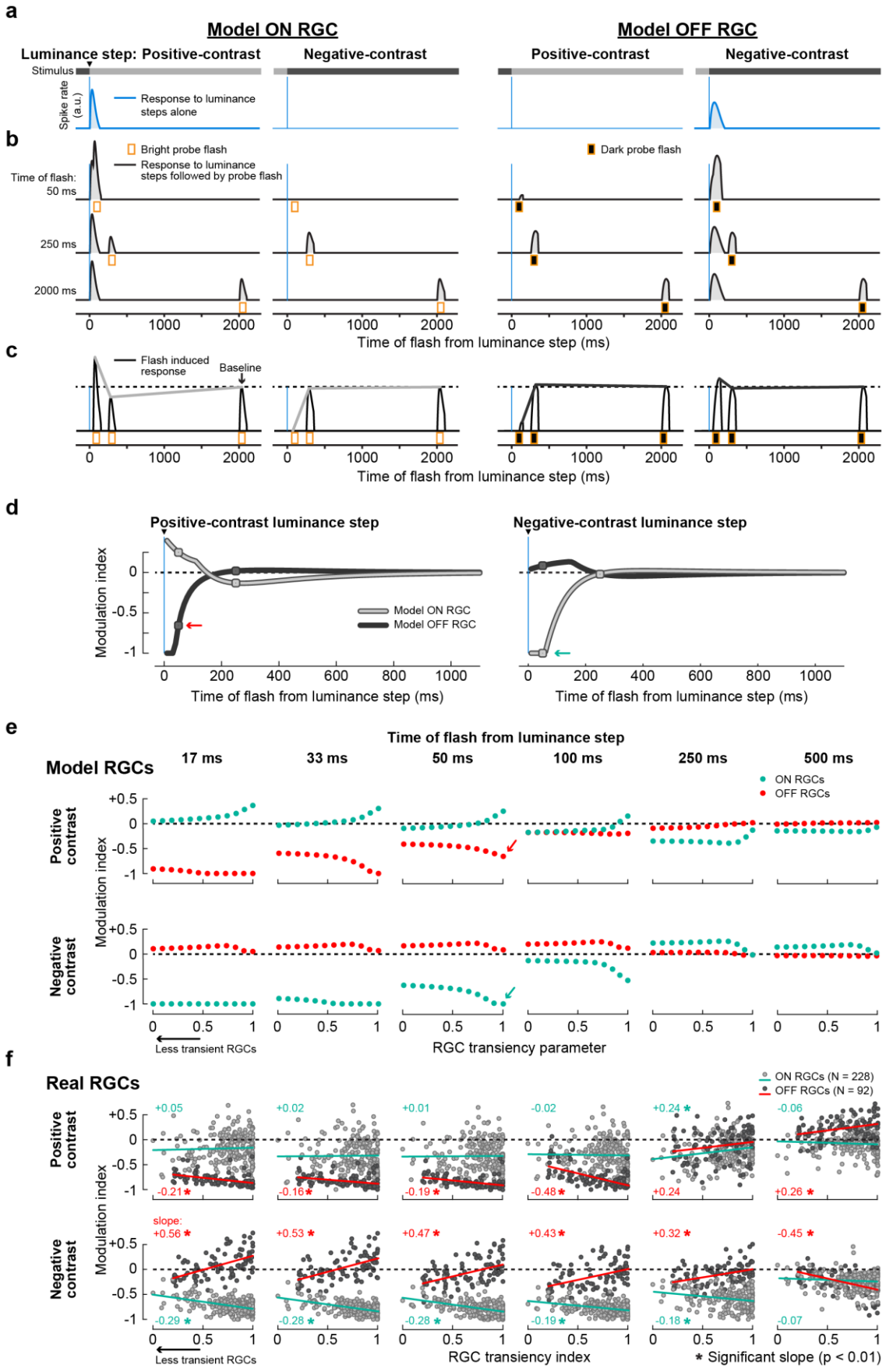
758 **Figure 5**



759

760 **Figure 5 Cone output in response to probe flashes following luminance steps.**
761 **a,b.** Cone responses (baseline normalized iGluSnFR indicator fluorescence signal) to
762 positive- and negative-contrast (+0.4 and -0.4 on Michelson scale) luminance steps
763 alone **(a)** and to luminance steps followed by probe flashes at 17, 250 and 2000 ms
764 (orange bars below the intensity bar shows timing of probe flashes) **(b)**. Probe flashes
765 were either bright or dark (+0.33 or -0.33 Michelson contrast respectively; 100 ms
766 long). In **b**, responses to step-alone (blue) and individual step → flash pairs (dark gray)
767 are overlaid. Dashed blue lines: timing of luminance step; orange circles: peak cone
768 response to flashes; horizontal dashed line: cone level prior to the luminance step.
769 **c.** Flash-induced responses, isolated by subtracting luminance step alone responses
770 (blue) from individual composite luminance step-probe flash responses (dark gray) in
771 **b**. Lines connecting the response peaks highlight the time courses of suppression
772 relative to baseline flash-induced (2000 ms) responses.

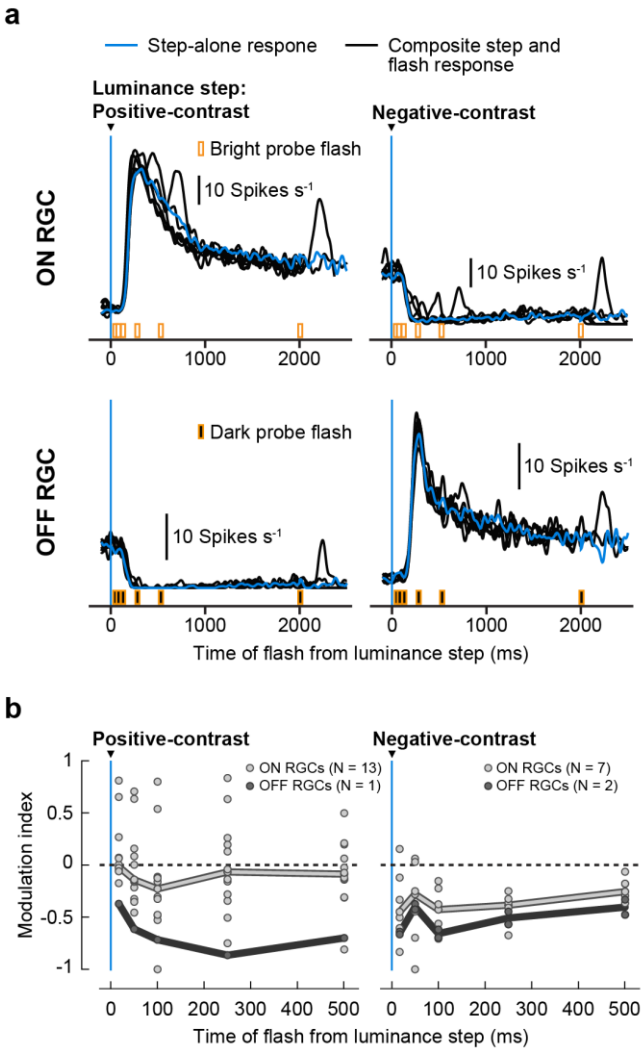
773 **Figure 6**



774

Figure 6 Model RGC responses to probe flashes following luminance steps.
a,b. Spiking response of model ON (columns 1-2) and OFF (columns 3-4) RGCs to luminance steps alone (**a**, blue) and to luminance steps followed by probe flashes (**b**, black) at 50, 250 and 2000 ms (different rows; analogous to real RGCs in Fig. 1b, c). Luminance steps are depicted by intensity bars in **a**. First column in each cell type: responses following a positive-contrast luminance step; second column: responses following a negative-contrast luminance step. Vertical blue lines: timing of luminance step; orange bars: timing of probe flashes. Note the ON RGC and OFF RGC did not spike in response to negative-contrast (column 2) and positive-contrast (column 3) luminance steps respectively.
c. Flash-induced responses, after subtracting **a** from **b**, overlaid to show the modulation of probe flash responses at different times (analogous to real RGCs in Fig. 1d). Lines connecting the response peaks highlight the time courses of suppression relative to baseline flash-induced responses (2000 ms).
d. Modulation indices for probe flashes in ON (light gray) and OFF model RGCs (dark gray), following positive-contrast (left panel) and negative-contrast (right panel) luminance steps. Modulation indices were calculated based on model responses to probe flashes presented at 10 ms intervals after luminance steps, and baseline as response to a flash at 2000 ms. Circle markers indicate modulation indices based on probe flashes at 50 and 250 ms shown in **b, c**. Cyan and red arrows highlight the suppression of opposite-contrast flashes at 50 ms in ON and OFF RGCs, respectively.
e. Modulation indices of model ON (cyan) and OFF (red) RGCs plotted as a function of RGC transiency. Individual panels correspond to different flash times (chosen from real-RGC recordings, Fig. 4) after positive-contrast (top row) and negative-contrast (bottom row) luminance steps. In **a-d**, transiency parameter was set to 1. Arrows highlight the same data as in **d**.
f. Same as in **e** but for real ON (light gray circles; N = 228; cyan line: linear regression fit) and OFF RGCs (dark gray circles; N = 92; red line: linear regression fit). These RGCs are a subset of the population data shown in Fig. 4b for which we could compute a transiency index (Methods). Consistent with model predictions, suppression after negative-contrast steps was weaker in less transient ON RGCs (bottom row, blue regression line has negative slope) and, after positive-contrast steps, suppression was weaker in less transient OFF RGCs (top row, red regression line has negative slope). Numbers in each panel indicate the slope of the fits and asterisk symbol indicates statistically significant slope (slope $\neq 0$, $p < 0.01$, two-tailed t-test).
Fig. S12 shows model RGC responses based on real cone data of Fig. 5 instead of model cone responses.

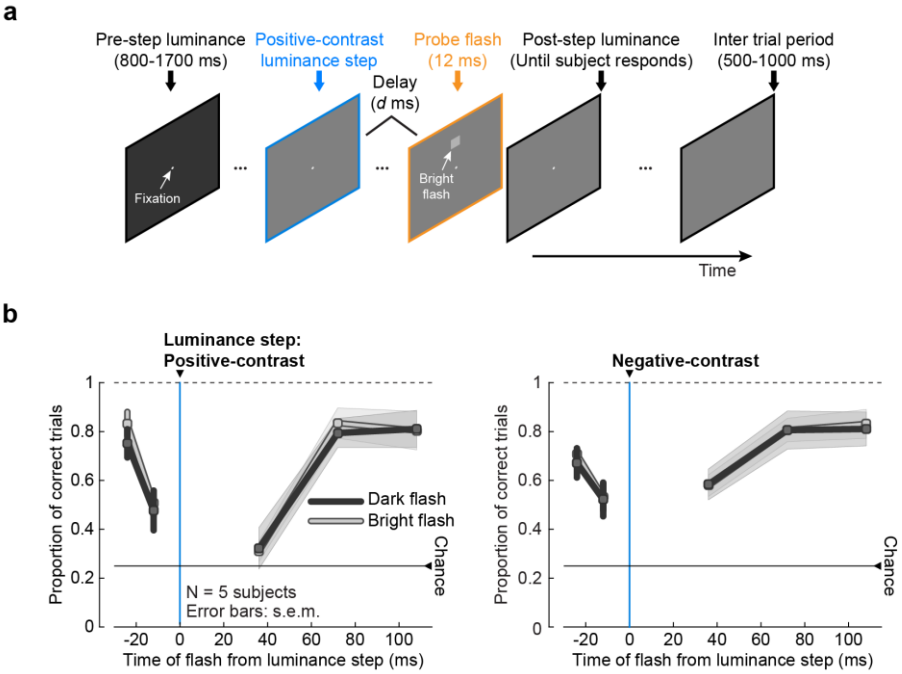
812 **Figure 7**



813

814 **Figure 7 Retinal saccadic suppression in Macaque RGCs.**
815 **a.** Average activity of an example macaque ON RGC (top) and OFF RGC (bottom) to
816 positive-contrast (left column) and negative-contrast (right column) luminance steps
817 alone (blue traces) and luminance steps followed by probe flashes (black traces). ON
818 RGCs were analyzed for bright probe flashes and OFF RGCs for dark probe flashes.
819 Responses were averaged across the different positive-contrast (0.05 to 0.5
820 Michelson contrast, N = 10 sequences) and negative-contrast luminance steps (-0.05
821 to -0.5 Michelson contrast, N = 10 sequences).
822 **b.** Median modulation index (thick lines) of macaque ON (light gray) and OFF (dark
823 gray) RGCs for probe flashes presented after positive-contrast (left panel; N = 13 ON
824 RGCs, N = 1 OFF RGC) and negative-contrast luminance steps (right panel; N = 7 ON
825 RGCs, N = 2 OFF RGC). Circles represent modulation indices of individual RGCs.
826 This experiment was performed under scotopic light conditions.

827 **Figure 8**



828

829 **Figure 8 Perceptual suppression following luminance steps.**
830 **a.** Example visual task trial. Subjects fixated a small spot on a uniform background
831 with a set luminance (pre-step luminance) for a random duration (800-1700 ms).
832 Background luminance then increased (like shown here) or decreased (positive or
833 negative-contrast luminance step, respectively). At one of 5 times relative to the
834 luminance step (-24, -12, 36, 72, or 108 ms), a luminance pedestal (probe flash,
835 147.8 x 147.8 min arc) was applied for ~12 ms at one of four locations relative to the
836 fixation spot: 7 deg above (shown here), below, to the right, or to the left. The probe
837 flash was brighter (shown here) or darker than the current screen luminance. The
838 background remained at the post-step luminance until the subject responded with the
839 perceived location of the flash, plus an additional 500-1000 ms but without the fixation
840 spot, allowing the subject to relax. The current luminance was the pre-step luminance
841 of the consecutive trial.
842 **b.** Performance of human subjects (mean ± s.e.m., N = 5 subjects), to correctly
843 localize a dark (dark gray) or a bright (light gray) probe flash presented at different
844 times relative to positive-contrast (left panel) and negative-contrast (right panel)
845 luminance steps (blue line). Each subject's responses were averaged across the
846 different positive-contrast (0.3 to 0.56 Michelson contrast) and negative-contrast (-0.3
847 to -0.56 Michelson contrast) luminance steps. Perceptual performance was reduced
848 around the time of luminance steps, reflecting suppression, irrespective of the
849 combination of luminance step polarity and flash polarity. There were no statistically
850 significant differences in suppression of dark and bright probe flashes (two-tailed
851 Wilcoxon rank-sum test). Note that in the right panel, the suppression profile for bright
852 probe flashes almost completely overlaps the suppression profile of dark probe
853 flashes.

854 # STAR METHODS

855 ## Key resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
AAV2.7m8.hSyn.iGluSnFR	Dalkara lab (Institut de la Vision), (Dalkara et al., 2013)	N/A
Chemicals, Peptides, and Recombinant Proteins		
SR-95531	Sigma	N/A
Picrotoxin	Sigma	N/A
Strychnine	Sigma	N/A
Experimental Models: Organisms/Strains		
Mouse: PV-Cre (strain: B6;129P2-Pvalb$^{tm1(cre)Arbr}$/J)	Jackson laboratories	RRID: IMSR_JAX:008069
Mouse: Thy-S-Y (strain: C57BL/6-tg (ThystopYFPJS))	FMI Basel	N/A
Mouse: BL6 (C57BL/6)	Jackson laboratories	RRID: IMSR_JAX:000664
Pigs: domestic pigs	Department of Experimental Surgery at the Medical Faculty of the University of Tübingen	N/A
Software and Algorithms		
MATLAB	MathWorks	RRID: SCR_001622
Spike sorting algorithm for high-density microelectrode array recordings	(Diggelmann et al., 2018)	N/A
Spike sorting post hoc processing: UnitBrowser	(Idrees et al., 2016)	N/A
Other		
Low density multielectrode array setup: 60pMEA200/30ir-Ti-gt	Multichannel Systems	N/A
High density multielectrode array setup: HiDens CMOS	Hierlemann Lab, (Frey et al., 2009)	N/A
High density multielectrode array setup: MaxOne	MaxWell Biosystems, (Müller et al., 2015)	N/A
High density multielectrode array setup	Chichilnisky et al (Chichilnisky and Baylor, 1999)	N/A

MOM-type two-photon microscope setup	Designed by W. Denk, MPI, Heidelberg; purchased from Sutter Instruments / Science Products	

Contact for reagent and resource sharing

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Thomas A. Münch (thomas.muench@gmail.com).

Experimental model and subject details

Animals

We performed electrophysiological experiments on ex vivo mouse, pig and macaque retinae; and imaging experiments on ex vivo mouse retinae.

Mouse and pig ex vivo retinae experiments were performed in Tübingen, in accordance with German and European regulations, and animal experiments were approved by the Regierungspräsidium Tübingen. Macaque ex vivo retina experiment was performed at Stanford University. Eyes were removed from a terminally anesthetized macaque rhesus monkey used by other laboratories in the course of their experiments, in accordance with the Institutional Animal Care and Use Committee guidelines of Stanford University.

For mouse retina electrophysiology, we used 47 retinae from 15 male and 30 female *PV-Cre x Thy-S-Y* mice (*B6;129P2-Pvalb*$^{tm1(cre)Arbr}$*/J × C57BL/6-tg (ThystopYFPJS))*, 3-12 months old, which are functionally wild type (Farrow et al., 2013; Münch et al., 2009; Tikidji-Hamburyan et al., 2015). Additionally, we recorded the cone output from 4 retinae obtained from two C57BL/6 male mice, 9 to 10 weeks old. We housed mice on a 12/12 h light/dark cycle, in ambient temperatures between 20-22 °C and humidity levels of 40%.

We also replicated experiments on 9 pig retinae obtained from domestic female pigs after they had been sacrificed during independent studies at the Department of Experimental Surgery at the Medical Faculty of the University of Tübingen. Pigs were

881 anesthetized using atropine, azaperone, benzodiazepine (midazolam), and ketamine,
882 and then sacrificed with embutramide (T61). Before embutramide administration,
883 heparin was injected.

884 One experiment was conducted with a retina extracted from a macaque rhesus
885 monkey.

886 **Humans**

887 Human psychophysics experiments were performed in Tübingen. Human subjects
888 provided written, informed consent, and they were paid 10 Euros per session of 60
889 minutes each, for three sessions. Human experiments were approved by ethics
890 committees at the Medical Faculty of Tübingen University, and they were in
891 accordance with the Declaration of Helsinki.

892 In total, we collected data from 5 subjects (24-29 years old; one female).

893 Method details

894 **Retina electrophysiology (mouse and pig): procedure and laboratory setup**

895 Mice were dark adapted for 4-16 h before experiments. We then sacrificed them under
896 dim red light, removed the eyes, and placed eyecups in Ringer solution (in mM: 110
897 NaCl, 2.5 KCl, 1 $CaCl_2$, 1.6 $MgCl_2$, 10 D-glucose, and 22 $NaHCO_3$) bubbled with 5%
898 CO_2 and 95% O_2. We removed the retina from the pigment epithelium and sclera while
899 in Ringer solution.

900 Pigs were dark-adapted for 15-20 min before sacrifice. Immediately after
901 veterinary-confirmed sacrifice, the eyes were enucleated under dim red light, and the
902 cornea, lens, and vitreous were removed. Eyecups were kept in CO_2-independent
903 culture medium (Gibco) and protected from light. We transported eyecups to our
904 laboratory and cut pieces from mid-peripheral or peripheral retinae. Only those retinae
905 which showed ganglion cell responses to light stimuli were used in our experiments.

906 We recorded mouse and pig retinal ganglion cell (RGC) activity using either low- or
907 high-density multi-electrode arrays (MEAs). The low-density setup consisted of a

908 perforated 60-electrode MEA (60pMEA200/30ir-Ti-gt, Multichannel Systems (MCS),
909 Reutlingen, Germany) having a square grid arrangement and 200 μm inter-electrode
910 distance. We whole mounted an isolated retina on a nitrocellulose filter (Millipore) with
911 a central 2 x 2 mm hole. The mounted retina was placed with the RGC side down into
912 the recording chamber, and good electrode contact was achieved by negative
913 pressure through the MEA perforation. We superfused the tissue with Ringer solution
914 at 30-34 °C during recordings, and we recorded extracellular activity at 25 kHz using
915 a USB-MEA-system (USB-MEA 1060, Multichannel Systems) or a memory-card
916 based system (MEA1060, Multichannel Systems). Data was acquired using MC Rack
917 version 4.6.2 (Multichannel Systems). A detailed step-by-step approach is provided in
918 (Reinhard et al., 2014).

919 The high-density MEA setup consisted of either a HiDens CMOS MEA (Frey et al.,
920 2009) (developed by the lab of Andreas Hierlemann, Basel, Switzerland) or a MaxOne
921 system (Müller et al., 2015) (Maxwell Biosystems, Basel, Switzerland). The HiDens
922 CMOS MEA featured 11,011 metal electrodes with inter-electrode (center-to-center)
923 spacing of 18 μm placed in a honeycomb pattern over an area of 2 x 1.75 mm. Any
924 combination of 126 electrodes could be selected for simultaneous recording. The
925 MaxOne MEA featured 26,400 metal electrodes with center-to-center spacing of 17.5
926 μm in a grid-like arrangement over an area of 3.85 x 2.1 mm. In this system, up to
927 1024 electrodes could be selected for simultaneous recordings. For each experiment,
928 a piece of isolated retina covering almost the entire electrode array was cut and placed
929 RGC-side down in the recording chamber. We achieved good electrode contact by
930 applying pressure on the photoreceptor side of the retina by carefully lowering a
931 transparent permeable membrane (Corning Transwell polyester membrane, 10 μm
932 thick, 0.4 μm pore diameter) with the aid of a micromanipulator. The membrane was
933 drilled with 200 μm holes, with center-center distance of 400 μm, to improve access
934 of the Ringer solution to the retina. We recorded extracellular activity at 20 kHz using
935 FPGA signal processing hardware. In the case of the HiDens CMOS MEA, data were
936 acquired using custom data acquisition software, called MEA 1k Scope (developed by
937 the lab of Andreas Hierlemann, Basel, Switzerland). In the case of the MaxOne MEA,
938 data were acquired using MaxLab software provided by Maxwell Biosystems, Basel,
939 Switzerland.

940 In total, we performed 59 recordings, 47 from mouse and 12 from pig retinae. 24 of
941 the 59 recordings were done using low-density MEAs. Once a basic experimental
942 protocol was established, we shifted to HiDens CMOS MEA providing much higher
943 throughput. 12 experiments were done using this setup. We upgraded to the MaxOne
944 MEA for even higher throughput and did 23 recordings using this setup. A subset of
945 the data collected from 32 of the 59 recordings (20 from mouse and 12 from pig
946 retinae), was also used in our previous study (Idrees et al., 2020). Here, we show
947 further in-depth analysis of that data.

948 We presented light stimuli to the retinal piece that was placed on the MEA using a DLP
949 projector running at 60 Hz (Acer K11 for low-density MEA experiments and Lightcrafter
950 4500 from EKB Technologies Ltd. with internal red, green and blue light-emitting
951 diodes, for high-density MEA experiments). 60 Hz is above the flicker fusion frequency
952 of both mouse and pig retinae; therefore, the framerate of these projectors was
953 adequate for our purposes. The Acer K11 projector had a resolution of 800 x 600
954 pixels covering 3 x 2.25 mm on the retinal surface. Lightcrafter 4500 had a resolution
955 of 1280 x 800 pixels, extending 3.072 x 1.92 mm on the retinal surface. We focused
956 images onto the photoreceptors using a condenser (low-density MEA recordings,
957 illumination from below) or a 5x objective (high-density MEAs, illumination from
958 above). In each case, the light path contained a shutter and two motorized filter wheels
959 with a set of neutral density (ND) filters (Thorlabs NE10B-A to NE50B-A), having
960 optical densities from 1 (ND1) to 5 (ND5). The filters allowed us to adjust the absolute
961 light level of the stimulation.

962 We measured the spectral intensity profile (in μW cm^{-2} nm^{-1}) of our light stimuli with a
963 calibrated USB2000+ spectrophotometer (Ocean Optics) and converted the physical
964 intensity into a biological equivalent of photoisomerizations per rod photoreceptor per
965 second (R*rod^{-1}s^{-1}), as described before (Tikidji-Hamburyan et al., 2015). Light
966 intensities of the projector output covered a range of 3 log units (i.e. 1000-fold
967 difference between black and white pixels, over an 8-bit range). We linearized the
968 projector output, and we used only grayscale images of limited contrast, spanning at
969 most the range from 0 to 120 in the 8-bit range of the projector (see stimulus
970 description below for details). Absolute light intensities were set to the mesopic level,
971 where a stimulus intensity of '30' in our 8-bit DLP projector scale (0-255) corresponded

972 to 225 to 1000 $R^*rod^{-1}s^{-1}$, depending on the experimental rig used for the experiment
973 (i.e. different DLP projectors and MEAs). We pooled all data from the different rigs as
974 separate individual analyses from the individual setups revealed no effects of
975 recording conditions in the different setups. For experiments of Fig. S4, we also
976 recorded at scotopic light levels where a stimulus intensity of '30', corresponded to 23
977 $R^*rod^{-1}s^{-1}$ at scotopic level.

978 **Retina electrophysiology (macaque): procedure and laboratory setup**

979 In one experiment, we recorded the activity of macaque retinal ganglion cells. For this
980 experiment we used a high-density MEA, as described previously (Chichilnisky and
981 Baylor, 1999; Field et al., 2007). Following enucleation, the anterior portion of the eye
982 and vitreous were removed. The eye was stored in a dark container in oxygenated
983 Ames' solution (Sigma, St. Louis, MO) at 33°C, pH 7.4. Under infrared illumination, a
984 small piece of retina approximately 1x1 mm, from a retinal region with eccentricity
985 around 12 mm (4.0-17mm temporal equivalent eccentricity; (Chichilnisky and Kalmar,
986 2002)), was dissected and placed ganglion cell side down on a MEA for recording.
987 The retina pigment epithelium remained attached during the recording; the retina was
988 perfused with oxygenated Ames' solution. A custom planar large-scale MEA (Field et
989 al., 2007; Litke et al., 2004) with a hexagonal outline of 519 electrodes at 30 μm pitch
990 was used. Recorded voltages were band-pass filtered, amplified, and digitized at 20
991 kHz using custom electronics (Litke et al., 2004). Spike sorting process was described
992 previously (Field et al., 2007).

993 Visual stimulation was performed with the optically reduced image of a
994 gamma-corrected OLED microdisplay (eMagin) refreshing at 60.35 Hz focused on the
995 photoreceptor outer segments. The visual stimulus was delivered through the
996 mostly-transparent electrode array. The power of each display primary was measured
997 at the preparation with a calibrated photodiode (UDT Instruments). At the mean
998 background illumination level, the photoisomerization rates for the rods and the L, M,
999 and S cones were approximately 29, 9, 9, and 2 $P^*receptor^{-1}s^{-1}$, respectively (see Li
1000 et al., 2014), placing the retina in scotopic regime.

1001 **Retina electrophysiology: pharmacology**

1002 In several MEA experiments, we used pharmacological agents to block specific
1003 receptors in the mouse retina. To block $GABA_A$ receptors selectively, we used 5 μM
1004 SR-95531 (gabazine, an antagonist of $GABA_A$ receptors; Sigma). To block both
1005 $GABA_A$ and $GABA_C$ receptors, we used 100 μM picrotoxin (an antagonist of $GABA_A$
1006 and $GABA_C$ receptors; Sigma). To block glycine receptors, we used 1 μM Strychnine
1007 (antagonist of Glycine receptors).

1008 We first prepared a 1000x stock solution of these pharmacological blockers as follows:
1009 SR-95531 was dissolved in water at a concentration of 5 mM; picrotoxin was dissolved
1010 in DMSO at a concentration of 100 mM; Strychnine was dissolved in Chloroform at a
1011 concentration of 1mM. During the experiments, we pipetted the stock solution to the
1012 Ringer solution in a 1:1000 ratio. Wash-in was performed for 20 min.

1013 **Cone photoreceptors imaging: laboratory setup**

1014 To record the output of cone photoreceptors in the mouse retina, we measured the
1015 glutamate release using an intensity based glutamate-sensitive fluorescent reporter,
1016 iGluSnFR (Marvin et al., 2013) expressed in horizontal cell processes post-synaptic to
1017 cone terminals, using a viral approach. We recorded the cone output from 4 retinae
1018 obtained from two C57BL/6 male mice, 9 to 10 weeks old. Below, we reproduce the
1019 methods, previously described in Szatko et al (Szatko et al., 2020).

1020 We dark-adapted the mice for ≥1 h before the experiments. They were then
1021 anaesthetized using isoflurane (Baxter) and sacrificed by cervical dislocation. The
1022 eyes were enucleated and hemisected in carboxygenated (95% O_2 and 5% CO_2)
1023 artificial cerebrospinal fluid (ACSF) solution containing (in mM): 125 NaCl, 2.5 KCl, 2
1024 $CaCl_2$, 1 $MgCl_2$, 1.25 NaH_2PO_4, 26 $NaHCO_3$, 20 glucose, and 0.5 L-glutamine (pH
1025 7.4). We then moved the tissue to the recording chamber where it was continuously
1026 perfused with carboxygenated ACSF at ~36 °C. In these experiments, ACSF
1027 contained ~0.1 μM Sulforhodamine-101 (SR101, Invitrogen) to reveal blood vessels
1028 and any damaged cells in the red fluorescence channel (Euler et al., 2019). All
1029 procedures were carried out under very dim red (>650 nm) light.

1030 iGluSnFR was expressed in the retina by viral transduction of
1031 AAV2.7m8.hSyn.iGluSnFR, generated in the Dalkara lab (Institut de la Vision) as
1032 described in (Dalkara et al., 2013; Khabou et al., 2016). The iGluSnFR plasmid
1033 construct was provided by J. Marvin and L. Looger (Janelia Research Campus, USA).
1034 A volume of 1 μL of the viral construct was injected into the vitreous humour of the
1035 mice, anaesthetized with 10% Ketamine (Bela-Pharm GmbH & Co. KG) and 2%
1036 xylazine (Rompun, Bayer Vital GmbH) in 0.9% NaCl (Fresenius). For the injections,
1037 we used a micromanipulator (World Precision Instruments) and a Hamilton injection
1038 system (syringe: 7634-01, needles: 207434, point style 3, length 51 mm, Hamilton
1039 Messtechnik GmbH). Imaging experiments were performed 3-4 weeks after injection.
1040 In the outer retina, iGluSnFR was predominantly expressed in horizontal cells. As the
1041 expression tended to be weaker in the central retina, most scan fields were acquired
1042 in the medial to peripheral ventral or dorsal retina.

1043 To record the iGluSnFR signal, we used a MOM-type two-photon microscope setup
1044 (designed by W. Denk, MPI, Heidelberg; purchased from Sutter Instruments / Science
1045 Products). The design and procedures have been previously described in (Euler et al.,
1046 2009, 2019). In brief, the system was equipped with a mode-locked Ti:Sapphire laser
1047 (MaiTai-HP DeepSee, Newport Spectra-Physics), two fluorescence detection
1048 channels for iGluSnFR (HQ 510/84, AHF/Chroma) and SR101 (HQ 630/60, AHF), and
1049 a water immersion objective (W Plan-Apochromat 20x /1.0 DIC M27, Zeiss). The laser
1050 was tuned to 927 nm for imaging iGluSnFR. For image acquisition, we used custom
1051 made software (ScanM by M. Müller and T. Euler) running under IGOR Pro 6.3 for
1052 Windows (Wavemetrics), taking time-lapsed 128 x 128 pixel image scans at 3.9 Hz in
1053 the outer plexiform layer (OPL).

1054 For light stimulation in cone imaging experiments, we used the Lightcrafter (LCr;
1055 DPM-E4500UVBGMKII), a DLP projector from EKB Technologies Ltd. with internal UV
1056 and green light-emitting diodes (LEDs). The light from the DLP projector was focused
1057 through the objective. To optimize spectral separation of mouse M- and S- opsins,
1058 LEDs were band-pass filtered (390/576 Dualband, F59-003, AHF/Chroma). LEDs of
1059 the DLP projector were synchronized with the microscope's scan retrace. Stimulus
1060 intensity (as isomerization rate, $P^{*}cone^{-1}s^{-1}$) was calibrated to range from ~500 (black
1061 image) to ~20,000 for M- and S-opsins. In addition, a steady illumination component

of ~10^4 $P^*cone^{-1}s^{-1}$ was present during the recordings because of two-photon excitation of photopigments. The overall light intensity falling on to the retina was therefore in the low photopic regime. The light stimulus was centered to the recording field before every experiment. For all experiments, the retinal tissue was kept at a constant mean stimulator intensity level for at least 15 s after the laser scanning started and before stimuli were presented.

Human psychophysics: laboratory setup

We used a similar laboratory setup as described previously (Idrees et al., 2020). 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 (Hafed, 2013), 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 backgrounds in the luminance step experiment (Fig. 8) were always presented at an average luminance of 49.84 cd m^{-2}, and the monitor was linearized (8-bit resolution) such that equal luminance increments and decrements for luminance steps were possible around this average. In total, we collected data from 5 subjects (24-29 years old; one female). A subset of the data from 4 subjects was used in our previous study (Idrees et al., 2020). Here, we perform novel analyses of the complete dataset, in addition to one new subject.

Visual stimuli: Retina electrophysiology (Figs. 1-4, 7, S3-S10, S13)

In retina electrophysiology experiments, we used two broad visual stimulation paradigms: a saccade (texture displacements) paradigm (Fig. S1a), and a luminance step paradigm (Fig. S1b), described in detail below. In different experiments we used different spatial and/or pharmacological manipulations of these two paradigms.

1087 ***Saccade (texture displacements) paradigm***

1088 Background textures

1089 We created background textures (Fig. S2a) by convolving a random binary (i.e. white
1090 or black) pixel image with a two-dimensional Gaussian blurring filter (Schwartz et al.,
1091 2012) defined by the kernel

1092 $$G(x,y) = e^{\frac{-(x^2+y^2)}{2\sigma^2}} \qquad \text{(Equation 1)}$$

1093 The parameter σ of the kernel influenced the amount of blurring. This resulted in
1094 textures having effectively low-pass spectral content (Fig. S2b) with a cutoff frequency
1095 depending on σ. For easier interpretation, we define the spectral content of these
1096 textures by a spatial scale. Intuitively, the spatial scale approximates the size of the
1097 smallest dark and bright image blobs of the texture (Fig. S2a). Quantitatively, the
1098 spatial scale is defined as the $2*\sigma$ parameter of the Gaussian blurring filter. We
1099 generated textures with four different spatial scales: 25, 50, 150 and 300 μm, that
1100 resulted in dark and bright image blobs approximating a range of receptive field sizes
1101 between bipolar cells (texture with spatial scale 25 μm, see (Zhang et al., 2012)) and
1102 RGCs (textures with spatial scale 150 and 300 μm). In other words, coarser textures
1103 matched the resolution of RGCs, and finer textures matched the resolution of one
1104 processing stage earlier, the retinal bipolar cells. Calculating power spectra for the
1105 textures (Fig. S2b) confirmed that the spatial scale and hence the cutoff frequencies
1106 were consistent with this design aim. In different experiments, we used textures of all
1107 or a subset of the different spatial scales.

1108 We normalized the pixel intensities in the textures to have uniform variations in
1109 luminance around a given mean. We used pixel intensities (from our 8-bit resolution
1110 scale) ranging from 0 to 60 around a mean of 30, or ranging from 30 to 90 around a
1111 mean of 60 (see sub-section *Saccades and probe flashes* for when each paradigm
1112 was used).

1113 Saccades and probe flashes

1114 To simulate saccades in our ex vivo retina electrophysiology experiments, we
1115 displaced the texture across the retina in 6 display frames (100 ms at 60 Hz refresh
1116 rate). For easier readability, we usually 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 3100-9300 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 (Sakatani and Isa, 2007). Similar to primates, mice also have oculomotor behavior, even under cortical control (Itokazu et al., 2018). For example, they make, on average, 7.5 saccade-like rapid eye movements per minute when their head is fixed (Sakatani and Isa, 2007) (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, retinal “saccadic suppression” is not critically dependent on the details of movement kinematics, as it is triggered by visual transients (Fig. 4, also see Figs. 4, 5 in Idrees et al., 2020).

Each trial consisted of successive sequences (Fig. 1a, S1a) that combined a “saccade” with a probe flash, as follows: there was first a “pre-saccade” fixation of 2 seconds, where the texture remained static over the retina, then a 100 ms “saccade”, followed by “post-saccade” fixation where the texture again remained static over the retina but now with a shifted texture. At a certain time from “saccade” onset (delay *d*, range: 50 ms to 2100 ms), we presented a probe flash (see below). Following the probe flash, the texture remained static at the post-saccade fixation position for another 2 seconds before the next saccade of the successive sequence occurred. The post-probe-flash fixation of one sequence was therefore also the pre-saccade fixation of the next sequence. This way the texture remained visible during the entire trial, being translated during saccades of the successive sequences. In a single trial, 39 such sequences occurred. In each successive sequence, the direction and amplitude of the saccade was pseudo-randomly determined by the range of allowed saccade amplitudes and directions. The texture always landed at unique locations within a trial.

The end result was that, within a single trial, RGCs experienced a wide spectrum of saccade amplitudes, directions and contrasts across these 39 saccades. As such, by analyzing the average effects of the 39 saccades on RGC responses to probe flashes, we captured a wide range of saccade-induced kinematics and luminance changes over the RGC receptive fields.

In most cases, the probe flash had a duration of 2 frames (~33 ms). We used 1 frame (~16 ms) in a subset of earlier experiments (mouse: 709 of 1616 cells; pig: 116 of 228 cells). Results were pooled across these paradigms as the effects were indistinguishable. The probe flash was a full-screen positive ("bright") or negative ("dark") stimulus transient.

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 replaced the texture with a homogeneous bright (pixel intensity of 60 in our 8-bit projectors) or dark (pixel intensity of 0) full-screen (in these experiments, the textures themselves had intensities ranging from 0 to 60 pixel intensity; see *Background textures* above). This way, the flash contrast from the underlying background luminance was variable across space (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. This way, spatial 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). Out of the 1616 RGCs that we analyzed for saccadic suppression across all experiments where texture displacements were used as saccades (irrespective of the spatial or pharmacological manipulations), 1129 RGCs experienced such probe flashes, whereas the rest (487 RGCs) 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 to 0 in the first frame of the flash then going to 60 (on our 8-bit projectors) 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 the 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 flash at all (Fig. S1a shows schematic of one such trial). Moreover, we repeated the total number of conditions (e.g. the interleaved 15 trials) 4 times per session, and we averaged responses across those 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 texture scales). Therefore, the total number of saccades shown in any given experiment could be computed by #trials x 39 saccades per trial x #textures x #repetitions. A typical experiment lasted 10-12 hours. If the combination of conditions would have required even longer recordings in a given experiment, we reduced the number of conditions (e.g. we presented flashes at fewer delays, or used fewer texture scales).

<u>Full-field saccades</u>

In the full-field saccades experiments, saccades and probe flashes occurred over the entire retina. This was our main experimental paradigm that we used for characterizing how saccades modulate RGC responses to probe flashes. This paradigm was also used as a control in experiments in which we applied different spatial or pharmacological manipulations of this paradigm to probe for the spatial origins and mechanisms of saccadic suppression. Further, results from this paradigm served as a baseline standard across different experimental rigs. This paradigm was used in a total of 32 retinal recordings (32 retinae from 30 mice). In different recordings, we used a subset of the texture scales and probe flash delays. This explains the different values of N seen for different conditions in, for example, Figs. 1e, S3. However, to ensure comparison, some conditions always overlapped across different recordings. This paradigm was also used in 12 recordings with retinae from 6 pigs.

1212 Periphery saccades (global component of suppression)

1213 In this manipulation, we restricted saccades to the RGC's receptive fields periphery
1214 (i.e. its far surround). This spatial manipulation was used to investigate the spatial
1215 origins of the global component of suppression (Figs. 2a-b, S5). We performed 13
1216 recordings (13 retinae from 13 mice) with this paradigm, always with the high-density
1217 multielectrode array system (MaxOne by MaxWell), as it provided a large electrode
1218 area (~2 x 4 mm^2) for the retina to be placed on. The recording region was typically
1219 either a high density block of electrodes (inter-electrode spacing: ~17.5 μm) or a block
1220 with one-electrode spacing (inter-electrode spacing: ~35 μm). The recording region
1221 was selected close to the center of the electrode array. We centered a large square
1222 mask (1000 x 1000 μm^2) over the recording region to restrict the texture and saccade
1223 presentation to the periphery of RGC receptive fields (Fig. S5b). The mask had a
1224 homogenous intensity corresponding to the mean luminance of the texture. At different
1225 times relative to texture displacements, full-field probe flashes were presented, similar
1226 to experiments with full-field texture displacements. The intensity of each pixel of the
1227 stimulus (both the mask and the texture regions) was adjusted for the probe flashes,
1228 either decreased or increased by a pixel value of 30 (on our 8-bit projectors) for dark
1229 and bright probe flashes, respectively. In all periphery saccade experiments, we used
1230 probe flash duration of ~33 ms, and a coarse texture background of spatial scale 300
1231 μm.

1232 Checkerboard mask paradigm (local component of suppression)

1233 In this spatial manipulation we presented saccades and flashes in small square
1234 regions spread equidistantly over the entire retina. Each square region measured
1235 100 x 100 μm^2, separated from adjacent squares by an edge-edge gap of 100 μm.
1236 The gap was kept at mean luminance throughout the experiment. Saccades and
1237 flashes could either be presented in all the regions (similar to full-field saccades,
1238 except for the gap), or in alternate regions (Fig. 2e, S7a), arranged like in a
1239 checkerboard. This paradigm was used to investigate the origins of the local
1240 component of suppression (Figs. 2f, S7a-c). We performed 4 recordings (4 retinae
1241 from 3 mice) with this paradigm, always with the low density MCS MEA rig. In all
1242 experiments with this spatial manipulation, we used probe flash duration of ~17 ms,
1243 and a coarse texture background of spatial scale of 150 μm.

1244 ***Luminance step paradigm***

1245 In this paradigm (Fig. S1b), we used no textures at all. The screen was always a
1246 homogenous gray field, and the visual event of a "saccade" was replaced by an
1247 instantaneous step to a different gray value. The gray backgrounds had intensities
1248 between 30 and 90 (on our 8-bit projector). The instantaneous step in intensity caused
1249 either a positive contrast luminance step (in the range of +0.03 to +0.50 Michelson
1250 contrast) or a negative contrast luminance step (-0.03 to -0.50 Michelson contrast).
1251 This paradigm was used to characterize the stimulus-stimulus interactions that
1252 ultimately trigger retinal saccadic suppression (Figs. 4, 7, S8-10). We performed a total
1253 of 4 recordings (4 retinae from 4 mice) with this paradigm, always with the high density
1254 MaxOne MEA rig. A trial consisted of either 56 or 156 successive sequences
1255 (Fig. S1b) that each combined a luminance step with a probe flash, as follows: there
1256 was first a “pre-step” fixation of 2 seconds where the retina was exposed to a fixed
1257 gray level (analogous to “pre-saccade” fixation in texture displacements), then an
1258 instantaneous switch to “post-step” fixation (analogous to “post-saccade”). At a certain
1259 time from the luminance step (delay: 17, 33, 50, 100, 250, 500, 1000 or 2000 ms), we
1260 presented a 2-frame (~33 ms) dark (-0.33 Michelson contrast) or bright (+0.33
1261 Michelson contrast) probe flash. Some sequences contained no probe flash, the next
1262 luminance step then happened 4 seconds after the previous one. In a given
1263 experiment, we had 17 trials representing the 17 conditions: 8 flash delays x 2 probe
1264 flash polarities + 1 condition with no probe flash. Similar to the saccade paradigm, we
1265 always interleaved all conditions; that is, in any one of the 17 necessary trials, each
1266 luminance step could be followed by a bright or a dark probe at any of the 8 delays, or
1267 no probe flash. Moreover, we repeated the 17 trials at least 4 times.

1268 A shorter version of this paradigm was used in our macaque retina recording (Fig. 7).
1269 Here, a trial consisted of 20 successive sequences. The 20 luminance steps induced
1270 contrasts in the range -0.5 to +0.5. Flashes of ~33 ms were presented with a delay of
1271 17, 50, 100, 250, 500 and 2000 ms after the luminance step.

1272 ***Other stimuli***

1273 Finally, we used other stimuli unrelated to the main experiments to help us
1274 characterize RGC properties (e.g. response polarity, latency, transiency, and spatial
1275 receptive fields). These stimuli had the same mean intensities and intensity ranges as

the textures or luminance steps 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, or the luminance step experiment, 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 linear filters representing the temporal receptive fields of RGCs 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 linear filters representing the spatial receptive fields of RGCs.

Visual stimuli: Cone photoreceptors imaging (Figs. 5, 6)

For cone imaging experiments, we used a minimalistic version of the luminance step paradigm used in retina electrophysiology. A homogeneous background alternated between a darker (pixel intensity 50 on 8-bit projector) and brighter gray value (pixel intensity 120 on 8-bit projector); the transitions between these two background values represented positive and negative contrast of 0.4 Michelson contrast. At various times after the luminance step (50, 250, and 2000 ms) we presented a probe flash (100 ms duration, +0.33 or -0.33 Michelson contrast). The probe flash at 2000 ms served as the baseline. The next background transition always happened 2 s after the preceding probe flash. The combination of the two luminance steps and the two probe flash polarities yielded a total of four combinations: negative-contrast luminance step followed by dark flash; negative-contrast luminance step followed by bright flash; positive-contrast luminance step followed by dark flash; and positive-contrast luminance step followed by bright flash. A single trial (Fig. S1c) was composed of the four step-flash combinations occurring three times (for the three delays with which the

flash was presented); and the negative- and positive-contrast luminance step without a flash. Within a trial, these conditions were randomized. A trial was repeated three times. The luminance steps and the flashes were presented within a 700 μm disc region centered over the scan field.

Visual stimuli: Human psychophysics (Fig. 8)

In the human psychophysics experiment (Fig. 8), we mimicked the retinal luminance step experiments of Fig. 4. The paradigm (Fig. 8a) was similar to the one described in (Idrees et al., 2020). Subjects fixated a central fixation spot over a gray background that remained there for the entire duration of a trial. The background had one of 8 luminances (22.4, 30.24, 38.08, 45.92, 53.76, 61.6, 69.44, 77.28 cd m^{-2}). After a random initial fixation duration (800-1700 ms after fixation spot onset), the luminance of the background was changed suddenly (in one display frame update) to one of the remaining 7 luminances, inducing a positive-contrast luminance step or a negative-contrast luminance step. In our analysis, we used the luminance steps that induced contrasts in the range +0.3 to +0.56 Michelson contrast and -0.3 to -0.56 Michelson contrast. At one of 5 different possible times relative to the time of the luminance step (-24, -12, 36, 72, or 108 ms), a luminance pedestal (probe flash) was applied briefly 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 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 to represent bright and dark probe flashes. We ensured that the contrast of the flash (relative to the currently displayed background luminance) was always the same across all trials: +0.033 for a bright flash, and -0.033 for a dark flash. Following each trial, the fixation spot was removed from the background to allow the subjects to relax. This inter trial period lasted for 500-1000 ms. The next trial happened consecutively, in a way that the current luminance of the background was used as the pre-step luminance. Subjects maintained fixation

1339 throughout all trials (except the inter trial period) and simply reported the locations of
1340 the brief flashes. Each subject performed 3 sessions, with 1120 trials per session.

1341 Quantification and statistical analysis

1342 Data analysis: Retina electrophysiology

1343 *MEA recordings preprocessing*

1344 Low-density MEA recordings were high-pass filtered at a 500 Hz cutoff frequency
1345 using a tenth-order Butterworth filter. We extracted spike waveforms and times using
1346 thresholding, and we semi-manually sorted spikes using custom software. For
1347 high-density MEA recordings, we performed spike sorting by an offline automatic
1348 algorithm (Diggelmann et al., 2018) and assessed the sorted units using a custom
1349 developed tool, the UnitBrowser (Idrees et al., 2016). We judged the quality of all units
1350 using inter-spike intervals and spike shape variation. Low quality units, such as ones
1351 with high inter-spike intervals, missing spikes, or contamination, were discarded. All
1352 firing rate analyses were based on spike times of individual units. In total, we extracted
1353 3,510 high quality units after the spike sorting (referred to as RGCs from now on), from
1354 recordings of mouse retina. From pig retina recordings, we extracted 376 RGCs and
1355 from macaque retina we extracted 57 RGCs after the spike sorting. However, as we
1356 mention below, only a subset of these could be analyzed for saccadic suppression.

1357 *RGCs characterization: Receptive fields, ON-OFF index, Transiency index*

1358 We first characterized the properties of RGCs. We calculated linear filters in response
1359 to full-field Gaussian flicker and binary checkerboard flicker by summing the 500-ms
1360 stimulus history before each spike. The linear filters allowed determining cell polarity.
1361 Specifically, the amplitude of the first peak of the filter was used: If the peak was
1362 positively deflected, the cell was categorized as an ON cell; if negatively deflected, the
1363 cell was an OFF cell. ON cells were later always analyzed with respect to their
1364 responses to bright probe flashes, and OFF cells were analyzed with dark probe
1365 flashes. We determined the spatial receptive fields of RGCs by calculating the linear
1366 filters for each region (checker) defined by the binary checkerboard flickering stimulus.
1367 The modulation strength of each linear filter, measured as the standard deviation (s.d.)
1368 along the 500 ms temporal kernel, is an estimate for how strongly that region drives

1369 ganglion cell responses. We fitted the resulting 2D-map of s.d. values with a two
1370 dimensional Gaussian and took the 2-σ ellipse (long axis) as the receptive field
1371 diameter. For all other figures and analyses, we converted spike times to estimates of
1372 firing rate by convolving these times with a Gaussian of σ = 10 ms standard deviation
1373 and amplitude 0.25 $\sigma^{-1}e^{1/2}$.

1374 For each RGC, we used responses to full-field contrast steps to calculate an ON-OFF
1375 index, a transiency index, and a response latency index. These indices were used to
1376 characterize the properties of RGCs that we included in our analyses. The ON-OFF
1377 index was calculated by dividing the difference between ON and OFF step peak
1378 response by their sum. The resulting index values ranged between -1 (OFF) and +1
1379 (ON) and were then scaled to span between 0 (OFF) and +1 (ON). The transiency
1380 index was defined as the ratio of the response area within the first 400 ms and the
1381 total response area spanning 2000 ms. The resulting index had a value of 1 for pure
1382 transient cells. Response latency was calculated as the time from stimulus onset to
1383 90% of peak response. This value was normalized to the maximum response latency
1384 in our dataset to create the response latency index.

1385 ***Modulation index***

1386 To quantify retinal saccadic suppression, we first determined a baseline response,
1387 defined as the response to a probe flash approximately 2 s after texture displacement
1388 onset or 2 s after luminance step (delay between 1967 to 2100 ms, depending on the
1389 specific flash times used in a specific experiment). This baseline response was
1390 compared to responses of the same cell to the same flash when it occurred at an
1391 earlier time (i.e. closer in time to the “saccade”). Usually, the saccade-like texture
1392 displacements themselves caused significant neural responses (saccade-alone
1393 response, e.g. Fig. 1b), and the responses to the flashes were superimposed on these
1394 “saccade-responses”. We therefore first isolated the component of the responses
1395 caused by the flashes by subtracting the saccade-alone responses from the composite
1396 saccade and flash responses. We refer to this isolated component as the
1397 flash-induced responses.

1398 To get a robust estimate of the response to saccades-alone (i.e. without any flashes),
1399 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, across all saccades. 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. To quantify the modulation at a population level, we averaged the modulation indices of the individual RGCs in that population. For some analyses, we also calculated modulation indices of RGCs for each of the 39 individual "saccades" using the same procedure.

In some cells, individual saccades from the sequence of 39 were discarded. For example, imagine that "saccade No 3" gets discarded. This would happen when the baseline response strength (response to the probe flash with 2 s delay) after saccade No 3 is weak (specifically: peak amplitude less than 60% of the median of all 39 baseline response strengths). We did this to ensure that our modulation indices were not marred by a denominator approaching zero (e.g. if both flash and baseline responses were weak). We did, however, re-include some sequences. For example, if the probe flash with delay 100 ms after saccade No 3 triggered a strong response (specifically: peak amplitude larger than the median baseline response peak across the 39 saccades), then saccade No 3 would be re-included for the condition "100 ms delay". 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

1432 gave a robust response. For example, this could happen if a cell did not respond to
1433 the baseline flash but the saccade enhanced the response to a flash following it.

1434 Finally, to perform statistics, we applied tests both at the level of individual cells and
1435 at the level of the population. At the individual cell level, we determined whether the
1436 modulation index for a probe flash presented at a certain delay was significantly
1437 different from 0 (i.e. “Is the response of this cell modulated by the ‘saccade’?”). For
1438 this, we performed a one-tailed sign test of the null hypothesis that the 39 individual
1439 modulation indices (or its subset in case weak sequences were discarded, as
1440 described above) came from a distribution with zero median. The alternative
1441 hypothesis was that the median was below (for negative modulation index) or above
1442 (for positive modulation index) zero. The modulation index was considered significant
1443 (i.e. the flash response was modulated by the saccade) at $p < 0.05$. However, we did
1444 not consider cells significantly modulated if the test had a power ($1-\beta$) smaller than
1445 0.8, which could happen if we previously had to exclude too many sequences ($N \leq 39$).
1446 At the population level, we determined whether the retinal output as a whole was
1447 modulated by saccades. For this, we performed a two-tailed Wilcoxon signed rank test
1448 of the null hypothesis that the median of the distribution of modulation indices did not
1449 differ from 0. Lastly, we tested whether the population modulation index was
1450 significantly different across populations of ON and OFF RGCs or across different
1451 paradigms. For this, we performed a two-tailed Wilcoxon rank-sum test of the null
1452 hypothesis that the median of the distribution of modulation indices did not differ
1453 across the two populations being tested.

1454 Since the modulation index was based on responses to the brief probe flashes, it could
1455 only be computed for RGCs that did respond to these brief flash stimuli. In our analysis,
1456 we included all such RGCs. Of the spike sorted RGCs across all paradigms, we
1457 included: 2002 of 3510 in mice; 228 of 376 in pigs; and 15 of 57 in macaque.

1458 ***Saccade (texture displacement) paradigm***

1459 <u>Full-field saccades</u>

1460 We analyzed 1010 mouse RGCs (633 ON; 377 OFF) and 228 pig RGCs (197 ON; 31
1461 OFF) for saccadic suppression using the full-field version of the saccade paradigm. A
1462 subset of data from 688 of the 1010 mouse RGCs and a subset of data from all the

228 pig RGCs was presented previously (Idrees et al., 2020). Here, we perform novel analyses on the complete datasets from the RGCs recorded previously, in addition to analyzing the newly recorded RGCs. For each RGC, we quantified the modulation index for a full-field probe flash presented at different times from saccade onset.

A subset of these RGCs were also tested for saccadic suppression while blocking GABAergic and glycinergic inhibition. For 82 ON and 30 OFF RGCs, we had a direct comparison with and without $GABA_{A,C}$ blockers (5 μM SR-9553 + 100 μM Picrotoxin) (Fig. 2c, S6a). For 51 ON and 13 OFF RGCs, we had a direct comparison with and without $GABA_{A,C}$ blockers in addition to glycine blocker (1μM Strychnine; Fig. 2d, S6b).

In yet another subset of RGCs (72 ON; 49 OFF), we also analyzed saccadic suppression at scotopic light level (Fig. S4), which was 1 log unit dimmer than the light level at which all other recordings were performed. For these cells, we had a direct comparison of responses at scotopic and mesopic levels.

<u>Periphery saccades (global component of suppression)</u>

For each recorded RGC, we computed a masking factor (post hoc) to quantify how well its receptive field was covered by the 1000 x 1000 μm^2 mask. We first determined the spatial receptive field center of each RGC (described above in *RGCs characterization*). The masking factor was defined as the multiple of σ of the two dimensional Gaussian fit for which the ellipse just touched the mask boundary (Fig. S7d). Cells with receptive field centers within the mask were defined to have a positive masking factor while those lying outside were given a negative masking factor. The magnitude of the factor increased with distance from the edge of the mask. This way, cells for which the mask covered their receptive field centers and immediate surround had masking factors > 2 (these were the cells included in analysis shown in Figs. 2b, S5); cells with a mask covering only the receptive field center had masking factors between 1 and 1.5. Cells located close to the mask's edge, with masking factors between -1 to +1, had their receptive field centers partially exposed to the saccade. Finally, cells lying outside the mask where the receptive field center always experienced saccades had masking factors < -1. A total of 642 RGCs (401 ON; 241 OFF) from 13 experiments were recorded with this spatial layout of the saccade paradigm. Cells for which clear receptive fields could not be calculated were excluded

1494 from any further analysis. The exact number of cells for different conditions within this
1495 paradigm are reported in the results section. For each RGC, we calculated the
1496 modulation index for flashes presented at different times from saccade onset, in the
1497 same manner as described under the heading *Modulation index*. Fig. S7e shows the
1498 modulation index of individual RGCs as a function of its masking factor. Only a subset
1499 of RGCs with masking factors in the range -3 to +5 were included in this analysis. The
1500 median modulation index was calculated by taking the median of modulation indices
1501 of RGCs within a 1.2 masking factor window, sampled at intervals of 0.1 (running
1502 median).

1503 In a subset of these experiments, we analyzed the effects of blocking GABAergic
1504 inhibition (using 5 μM SR-95531) on the modulation of probe flash responses in RGCs
1505 with masking factors > 2, i.e. with receptive field centers and immediate surround
1506 effectively masked. This included 62 ON RGCs and 35 OFF RGCs with robust
1507 responses to brief probe flashes with the pharmacological agent; only these RGCs
1508 were analyzed for saccadic suppression under the presence of SR-95531 (Fig. 2b). In
1509 a subset of these RGCs (29 ON; 25 OFF), we had a direct comparison of modulation
1510 indices with the control condition, i.e. periphery saccades in the absence of SR-95531
1511 (Fig. S5d).

1512 <u>Checkerboard mask regions saccades</u>

1513 In experiments where we investigated the local component of suppression
1514 (Figs. 2e-f, S7b, c), we had three stimulus presentation settings (Fig. S7a): (1)
1515 saccade and flash presented in all squares regions; (2, 3) saccades and flashes in
1516 alternate regions of a hypothetical binary checkerboard. For each recorded RGC we
1517 calculated the modulation index for the following two scenarios: saccades and flashes
1518 in all regions (Fig. S7a presentation setting 1); and saccades excluded from the
1519 receptive field center (Fig. S7a presentation setting 1 or 2 depending on the RGCs
1520 spatial position relative to the presentation area). For this, we first calculated the
1521 spatial receptive field center (see heading *RGCs characterization*; receptive field
1522 center was defined as the 1-σ ellipse of the 2D Gaussian fit) for each RGC. We then
1523 calculated the Euclidean distance between the receptive field center and the closest
1524 flash region for both presentation settings 2 and 3. For further analysis, we used the
1525 presentation setting with the shortest distance between a flash region and the

1526 receptive field center. In case this flash region is perfectly centered over the receptive
1527 field center, saccades will be excluded from a region of at most 300 μm diameter
1528 centered over the receptive field center. In most cases, the RGCs were indeed well
1529 centered within a flash region since the center of each region coincided with electrodes
1530 in the low density MEA. Nonetheless, for each RGC we calculated the intersection
1531 between its receptive field area and saccade regions in pixels (where each pixel
1532 corresponded to 3.75 μm on the retinal surface). RGCs for which more than 15% of
1533 their receptive field area intersected with the saccade regions were excluded from
1534 further analysis. In the end, a total of 51 RGCs (32 ON; 38 OFF) from 4 retinae
1535 recordings were used for further analysis of saccadic suppression.

1536 ***Luminance step paradigm***

1537 To quantify retinal "saccadic suppression" with the luminance-step paradigm (Fig. 4,
1538 Fig. 7), we used the same analyses and statistical procedures to those described
1539 above for the saccade (texture displacement) paradigm. The only difference was that
1540 instead of 39 successive sequences in a trial, we now had either 56 or 156 successive
1541 sequences (or 20 in case of macaque retina experiment), spanning a contrast range
1542 of ±0.03 to ±0.5 Michelson contrast. Similar to the texture displacement paradigms,
1543 the modulation index was based on responses to brief probe flashes (~33 ms flash
1544 duration), and it could therefore only be computed for cells that did respond to these
1545 flash stimuli (N = 366 of 668 spike sorted RGCs from 4 mouse retinae; N = 15 of 57
1546 spike sorted RGCs from a macaque retina). The modulation index for ON RGCs
1547 (N = 259 mouse RGCs; N = 13 macaque RGCs) was calculated from responses to
1548 bright probe flashes, and that for OFF RGCs (N = 107 mouse RGCs; N = 2 macaque
1549 RGCs) was calculated from responses to dark flashes. A subset of data from all mouse
1550 RGCs was presented previously (Idrees et al., 2020). Here, we perform novel analyses
1551 on the complete dataset of the same RGCs.

1552 In a subset of luminance step experiments (2 mouse retinae), we analyzed the effects
1553 of blocking GABAergic and glycinergic inhibition on the modulation of probe flash
1554 responses. 115 ON RGCs showed robust responses to the brief probe flashes with
1555 and without the pharmacological blockers. However, none of the OFF RGCs
1556 responded robustly to the baseline probe flash in the presence of pharmacological

1557 blockers and therefore OFF RGCs were excluded in quantification of saccadic
1558 suppression in the presence of pharmacological blockers.

1559 We also analyzed if the modulation of flash-induced responses depended on the
1560 strength of the response to the preceding luminance step. This analysis was done to
1561 establish whether suppression of flash responses resulted from saccade-induced
1562 saturation or adaptation of ganglion cell responses (Fig. S10). For each RGC, we
1563 calculated an association index which quantified the monotonic relationship between
1564 response to the luminance step and response to subsequent flashes. We first binned
1565 the responses across the 56 or 156 step-flash sequences (Fig. S1b) based on the
1566 contrast induced by the luminance step in each sequence. Bin width was set to 0.025
1567 Michelson contrast. Then, within each bin, we averaged the responses to luminance
1568 steps alone (Fig. S9a, b, top row) and to luminance steps followed by probe flashes.
1569 For each probe flash delay and contrast bin, we quantified the strength of the response
1570 induced by the luminance step preceding each probe flash (we integrated the average
1571 response to the luminance steps followed by probe flashes, up until the response to
1572 that probe flash). The modulation index was calculated as usual. We then calculated
1573 the Spearman's rank correlation coefficient (R) between the modulation index and the
1574 response strength induced by the luminance step, across all the contrast bins. This
1575 can be visualized from the insets in Fig. S9. For ease in interpretation of the results,
1576 we termed the resulting correlation coefficient as the association index. Intuitively, this
1577 association index describes the monotonic relationship between the step response
1578 strength and the strength of suppression and can be interpreted as follows: the larger
1579 the magnitude of the association index, the stronger is the monotonic relation between
1580 the two quantities. A negative value indicates that stronger step responses are
1581 associated with decreasing (more negative) modulation indices (i.e. weaker flash
1582 responses → suppression). A positive value indicates that stronger step responses
1583 are associated with increasing modulation indices (i.e. stronger flash responses →
1584 less suppression or even enhancement). In the example cell of Fig. S9a, the
1585 association index has large negative values for flashes immediately after the positive
1586 luminance step, suggesting that a stronger step response is indeed strongly correlated
1587 with stronger suppression of subsequent flashes. A robust calculation of association
1588 index was only possible for luminance steps that activated the RGC (i.e.

positive-contrast luminance steps in ON RGCs, negative-contrast luminance steps in OFF RGCs). Fig. S10 shows the association index for each RGC and flash time as a function of the cell's modulation index.

In Fig. 6f, we plot modulation index as a function of RGC transiency index (see heading *RGCs characterization* for details on transiency index). The RGCs shown in this sub-figure were a subset of the RGCs analyzed with the luminance step paradigm for which we could also compute a transiency index. The relation between RGC transiency and modulation index for each condition was modeled using a linear regression least-squares fit through the ON and OFF RGC population. To determine if the slope of the resulting line was statistically significant non-zero, we conducted a t-test of this slope.

Data analysis: Cone photoreceptors

Quantifying cone responses

We analyzed data from 11 scan fields recorded from four retinae (2 mice). Each scan field was 128 x 128 pixels, which on the retinal surface was 94 x 94 μm^2, 110 x 110 μm^2, or 132 x 132 μm^2, depending on the zoom factor used. In each scan field, we identified regions of interest (ROIs) as a group of neighboring pixels with correlated fluorescence signals in time. Only ROIs with diameters corresponding to the cone axon terminal diameter (3-7 μm) were considered for further analysis. The output signal of the ROIs (baseline normalized iGluSnFR indicator fluorescence signal), represented the changes in glutamate release at the cone terminals. A total of 931 ROIs were extracted from the 11 scan fields. Identifying the ROIs and extracting their output signal were automated using custom IGOR Pro scripts.

Within a scan field, each ROI was sampled every 256 ms (3.9 Hz sampling rate). This interval was greater than the duration of probe flashes (100 ms) in these experiments. Therefore, the measured signal of many ROIs might not capture the peak response to the probe flashes. A conventional upsampling method, such as interpolation, could also underestimate the peak response in this case. However, since all ROIs (within and across scan fields) experienced the same visual stimulus (Fig. S1c), but were sampled at different points in time, we temporally 'stitched' the output from these ROIs.

1619 The resulting signal had a sampling interval of 2 ms, where the signal in a specific time
1620 bin was computed from ROIs sampled within that specific time window. In this
1621 'stitching' approach, we first baseline-adjusted the output signal of each ROI for a trial
1622 by subtracting the baseline activity (calculated as the average ROI output across 1 s
1623 epoch prior to the first luminance step in the trial). Then, for each 2-ms time bin, we
1624 averaged the response across those ROIs that were sampled within that time window.
1625 This resulted in an output vector of the same duration as the trial but with a sampling
1626 interval of 2 ms. The output vector was empty for time bins where no ROIs were
1627 sampled. We therefore convolved this output vector with a moving average filter of
1628 size 80 ms to fill in the empty time bins and to also smooth out the stitching boundaries
1629 (boundaries between time bins filled with output from different ROIs). This method
1630 gave a much better temporal resolution than conventional upsampling techniques,
1631 robustly capturing the peak for the 100 ms duration probe flash, as shown in Fig. 5b.

1632 The 'stitched' signal was then cut into snippets that captured the relevant responses
1633 to our stimulus (e.g. step followed by probe flash). For this, we used stimulation trigger
1634 signals that marked the presentation of each luminance step. Each snippet was then
1635 baseline-adjusted by subtracting the average response over 800 ms prior to the
1636 luminance step. This way the output signal was 0 prior to a luminance step (Fig. 5).
1637 The output was averaged across the three repetitions of the trial. The normalized and
1638 averaged snippets represented the cone response to a particular stimulus sequence
1639 (for example Fig. 5a) and were used to fit parameters for model cones as described in
1640 the next section.

1641 ***RGC model***

1642 To describe cone responses (Fig. 5) and RGC responses (Fig. 7) to our luminance
1643 step paradigm, we used a phenomenological model of the retina, previously published
1644 in Drinnenberg et al., 2018.

1645 The original model related light intensity to retinal ganglion spiking activity by three
1646 layers of processing: first, the "light stimulus" was passed to model cone
1647 photoreceptors. Their activity was modulated by negative feedback from model
1648 horizontal cells. Second, the output of the model cones was passed to six inner retina
1649 pathways describing retinal processing by three different ON and three different OFF

bipolar cells (fast, intermediate and slow pathways). Third, the output from the model pathways were then fed into model RGCs to yield RGC spiking activity. This cascade modeled RGC spiking in response to the “light stimulus” passed to model cones.

The cone responses were describes as

$$r(t) = \frac{\alpha_c y(t)}{[1 + \beta_c z(t)]} - h(t) \qquad \text{(Equation 2)}$$

where

$$r(t) = V(t) - V_{dark} \qquad \text{(Equation 3)}$$

$V(t)$ and V_{dark} were the instantaneous and dark membrane potentials of the cone, respectively, $h(t)$ was the feedback signal from the horizontal cell, and α_c and β_c were numerical factors. The time-varying functions $y(t)$ and $z(t)$, were related to light input through linear convolution, as

$$y(t) = \int_{-\infty}^{t} dt' K_y(t - t') I(t') \qquad \text{(Equation 4)}$$

$$z(t) = \int_{-\infty}^{t} dt' K_z(t - t') I(t') \qquad \text{(Equation 5)}$$

where $I(t)$ was the incident light intensity (or, more precisely, R*/s). The kernels describing the cone response were given by

$$K_y(t) = \frac{t}{\tau_y} \frac{e^{-\frac{t}{\tau_y}}}{\tau_y} \qquad \text{(Equation 6)}$$

and

$$K_z(t) = \gamma K_y(t) + (1-\gamma)\frac{t}{\tau_z}\frac{e^{-\frac{t}{\tau_z}}}{\tau_z} \qquad \text{(Equation 7)}$$

where τ_z was larger than τ_y, and $0 \leq \gamma \leq 1$ ensured proper normalization. Note that $\int_0^\infty dt' K_y(t-t') = 1$ for all filters. The response of the horizontal cell was described by

$$h(t) = \alpha_c \int_{-\infty}^{t} dt' K_h(t-t') r(t') \qquad \text{(Equation 8)}$$

with

$$K_h(t) = \frac{t}{\tau_h}\frac{e^{-\frac{t}{\tau_h}}}{\tau_h} \qquad \text{(Equation 9)}$$

Here, instead of the cone model parameters used in the published model (Drinnenberg et al., 2018), we re-fitted the parameters of the model cone to reflect our measured data of cone output (Fig. 5b) which yielded faithful fits (Fig. S12) to our experimentally measured cone responses.

All the parameters of the outer retina component (Equations 12 and 13) of our circuit model were fit once to cone responses (Fig. 5b, S11) and then kept unchanged for all simulations reported in Fig. 6. The fitted values are given as follows (original values (Drinnenberg et al., 2018) are reported in brackets):

$\alpha_c = -3.342 * 10^{-5}\ (-9.602 * 10^{-6}), \beta_c = -1.273 * 10^{-6}\ (-1.148 * 10^{-5}), \gamma = 0.842\ (0.764), \alpha_h = 0.016\ (0.177), \tau_y = 48.98\, ms\ (50.64), \tau_z = 200ms\ (576.9), \tau_h = 1.232 * 10^3 ms\ (371)$.

The small α_h suggested that in our cone recordings, horizontal cell feedback had only a minor effect.

In (Drinnenberg et al., 2018) three different retinal pathways were modelled according to

$$b_{p,k}(t) = \left\lfloor -1^k \left(\int_{-\infty}^{t} dt' K_p(t-t') V(t') - \theta_{p,k} \right) \right\rfloor \qquad \text{(Equation 10)}$$

where $p = 1,2,3$ labeled the pathway based on its response properties (1 = fast, 2 = intermediate, 3 = slow), $k = 0$ for OFF pathways and $k = 1$ for ON pathways.

$$\lfloor x \rfloor = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \qquad \text{(Equation 11)}$$

was a thresholding non-linearity, and $\theta_{p,k}$ acted as a threshold.

The main difference between the pathways was the temporal characteristics of the filters K_p. In the current study we wanted to smoothly vary the transiency of the model ganglion cells. To this end, we based our bipolar pathway on the fast bipolar pathway $(p = 1)$ and modified its temporal characteristics to make it less transient. K_1 represented a high-pass filter which took the derivative of the cone potential on the order of 1 ms. We obtained K_1 by convolving the high-pass filter of the form

$$\mathrm{G}(t) = \sin\left(\frac{\pi t}{\mu}\right) \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}, with\ \mu = 3\ ms, \sigma = 1\ ms \qquad \text{(Equation 12)}$$

with an exponential function.

$$K_1(t) = \int_{-\infty}^{t} dt' \left(e^{-\frac{t-t'}{\tau_d}} \right) G(t') \qquad \text{(Equation 13)}$$

Higher values of the time constant, τ_d, of the exponential function decreased the pathway transiency. We set $\tau_d = 0.5$ as the default transient pathway (Fig. 6a-d). The 'transiency parameter' shown in In Fig. 6e was obtained by normalizing $15 \geq \tau_d \geq 0.5$ in the range 0 to 1, with 0 being less transient. In the original model (Drinnenberg et al., 2018), $K_1(t) = G(t)$.

The threshold $\theta_{1,k}$ was set to $-1^k \cdot 0.1$.

$b_{1,k}(t)$, the output of fast inner retina models was used as the input to the model RGCs used in this study. The spiking rate of the model RGC was obtained as the thresholded input and a temporally coarse version of the input's derivative,

1715 $$R_{1,k}(t) = \left\lfloor (1-\alpha) b_{1,k}(t) + \alpha \left(\int_{-\infty}^{t} dt'\, K(t-t')\, b_{1,k}(t') \right) - \theta \right\rfloor \qquad \text{(Equation 14)}$$

1716

1717 where $K(t)$ was a biphasic filter similar in its form to $G(t)$. The threshold, θ, was a
1718 multiple of the peak response to any given input. We used the same parameters for
1719 the inner retina component as the ones described in the published model (Drinnenberg
1720 et al., 2018):

1721 (i) $Transient\ OFF,\ R_{1,0}(t)\colon \alpha = 0,\ \theta = 0;$

1722 (ii) $Transient\ ON, R_{1,1}(t)\colon \alpha = 0,\ \theta = 0;$

1723 All simulations were computed with a 1 ms sampling interval.

1724 Using the above cascade, we calculated model RGC's spike rate in response to the
1725 recorded cone output when subjected to "luminance steps alone" and "luminance
1726 steps followed by probe flash" stimuli (Fig. 6a, b). Similar to real RGC analysis, we
1727 calculated the "flash-induced responses" (Fig. 6c) by subtracting "response to
1728 luminance steps alone" (Fig. 6a) from "response to luminance steps followed by probe
1729 flash" (Fig. 6b). We then calculated the modulation index also in the same way as for
1730 the real RGCs: $(r_d - r_b)/(r_d + r_b)$ where r_d was the peak spiking rate of the flash-induced
1731 response for the flash presented with delay *d* from the luminance step, and r_b was the
1732 peak firing rate for the baseline flash-induced response (flash at 2000 ms).

1733 As a control, we replaced the model cone responses with the experimentally acquired
1734 cone responses (Fig. 5) thus forming a hybrid model. Before passing the cone
1735 response to the downstream model pathways, we passed it through a low pass filter
1736 to further smooth the fluctuations at the stitching boundaries in order to avoid
1737 discontinuities in calculation of its temporal derivative. For this smoothing, we
1738 convolved the cone output with a moving average filter of size 40 ms. Here, the model
1739 ON RGC responses (Fig. S12a, b, columns 1-2) were calculated using the cone
1740 responses to bright probe flashes (Fig. 5a, b, columns 1-2), and OFF RGC responses
1741 (Fig. S12a, b, columns 3-4) were calculated using cone responses to dark probe

flashes (Fig. 5a, b, columns 3-4). The resulting hybrid model RGC responses (Fig. S12) were consistent with the pure model responses shown in Fig. 6.

Data analysis: Human psychophysics

We analyzed eye movements in all trials and detected saccades using established methods (Chen and Hafed, 2013). We excluded from analysis trials in which a saccade or microsaccade happened anywhere in the interval from 200 ms before to 50 ms after a probe flash. At each flash time, we calculated the proportion of correct trials to obtain time courses of this perceptual measure. We obtained time course curves for each subject individually, and then averaged it across trials and different contrasts of the luminance steps. Reduced proportion of correct trials at any flash time indicated perceptual saccadic suppression. A subset of data from 4 of the 5 subjects was used in our previous study (Idrees et al., 2020). Here we perform novel analyses of the complete dataset.

We applied a two-tailed Wilcoxon rank-sum test to determine if the suppression after luminance steps differed across bright and dark probe flashes.

All data analyses were performed in MATLAB (The MathWorks Inc).

Data and software availability

Requests for data should be directed to and will be fulfilled by the Lead Contact, Thomas A. Münch (thomas.muench@gmail.com).

1761 ## References

1762 Amthor, F.R., Tootle, J.S., and Gawne, T.J. (2005). Retinal ganglion cell coding in
1763 simulated active vision. Vis. Neurosci. *22*, 789–806.

1764 Appleby, T.R., and Manookin, M.B. (2019). Neural sensitization improves encoding
1765 fidelity in the primate retina. Nat. Commun. *10*, 1–15.

1766 Baccus, S.A., Olveczky, B.P., Manu, M., and Meister, M. (2008). A Retinal Circuit
1767 That Computes Object Motion. J. Neurosci. *28*, 6807–6817.

1768 Barlow, H.B., Derrington, A.M., Harris, L.R., and Lennie, P. (1977). The effects of
1769 remote retinal stimulation on the responses of cat retinal ganglion cells. J. Physiol.
1770 *269*, 177–194.

1771 Baylor, D.A., and Hodgkin, A.L. (1974). Changes in time scale and sensitivity in turtle
1772 photoreceptors. J. Physiol. *242*, 729–758.

1773 Beeler, G.W. (1967). Visual threshold changes resulting from spontaneous saccadic
1774 eye movements. Vision Res. *7*, 769–775.

1775 Berry, M., Brivanlou, I., Jordan, T., and Meister, M. (1999). Anticipation of moving
1776 stimuli by the retina. Nature *398*.

1777 Breitmeyer, B. (2007). Visual masking: Past accomplishments, present status, future
1778 developments. Adv. Cogn. Psychol. *3*, 9–20.

1779 Bremmer, F., Kubischik, M., Hoffmann, K.P., and Krekelberg, B. (2009). Neural
1780 dynamics of saccadic suppression. J. Neurosci. *29*, 12374–12383.

1781 Brooks, B., and Holden, A.L. (1973). Suppression of visual signals by rapid image
1782 displacement in the pigeon retina: A possible mechanism for “saccadic” suppression.
1783 Vision Res. *13*, 1387–1390.

1784 Chen, C.-Y., and Hafed, Z.M. (2013). Postmicrosaccadic enhancement of slow eye
1785 movements. J. Neurosci. *33*, 5375–5386.

1786 Chen, C.-Y., and Hafed, Z.M. (2017). A neural locus for spatial-frequency specific
1787 saccadic suppression in visual-motor neurons of the primate superior colliculus. J.

1788 Neurophysiol. *117*, 1657–1673.

1789 Chen, E.Y., Marre, O., Fisher, C., Schwartz, G., Levy, J., da Silviera, R.A., and
1790 Berry, M.J. (2013). Alert response to motion onset in the retina. J. Neurosci. *33*,
1791 120–132.

1792 Chichilnisky, E.J., and Baylor, D.A. (1999). Receptive-field microstructure of blue-
1793 yellow ganglion cells in primate retina. Nat. Neurosci. *2*, 889–893.

1794 Chichilnisky, E.J., and Kalmar, R.S. (2002). Functional asymmetries in ON and OFF
1795 ganglion cells of primate retina. J. Neurosci. *22*, 2737–2747.

1796 Clark, D.A., Benichou, R., Meister, M., and Azeredo da Silveira, R. (2013).
1797 Dynamical Adaptation in Photoreceptors. PLoS Comput. Biol. *9*.

1798 Dalkara, D., Byrne, L.C., Klimczak, R.R., Visel, M., Yin, L., Merigan, W.H., Flannery,
1799 J.G., and Schaffer, D. V. (2013). In vivo-directed evolution of a new adeno-
1800 associated virus for therapeutic outer retinal gene delivery from the vitreous. Sci.
1801 Transl. Med. *5*.

1802 Diamond, M.R., Ross, J., and Morrone, M.C. (2000). Extraretinal control of saccadic
1803 suppression. J. Neurosci. *20*, 3449–3455.

1804 Diggelmann, R., Fiscella, M., Hierlemann, A., and Franke, F. (2018). Automatic spike
1805 sorting for high-density microelectrode arrays. J. Neurophysiol. *120*, 3155–3171.

1806 Drinnenberg, A., Franke, F., Morikawa, R.K., Jüttner, J., Hillier, D., Hantz, P.,
1807 Hierlemann, A., Azeredo da Silveira, R., and Roska, B. (2018). How Diverse Retinal
1808 Functions Arise from Feedback at the First Visual Synapse. Neuron 1–18.

1809 Duffy, F.H., and Lombroso, C.T. (1968). Electrophysiological evidence for visual
1810 suppression prior to the onset of a voluntary saccadic eye movement. Nature *218*,
1811 1074–1075.

1812 Enroth-Cugell, C., and Jakiela, H.G. (1980). Suppression of cat retinal ganglion cell
1813 responses by moving patterns. J. Physiol. *302*, 49–72.

1814 Euler, T., Hausselt, S.E., Margolis, D.J., Breuninger, T., Castell, X., Detwiler, P.B.,
1815 and Denk, W. (2009). Eyecup scope—optical recordings of light stimulus-evoked

1816 fluorescence signals in the retina. Pflügers Arch. - Eur. J. Physiol. *457*, 1393–1414.

1817 Euler, T., Franke, K., and Baden, T. (2019). Studying a light sensor with light:
1818 Multiphoton imaging in the retina. Neuromethods *148*, 225–250.

1819 Farrow, K., Teixeira, M., Szikra, T., Viney, T.J., Balint, K., Yonehara, K., and Roska,
1820 B. (2013). Ambient illumination toggles a neuronal circuit switch in the retina and
1821 visual perception at cone threshold. Neuron *78*, 325–338.

1822 Field, G.D., Sher, A., Gauthier, J.L., Greschner, M., Shlens, J., Litke, A.M., and
1823 Chichilnisky, E.J. (2007). Spatial properties and functional organization of small
1824 bistratified ganglion cells in primate retina. J. Neurosci. *27*, 13261–13272.

1825 Frey, U., Egert, U., Heer, F., Hafizovic, S., and Hierlemann, A. (2009).
1826 Microelectronic system for high-resolution mapping of extracellular electric fields
1827 applied to brain slices. Biosens. Bioelectron. *24*, 2191–2198.

1828 Fried, S., Münch, T., and Werblin, F. (2002). Mechanisms and circuitry underlying
1829 directional selectivity in the retina. Nature 411–414.

1830 Garvert, M.M., and Gollisch, T. (2013). Local and global contrast adaptation in retinal
1831 ganglion cells. Neuron *77*, 915–928.

1832 Geffen, M.N., De Vries, S.E.J., and Meister, M. (2007). Retinal ganglion cells can
1833 rapidly change polarity from off to on. PLoS Biol. *5*, 0640–0650.

1834 Gollisch, T. (2013). Features and functions of nonlinear spatial integration by retinal
1835 ganglion cells. J. Physiol. Paris *107*, 338–348.

1836 Hafed, Z.M. (2013). Alteration of Visual Perception prior to Microsaccades. Neuron
1837 *77*, 775–786.

1838 Hafed, Z.M., and Krauzlis, R.J. (2010). Microsaccadic suppression of visual bursts in
1839 the primate superior colliculus. J. Neurosci. *30*, 9542–9547.

1840 Ibbotson, M.R., Crowder, N.A., Cloherty, S.L., Price, N.S.C., and Mustari, M.J.
1841 (2008). Saccadic modulation of neural responses: possible roles in saccadic
1842 suppression, enhancement, and time compression. J. Neurosci. *28*, 10952–10960.

1843 Idrees, S., Franke, F., Diggelmann, R., Hierlemann, A., and Münch, T.A. (2016).
1844 UnitBrowser - A tool to evaluate and post-process units sorted by automatic spike
1845 sorting algorithms. Front. Neurosci. *10*.

1846 Idrees, S., Baumann, M.P., Franke, F., Münch, T.A., and Hafed, Z.M. (2020).
1847 Perceptual saccadic suppression starts in the retina. Nat. Commun. *11*, 1977.

1848 Itokazu, T., Hasegawa, M., Kimura, R., Osaki, H., Albrecht, U.-R., Sohya, K.,
1849 Chakrabarti, S., Itoh, H., Ito, T., Sato, T.K., et al. (2018). Streamlined sensory motor
1850 communication through cortical reciprocal connectivity in a visually guided eye
1851 movement task. Nat. Commun. *9*, 338.

1852 Kemmler, R., Schultz, K., Dedek, K., Euler, T., and Schubert, T. (2014). Differential
1853 regulation of cone calcium signals by different horizontal cell feedback mechanisms
1854 in the mouse retina. J. Neurosci. *34*, 11826–11843.

1855 Khabou, H., Desrosiers, M., Winckler, C., Fouquet, S., Auregan, G., Bemelmans,
1856 A.P., Sahel, J.A., and Dalkara, D. (2016). Insight into the mechanisms of enhanced
1857 retinal transduction by the engineered AAV2 capsid variant -7m8. Biotechnol.
1858 Bioeng. *113*, 2712–2724.

1859 Kleiser, R., Seitz, R.J., and Krekelberg, B. (2004). Neural correlates of saccadic
1860 suppression in humans. Curr. Biol. *14*, 386–390.

1861 Krekelberg, B. (2010). Saccadic suppression. Curr. Biol. *20*, R228–R229.

1862 Krishnamoorthy, V., Weick, M., and Gollisch, T. (2017). Sensitivity to image
1863 recurrence across eye-movement-like image transitions through local serial inhibition
1864 in the retina. Elife e22431.

1865 Krueger, J., and Fischer, B. (1973). Strong periphery effect in cat retinal ganglion
1866 cells. Excitatory responses in ON- and OFF-center neurones to single grid
1867 displacements. Exp. Brain Res. *18*, 316–318.

1868 Li, P.H., Field, G.D., Greschner, M., Ahn, D., Gunning, D.E., Mathieson, K., Sher, A.,
1869 Litke, A.M., and Chichilnisky, E.J. (2014). Retinal Representation of the Elementary
1870 Visual Signal. Neuron *81*, 130–139.

1871 Lin, B., and Masland, R.H. (2006). Populations of wide-field amacrine cells in the
1872 mouse retina. J. Comp. Neurol. *499*, 797–809.

1873 Litke, A.M., Bezayiff, N., Chichilnisky, E.J., Cunningham, W., Dabrowski, W., Grillo,
1874 A. a., Grivich, M., Grybos, P., Hottowy, P., Kachiguine, S., et al. (2004). What does
1875 the eye tell the brain?: Development of a system for the large-scale recording of
1876 retinal output activity. IEEE Trans. Nucl. Sci. *51*, 1434–1440.

1877 Marvin, J.S., Borghuis, B.G., Tian, L., Cichon, J., Harnett, M.T., Akerboom, J.,
1878 Gordus, A., Renninger, S.L., Chen, T.W., Bargmann, C.I., et al. (2013). An optimized
1879 fluorescent probe for visualizing glutamate neurotransmission. Nat. Methods *10*,
1880 162–170.

1881 Matin, E. (1974). Saccadic suppression: a review and an analysis. Psychol. Bull. *81*,
1882 899–917.

1883 Mayo, J.P., and Sommer, M.A. (2008). Neuronal Adaptation Caused by Sequential
1884 Visual Stimulation in the Frontal Eye Field. J. Neurophysiol. *100*, 1923–1935.

1885 Müller, J., Ballini, M., Livi, P., Chen, Y., Radivojevic, M., Shadmani, A., Viswam, V.,
1886 Jones, I.L., Fiscella, M., Diggelmann, R., et al. (2015). High-resolution CMOS MEA
1887 platform to study neurons at subcellular, cellular, and network levels. Lab Chip *15*,
1888 2767–2780.

1889 Münch, T.A., da Silveira, R.A., Siegert, S., Viney, T.J., Awatramani, G.B., and
1890 Roska, B. (2009). Approach sensitivity in the retina processed by a multifunctional
1891 neural circuit. Nat. Neurosci. *12*, 1308–1316.

1892 Noda, H., and Adey, W.R. (1974). Excitability Changes in Cat Lateral Geniculate
1893 Cells during Saccadic Eye Movements. Science (80-.). *183*, 543–545.

1894 Ölveczky, B.P., Baccus, S.A., and Meister, M. (2003). Segregation of object and
1895 background motion in the retina. Nature *423*, 401–408.

1896 Passaglia, C.L., Freeman, D.K., and Troy, J.B. (2009). Effects of remote stimulation
1897 on the modulated activity of cat retinal ganglion cells. J. Neurosci. *29*, 2467–2476.

1898 Reinhard, K., Tikidji-Hamburyan, A., Seitter, H., Idrees, S., Mutter, M., Benkner, B.,

1899 and Münch, T.A. (2014). Step-By-Step instructions for retina recordings with
1900 perforated multi electrode arrays. PLoS One *9*, e106148.

1901 Roska, B., and Werblin, F. (2003). Rapid global shifts in natural scenes block spiking
1902 in specific ganglion cell types. Nat. Neurosci. *6*, 600–608.

1903 Ross, J., Morrone, M.C., Goldberg, M.E., and Burr, D.C. (2001). Changes in visual
1904 perception at the time of saccades. Trends Neurosci. *24*, 113–121.

1905 Sakatani, T., and Isa, T. (2007). Quantitative analysis of spontaneous saccade-like
1906 rapid eye movements in C57BL/6 mice. Neurosci. Res. *58*, 324–331.

1907 Schiller, P.H. (1968). Single unit analysis of backward visual masking and
1908 metacontrast in the cat lateral geniculate nucleus. Vision Res. *8*, 855–866.

1909 Schreyer, H.M., and Gollisch, T. (2020). Nonlinearities in retinal bipolar cells shape
1910 the encoding of artificial and natural stimuli. BioRxiv 1–40.

1911 Schütz, A., Braun, D., and Gegenfurtner, K. (2011). Eye movements and perception:
1912 A selective review. J. Vis. *11*, 1–30.

1913 Schwartz, G.W., Okawa, H., Dunn, F. a, Morgan, J.L., Kerschensteiner, D., Wong,
1914 R.O., and Rieke, F. (2012). The spatial structure of a nonlinear receptive field. Nat.
1915 Neurosci. *15*, 1572–1580.

1916 Shapley, R., and Enroth-Cugell, C. (1984). Chapter 9 Visual adaptation and retinal
1917 gain controls. Prog. Retin. Res. *3*, 263–346.

1918 Szatko, K.P., Korympidou, M.M., Ran, Y., Berens, P., Dalkara, D., Schubert, T.,
1919 Euler, T., and Franke, K. (2020). Neural circuits in the mouse retina support color
1920 vision in the upper visual field. Nat. Commun. *11*, 3481.

1921 Tatler, B.W., Wade, N.J., Kwan, H., Findlay, J.M., and Velichkovsky, B.M. (2010).
1922 Yarbus, Eye Movements, and Vision. Iperception. *1*, 7–27.

1923 Tikidji-Hamburyan, A., Reinhard, K., Seitter, H., Hovhannisyan, A., Procyk, C.A.,
1924 Allen, A.E., Schenk, M., Lucas, R.J., and Münch, T.A. (2015). Retinal output
1925 changes qualitatively with every change in ambient illuminance. Nat. Neurosci. *18*,
1926 66–74.

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

Wei, W. (2018). Neural mechanisms of motion processing in the mammalian retina. Annu. Rev. Vis. Sci. *4*, 165–192.

Werblin, F. (2010). Six different roles for crossover inhibition in the retina: correcting the nonlinearities of synaptic transmission. Vis. Neurosci. *27*, 1–8.

Wurtz, R.H. (2008). Neuronal mechanisms of visual stability. Vision Res. *48*, 2070–2089.

Yarbus, A.L. (1967). Eye Movements and Vision (Boston, MA: Springer US).

Zaghloul, K.A., Manookin, M.B., Borghuis, B.G., Boahen, K., and Demb, J.B. (2007). Functional circuitry for peripheral suppression in Mammalian Y-type retinal ganglion cells. J Neurophysiol *97*, 4327–4340.

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

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