Atypical cortical feedback underlies failure to process contextual information in the superior colliculus of Scn2a$^{+/-}$ autism model mice

Authors: Leiron Ferrarese*[1], Hiroki Asari*[1]

Affiliations:
1. Epigenetics and Neurobiology Unit, EMBL Rome, European Molecular Biology Laboratory, Monterotondo, 00015, Italy

***Corresponding authors:** Leiron Ferrarese <leiron.ferrarese@embl.it>, Hiroki Asari <asari@embl.it>

Acknowledgments: This work was supported by research grants from EMBL (H.A.) and SFARI (#875447; H.A.); and EMBL Interdisciplinary, International and Intersectorial Postdoctoral Fellowship (EI3POD; L.F.). The EMBL Gene Editing and Embryology Facility is acknowledged for the mouse line rederivation; EMBL Genetic and Viral Engineering Facility for virus production; EMBL IT Support for provision of computer and data storage servers; and the LAR facility for taking care of animals. We thank all the Asari lab members as well as Cornelius Gross, Eleonora Satta and Michael Lombardo for many useful discussions.

Author contributions: L.F. and H.A. designed the study; L.F performed experiments and analyzed the results; L.F. and H.A. wrote the manuscript.

Competing interests: The authors declare no competing financial interests.

Key words: autism spectrum disorder; superior colliculus; implicit visual learning; Bayesian inference; hierarchical Gaussian filter.

Acronyms: AAV, adeno-associated virus; ANOVA, analysis of variance; AP, anterior-posterior; ASD, autism spectrum disorder; CNMF, constrained non-negative matrix factorization; CNO, clozapine N-oxide; DAPI, 4',6-diamidino-2-phenylindole; DREADD, designer receptors exclusively activated by designer drugs; DV, dorsal-ventral; ES, early stable; FDR, false discovery rate; GECI, genetically-encoded calcium indicator; hM4Di, modified human M4 muscarinic acetylcholine receptor for inhibitory DREADD; HGF, hierarchical Gaussian filter; IR, infrared; ISI,: interstimulus interval; L5, layer 5; LS, late stable; ML, medial-lateral; PFA, paraformaldehyde; PBS, phosphate buffer saline; RGC, retinal ganglion cell; SC, superior colliculus; s.e.m., standard error of the mean; Scn2a, sodium voltage-gated channel alpha subunit 2; V1, primary visual cortex; VL, volatile; WT, wild-type.

Abstract

Atypical sensory integration and contextual learning are common symptoms in autism spectrum disorder (ASD), but how sensory circuits are affected remains elusive. Here we focused on the early visual information processing, and performed *in vivo* two-photon calcium imaging and pupillometry of mice engaged in an implicit learning task in stable and volatile visual contexts. Wild-type (WT) mice showed stimulus-specific contextual modulation of the visual responses in the superior colliculus (SC) and pupil dynamics, whereas Scn2a-haploinsufficient ASD-model mice exhibited abnormal modulation patterns. In both genotypes, feedforward inputs from the retina to SC demonstrated no such contextual modulation. In contrast, feedback inputs from the primary visual cortex (V1) showed modulation patterns similar to those of SC cells in WT mice, but no modulation in Scn2a$^{+/-}$ mice. Furthermore, chemogenetic perturbation revealed that this top-down signaling from V1 to SC mediated the observed contextual modulation both at the neurophysiological and behavioral levels. These results suggest that the corticotectal input is critical for contextual sensory integration in SC, and its anomaly underlies atypical sensory learning in ASD.

Introduction

Bayesian inference provides a theoretical framework for understanding how the brain combines prior knowledge with incoming sensory information to shape perceptual experiences and guide behavior (Friston, 2009, 2012). In particular, a hierarchical Bayesian learning model has been successfully applied to explain atypical behaviors observed in individuals with autism spectrum disorder (ASD; Pellicano and Burr, 2012; Chrysaitis and Seriès, 2023): e.g., how overlearning and overestimation of the environmental volatility result in a reduced response to unexpected events in ASD (Lawson et al., 2017). However, the current experimental evidence supporting this

hypothesis is limited in both sample sizes and computational models (Chrysaitis and Seriès, 2023).

How do neuronal circuit functions underlie such differences in learning behavior between individuals with and without ASD? The superior colliculus (SC) is a promising target to address this question. It is a midbrain hub for sensorimotor transformation implicated in abnormal visual processing in ASD (Johnson et al., 2016; Jure, 2022). For instance, gaze management and attentional focus are two major functions of SC that are often altered in ASD (Lomber et al., 2002; Kato et al., 2011). SC also plays a role in the initial automatic bias towards global visual processing, which is suggested to be affected in ASD subjects as they typically focus too much on details (Jure, 2022). Moreover, extensive connections to many other brain regions (May, 2006) support that SC interacts with emotional and motor networks, allowing animals to perceive and respond appropriately to sensory stimuli with attentional prioritization, such as the innate looming avoidance responses in mice (Dean et al., 1989; Yilmaz and Meister, 2013; Wei et al., 2015; De Franceschi et al, 2016). The involvement of SC in these processes underscores its relevance to the neurophysiological mechanisms of ASD.

Given the convergence of retinal and cortical inputs, SC is expected to be part of a hierarchical recursive system for visual inference in the brain (Friston, 2012; Jure, 2022). On the one hand, neurons in the superficial layers of SC receive direct visual input from the retina, and play a role in the processing of visual motion and orientation (Inayat et al., 2015; Lomber et al., 2001). On the other hand, superficial SC neurons also receive a prominent feedback signal from the primary visual cortex (V1; Zingg et al., 2017; Benavidez et al., 2021; Wang and Burkhalter, 2013), which is implicated to help integrate sensory and perceptual information. Recent studies have revealed that the V1-to-SC projection provides contextual gain modulation, enhancing the response amplitude of SC neurons to looming stimuli, while also acting as a negative gain modulator for the saliency of complex visual scenes (Zhao, 2014; Ahmadlou et al., 2017).

Furthermore, the corticotectal synapses show nonlinear properties, suggesting their role in regulating sensory-guided innate behavior and modulating attention and learning processes (Zhao, 2014; Ruediger and Scanziani, 2020; Tang and Higley, 2020), although some controversies still remain on the modulatory role of the corticotectal inputs (Ogasawara et al., 1984; Hoffmann and Straschill, 1971; Wang et al., 2010).

As an increasing number of ASD-associated genes have been identified (Sanders et al., 2012; Satterstrom et al., 2020), animal models of ASD have been developed to investigate the underlying neurobiological mechanisms (Patel et al., 2018). One of the well-established ASD mouse models is the haploinsufficiency for *sodium voltage-gated channel alpha subunit 2* (Scn2a$^{+/-}$ mice; Ben-Shalom et al., 2017). The *Scn2a* gene encodes the $Na_V1.2$ sodium channels that help regulate dendritic excitability of mature cortical pyramidal neurons (Planells-Cases et al., 2000), including those forming the corticotectal pathway (Zingg et al., 2017; Spratt et al., 2019). Haploinsufficiency of $Na_V1.2$ reduces the backpropagation of action potentials into dendrites, impairing synaptic plasticity and strength in layer 5 (L5) corticofugal pyramidal neurons (Spratt et al., 2019). Therefore, Scn2a$^{+/-}$ mice are expected to have atypical corticotectal projections, leading to abnormal neuronal modulations within the SC circuitry.

Here we developed an implicit visual learning task, similar to the ones used in human studies (Mathys et al., 2014; Lawson et al., 2017), but for mice to form probabilistic associations of cue-outcome stimulus sequences in stable and volatile environments. We then combined *in vivo* two-photon calcium imaging, chemogenetic perturbation, pupillometry, and computational modelling to characterize both neurophysiological and behavioral phenotypes of mice engaged in this task. Over the course of implicit learning, wild-type (WT) mice exhibited a stimulus-specific context-dependent modulation in SC as well as in the corticotectal input from V1 and the pupil dynamics, but not in the input from the retina. A hierarchical Bayesian learning model indicated that a larger fraction of WT SC neurons carries information about the high-level environmental

volatility rather than the cue-outcome contingency *per se*. In contrast, Scn2a$^{+/-}$ mice demonstrated anomalous contextual modulation in SC, and no modulation in the V1-to-SC projections, retinal signals, and pupillary responses. Specifically, these SC neurons showed no adaptation to an increased outcome predictability in the stable environment, and largely represented the low-level outcome expectations in a way consistent with the observations in ASD subjects from the perspective of the hierarchical Bayesian learning model (Lawson et al., 2017). We further found that chemogenetically inhibiting the visual cortical feedback to SC abolished contextual modulations in both WT and Scn2a$^{+/-}$ mice, both at the physiological and behavioral levels. These results suggest that the corticotectal input plays a critical role in integrating contextual information in SC, and that anomaly in this process underlies atypical sensory learning in ASD.

Results

Implicit visual learning paradigm to test behavioral and physiological anomalies in a mouse model of autism spectrum disorder

Individuals with ASD often struggle to learn about probabilistically aberrant events (Lawson et al., 2017). Understanding the neuronal processes underlying implicit learning and the integration of contextual information will thus provide valuable insights into the neurobiological mechanisms of ASD (Chrysaitis and Seriès, 2023). In this study, we focused on the early visual system of mice to study a) how neuronal circuit computation contributes to implicit learning about visual environmental contexts; and b) how it is altered in a mouse model of ASD (Scn2a$^{+/-}$). We thus began by designing an implicit visual learning paradigm based on Bayesian principles (Mathys, 2011): i.e., probabilistic associative learning of cue-outcome stimulus sequences, much as those experimental paradigms used in human studies (Lawson et al., 2017).

Our implicit learning task was carefully tailored with visual stimuli to elicit robust responses in the mouse SC (Cang et al., 2018), such as static gratings with different orientations (Basso and

May, 2017; De Franceschi and Solomon, 2018) and looming and sweeping stimuli that trigger SC-mediated arousal and innate defensive behavior (Dean et al., 1989; Yilmaz and Meister, 2013; Wei et al., 2015; De Franceschi et al., 2016). Specifically, we used vertical and horizontal gratings as neutral cues; and a combination of sweeping and looming dark spots as aversive outcomes while a gray screen (omitted stimulus) as a non-aversive outcome (**Fig. 1a**). Since the task did not involve any reward or punishment, we exploited innate visual threats in the outcome stimulus to motivate animals to recognize stimulus sequences. To increase unpredictability, we introduced uncertainty by randomizing the type and direction of the visual threat at each presentation (**Fig. 1a**; see Methods for details).

The task followed a Bayesian update scheme modeled by a hierarchical Gaussian filter (HGF; Mathys, 2011; Lawson et al., 2017), consisting of a stable environment followed by a volatile one (**Fig. 1b**). During the stable environment (144 trials in total), the aversive outcome stimuli were primarily preceded by one type of neutral cue stimuli (vertical gratings; 83%). Mice were then expected to form probabilistic associations of these cue-outcome stimulus sequences (cue-outcome contingencies). At a small probability, however, the aversive stimuli followed the other type of cue stimuli (horizontal gratings; 17%). This allowed animals to determine whether an unexpected outcome indicates a deviation from a stable stimulus pattern or a change in the pattern itself. Importantly, the stability and precision of the sensory inputs (likelihood) should remain high in the stable environment, allowing the prior expectation to increase over trials (**Fig. 1c**). As a result, the posterior belief about the environment – or, the predictability given by the product of the likelihood and the prior – should become higher and closer towards the prior in the late-stable epoch (LS; last 72 trials) than in the early-stable epoch (ES; initial 72 trials). In the following volatile epoch (VL, 72 trials in total), we then introduced variations in the probability of presenting aversive stimuli after each neutral stimulus. This challenged the animal's expectations on the stimulus sequences. The environment became less precise, hence lowering likelihood,

while the prior belief remained relatively high and precise because updating the prior is slow in general. Thus, the posterior precision should decrease and shift towards the prior. Taken together, we then expect that defects in the prior formation, as implicated by reduced adaptability in ASD individuals (Turi et al., 2015, 2016; van Boxtel et al et al., 2016), should reveal anomalies in the LS epoch, whereas those in likelihood, arising from sensory hypersensitivity for example (Cannon et al., 2014), should affect the responses in the VL epoch.

Wild-type and Scn2a$^{+/-}$ mice exhibit distinct contextual modulation patterns in the visual responses of the superior colliculus

We began with neurophysiological assays on awake, head-fixed mice expressing genetically-encoded calcium indicators (GECIs) in SC somata delivered via adeno-associated viral vectors (AAVs; see Methods for details; **Fig. 2a,b**). Specifically, we exposed the mice to the implicit visual learning task as described above, and simultaneously performed *in vivo* two-photon calcium imaging of the superficial SC through a cranial window, together with pupillometry for behavioral monitoring (Dombeck et al., 2007; Feinberg and Meister, 2015; Molotkov et al., 2023; **Supplementary Fig. 1a**).

On average we recorded 140±39 cells per animal (mean ± standard deviation; 17 WT and 16 Scn2a$^{+/-}$ mice) and estimated their spiking activity using CaImAn (Pnevmatikakis et al., 2019; $Z(t)$ in ***z***-score, see Methods for details). As expected (Cang et al., 2018; De Franceschi and Solomon, 2018), superficial SC neurons generally responded well to the onset and/or offset of the static gratings presented as cue stimuli (**Supplementary Figs. 1 and 2**). About a half of these cells were responsive to the looming and/or sweeping stimuli presented as aversive outcome (74±28 cells per recording), whereas in most cells we found virtually no response to the omitted stimulus presented as non-aversive outcome (gray screen; **Supplementary Fig. 3**). We considered cells as responsive if the trial average of their visual response dynamics exceeded a predetermined threshold ($\max[z(t)] > 1$, where $z(t) = \langle Z(t) \rangle_{\text{trials}}$; see Methods for details). Due to

retinotopy in the early visual system, a) each aversive stimulus drove cells at different timings in our experimental design; and b) each cell showed distinct response patterns to different aversive stimuli. We thus combined all the observed responses to gain enough statistical power in our analysis.

To examine the SC dynamics during the implicit learning task, we compared the aversive stimulus responses at the population level across the three analysis epochs in our paradigm (ES, LS, and VL; **Fig. 2c-e**). For each responsive cell, we calculated the average response strength during the stimulus presentation period in each epoch ($\overline{z} = \langle z(t) \rangle_t$; e.g., **Fig. 2c**), and subtracted the mean across the epochs for normalization ($\Delta z = \overline{z} - \langle \overline{z} \rangle_{\text{epochs}}$; e.g., **Fig. 2e**). Using this normalized response, Δz, we found that SC neurons in WT animals responded significantly weaker to the aversive stimuli during the LS epoch than the ES or VL epoch (**Fig. 2e**). Such changes in the population response strength were relatively stronger for later components of the aversive stimulus responses than for the early response components (**Supplementary Fig. 4**), while not observed for the cue stimulus responses (**Supplementary Fig. 2**). These suggest that the responses of SC neurons in WT mice are modulated by the visual contexts in two ways. First, when the stimulus pattern is reliably predictable during the stable environment, the responses to the aversive stimuli become selectively and progressively attenuated over trials (adaptation). Second, when the stimulus pattern deviates from an animal's expectation during the volatile environment, the response of SC neurons is reset to the baseline and becomes as high as in the initial ES epoch before adaptation.

In contrast, Scn2a$^{+/-}$ mice displayed different context-dependent modulations in SC during the implicit visual learning task: no adaptation in the stable environment, but increased neuronal responses to the aversive stimuli while decreased responses to the cue stimuli during the volatile environment (**Fig. 2f-h** and **Supplementary Fig. 2**). Changes in the population response strength were relatively larger for the early response components than for the later ones (**Supplementary**

173 **Fig. 4**). These response patterns of SC neurons in Scn2a$^{+/-}$ mice were significantly different from
174 those observed in WT ($p(\text{epoch} \times \text{group}) = 0.0007$, 3-way nested analysis of variance (ANOVA)
175 on $\bar{z}$; see Methods for details). This suggests that, unlike WT animals, SC neurons in Scn2a$^{+/-}$
176 mice fail to selectively attenuate their aversive stimulus responses even when the stimulus pattern
177 is kept predictable. They are, however, still sensitive to increased volatility, and thus their
178 response during the VL epoch diverged from the baseline activity in a stimulus specific manner.
179 This is consistent with the observations on ASD subjects who tend to overestimate environmental
180 volatility in stable environments, and overreact to increased unpredictability in volatile
181 environments (Lawson et al., 2017).

182 Context-dependent modulation is absent in retinal output responses

183 How do SC neurons achieve such stimulus-specific context-dependent modulations? One
184 possibility is that the contextual modulation in SC is inherited from retinal ganglion cells (RGCs).
185 To investigate the contribution of bottom-up RGC input to SC, we performed *in vivo* two-photon
186 calcium imaging of RGC axon terminals in SC (Molotkov et al., 2023) while the animals were
187 engaged in the implicit visual learning task (**Fig. 3a**). On average we recorded 323±94 putative
188 axonal boutons per animal (mean ± standard deviation; 5 WT and 5 Scn2a$^{+/-}$ mice), of which
189 208±87 were responsive to the looming and/or sweeping stimuli ($\max[z(t)] > 0.75$). In both WT and
190 5 Scn2a$^{+/-}$ mice, RGC axon terminals in SC showed consistent responses to the aversive stimuli
191 regardless of the visual context (**Fig. 3b,e**), and we found no significant contextual modulation of
192 across ES, LS, and VL epochs at the population level (**Fig.3c,f**). No contextual modulation was
193 observed in the population responses to the omitted or cue stimuli, either (**Supplementary Figs.
194 2-4**). This excludes the possibility that the contextual modulation in SC arises from the retina.

195 **Corticotectal inputs show similar context-dependent modulation patterns as**
196 **local neurons in the superior colliculus of wild-type mice, but no modulation**
197 **in Scn2a$^{+/-}$ mice**

198 Another possible source for the contextual modulation in SC is the corticotectal projection.
199 Besides sensory input from RGCs, the superficial layer of SC receives substantial feedback
200 signals from L5 pyramidal neurons in V1 (Zingg et al., 2017). These V1 L5 neurons integrate
201 feedforward visual information through their basal dendrites and feedback contextual information
202 via the apical dendrites (Larkum, 2013), and provide output to SC for controlling the response
203 gain (Zhao et al., 2014; Ahmadlou et al., 2017; Tang and Higley, 2020). Cortical pyramidal
204 neurons in Scn2a$^{+/-}$ mice have impaired synaptic functions due to a reduced expression of the
205 $Na_V1.2$ sodium channel (Spratt et al., 2019). This may result in a failure to incorporate visual and
206 contextual information in V1, which in turn may lead to an anomalous modulation of SC as
207 observed in our paradigm (**Fig. 2** and **Supplementary Figs. 1-4**).

208 To examine the role of V1 feedback in the contextual modulation of SC, we first performed
209 *in vivo* two-photon calcium imaging of the axon terminals of V1 projection neurons in SC while
210 animals were engaged in the implicit visual learning task (**Fig. 4a**). On average, we recorded
211 412±104 putative axonal boutons per animal (mean ± standard deviation; 8 WT and 7 Scn2a$^{+/-}$
212 mice), with 167±51 boutons responsive to the looming and/or sweeping stimuli ($\max[z(t)] > 0.75$).
213 In WT mice, the aversive stimulus responses of the V1 axon terminals in SC showed significant
214 context-dependent modulation (**Fig. 4b,c**), especially for later response components
215 (**Supplementary Fig. 4**), resembling the dynamics observed in SC somata. The responses were
216 attenuated during the LS epoch compared to the ES epoch, consistent with adaptation during
217 stable environments, and returned to the baseline levels during the VL epoch, reflecting the
218 sensitivity to increased volatility. Unlike SC cells, however, these V1 axons showed a significant
219 reduction in the cue stimulus responses during the VL epoch (**Supplementary Fig. 2**), and

220 modest responses to the omitted stimulus, demonstrating adaptation during stable environments
221 but no recovery in the volatile environment (**Supplementary Fig. 3**). This highlights the stimulus
222 specificity of the contextual modulation and distinct information processing in different areas of
223 the early visual system.

224 In contrast, V1-to-SC projections in Scn2a$^{+/-}$ mice displayed no context-dependent
225 modulation (**Fig. 4d-f**). Unlike SC cells, these V1 axonal boutons failed to adapt to stable
226 environments or sensitize in volatile environments, and kept stable response levels throughout
227 the learning task regardless of the stimuli (**Supplementary Figs. 2-4**). Haploinsufficiency of
228 $Na_V1.2$ impairs synaptic plasticity and dendritic excitability in L5 corticofugal neurons (Planells-
229 Cases, 2000; Spratt et al., 2019; Zingg et al., 2017), including those projecting from V1 to SC.
230 Thus, the observed anomalies of the V1-to-SC projections in Scn2a$^{+/-}$ mice may arise from the
231 direct cell-intrinsic effects of the SCN2A mutation on those corticotectal neurons, or indirect
232 neuronal circuit effects on the upstream visual cortical processing.

233 Chemogenetically inhibiting corticotectal projection abolishes context-
234 dependent modulation in the superior colliculus

235 To further examine how V1 inputs affect the visual response of SC, we next employed a
236 chemogenetic approach to selectively inhibit the V1 L5 corticotectal neurons using a designer
237 receptor exclusively activated by designer drug (DREADD; Armbruster et al., 2007). Specifically,
238 we achieved selective expression of modified human M4 muscarinic acetylcholine receptors for
239 inhibitory DREADD (hM4Di) in the target neuronal population by injecting Cre-dependent AAV8
240 carrying the hM4Di receptor into V1 and AAV2retro carrying Cre-recombinase into SC (**Fig. 5a**).
241 Histological analyses confirmed that axonal projections were clearly visible in the ipsilateral SC of
242 all mice examined (e.g., **Fig. 5b**; n=7 each for WT and Scn2a$^{+/-}$ mice); and that these projections
243 originated mostly from the V1 L5 area, where labelled cells occupied 31±22% (mean ± standard
244 deviation), mostly in the medial-anterior part. Systemic administration of the DREADD actuator,

245 clozapine N-oxide (CNO), then activated hM4Di, hence effectively inhibited the V1 neurons
246 projecting to SC. As before, AAVs carrying GECIs were also injected into SC for monitoring the
247 visual responses of SC neurons with *in vivo* two-photon microscopy.

248 Chemogenetic inhibition of the projection from V1 to SC abolished the context-dependent
249 modulation of SC responses to the aversive stimuli in both WT and Scn2a$^{+/-}$ mice (**Fig. 5c-f**). After
250 the intraperitoneal injection of CNO, SC neurons did not show adaptation during the LS epoch or
251 recovery/enhancement during the VL epoch (e.g., **Fig. 5c,e**). As a result, the overall response
252 level of SC cells did not significantly change throughout the implicit visual learning task (**Fig. 5d,f**),
253 though the early response components of WT mice increased in the VL epoch, resembling the
254 dynamics of intact Scn2a$^{+/-}$ animals (**Supplementary Fig. 4**), and the cue stimulus responses
255 showed a tendency to decrease over trials after blocking V1 (**Supplementary Fig. 2**). Such
256 response patterns to the aversive stimuli across epochs were significantly different from those in
257 intact animals ($p(\text{epoch} \times \text{group}) = 0.0005$, 3-way nested ANOVA; see Methods for details).
258 Importantly, hM4Di expression in V1 or CNO administration alone had no effects on the contextual
259 modulation of SC cells in WT mice (**Supplementary Fig. 5**). We further confirmed via receptive
260 field analyses that the basic visual response properties of SC cells, RGC axons and V1 axons in
261 SC were overall similar between WT and Scn2a$^{+/-}$ mice (**Supplementary Fig. 6**). Taken together,
262 we suggest that feedback input from V1 plays a key role in modulating the aversive stimulus
263 response of SC neurons in a context-dependent manner, and that anomalies of the V1-to-SC
264 signals likely underlie the impairments of the context-dependent visual processing in the SC of
265 Scn2a$^{+/-}$ mice.

Pupil dynamics reflect context-dependent stimulus pattern recognition in wild-type but not in Scn2a$^{+/-}$ mice

Together with neuronal activities, we collected pupillometric data (**Fig. 6**) as a proxy to monitor cognitive processes of the animals (Loewenfeld, 1958). Here we in particular interpreted pupil dynamics as a reflection of an animal's stimulus pattern recognition (Suzuki et al., 2018; Ganea et al., 2020; Dobbins, 2021). Indeed, all mice exhibited stereotypical pupil dynamics during the implicit visual learning task: i.e., pupil contraction in response to the cue stimuli, and pupil dilation during the outcomes, especially when the aversive stimuli were presented (**Supplementary Figs. 1-3**). This consistent response over time across epochs suggests that they were attentive throughout the experiments.

In WT animals, we found a smaller magnitude of pupil dilation upon aversive stimulus presentations during the volatile environment than the stable environment (**Fig. 6d,h**), while no substantial change in the pupil dynamics was observed during the cue or omitted stimulus presentations (**Supplementary Figs. 2 and 3**). Such stimulus-specific context-dependent modulation of the pupil dynamics was not observed when the feedback signal from V1 to SC was chemogenetically blocked (**Fig. 6f,j**), and the resulting modulation pattern was significantly different from that of the negative controls ($p(\text{epoch} \times \text{group}) = 0.01$, 2-way ANOVA between animals expressing hM4Di with and without CNO, **Supplementary Fig. 5c**; $p(\text{epoch} \times \text{group}) = 0.04$, 2-way ANOVA between CNO-administered animals with and without hM4Di expression, **Supplementary Fig. 5f**).

In contrast, Scn2a$^{+/-}$ animals generally showed a smaller pupil size than WT animals throughout the task (24 Scn2a$^{+/-}$ mice, 0.62 (0.44) mm^2, median (interquartile range) of the median pupil area; 25 WT mice, 0.79 (0.28) mm^2; $p = 0.006$, Mann-Whitney U-test; **Fig. 6a-c**), and displayed no context-dependent modulation (**Fig. 6e,i and Supplementary Figs. 2 and 3**). Chemogenetic inhibition of the corticotectal projection did not affect the pupil dynamics in Scn2a$^{+/-}$

291 mice (**Fig. 6g,k**). The modulation pattern of the pupil dynamics across epochs was significantly
292 different between WT and the other three experimental groups ($p(\text{epoch} \times \text{group}) = 0.04$; 2-way
293 ANOVA; see Methods for details). This supports that the influence of the corticotectal signals
294 extends beyond the neuronal activity in SC, likely playing a role in the contextualization of the
295 visual scene even at the perceptual level; and that Scn2a$^{+/-}$ mice have altered V1-to-SC signaling,
296 resulting in a failure to properly recognize the stimulus patterns during the implicit learning task.

297 Taken together, we found that the modulation of the pupil dynamics (**Fig. 6**) was well linked
298 to that of the neuronal dynamics in SC and V1-to-SC projection (**Figs. 2-5**): i.e., the greater the
299 pupil dilation was, the weaker the responses of SC cells and V1 input were in our experimental
300 paradigm, much as was demonstrated in previous studies (Wang et al., 2012; Joshi et al., 2016).
301 Most importantly, Scn2a$^{+/-}$ mice showed impaired neurophysiological and behavioral phenotypes,
302 both due to anomalous signaling from V1 to SC. This highlights a crucial role of the top-down
303 corticotectal pathway in forming perceptual associations to recognize and properly react to visual
304 contexts (Tang and Higley, 2020), and a possible relationship between ASD and dysfunctions of
305 V1 L5 projection neurons and their upstream circuits (Marco et al., 2011; Johnson et al., 2016;
306 Jure, 2022).

307 Corticotectal signals in Scn2a$^{+/-}$ mice fail to convey visual contextual
308 information from the hierarchical Bayesian inference perspective

309 Thus far we made population-level analyses on our neurophysiological and behavioral data during
310 the implicit visual learning task, and identified that 1) V1 provides critical modulatory signals to
311 SC; and 2) anomalies of such corticotectal signals likely underlie the absence of stimulus-specific
312 context-dependent modulation in the SC and pupil dynamics of Scn2a$^{+/-}$ mice (**Figs. 2-6**). What
313 do those individual SC neurons and inputs to SC represent? To address this question, we
314 simulated how an ideal Bayesian observer should respond to the sequence of the cue-outcome
315 contingencies in our experimental paradigm (Mathys, 2011; Friston, 2009, 2012), and compared

the dynamics of the underlying model parameters with those of the observed neuronal responses on a trial-to-trial basis to gain insights from the Bayesian inference perspective (**Fig. 7**).

Among many Bayesian inference models (Sohn and Narain, 2021), here we employed an HGF model, similar to the one used to explain the differences between subjects with and without ASD in their behavioral responses to environmental uncertainty (Mathys et al., 2014; Lawson, et al., 2017). In particular, using variational Bayes inversion processes, the model infers hidden states that represent beliefs about the sensory input at three hierarchical levels (**Fig. 7a**; see Methods for details). The lowest level concerns directly on the outcome (aversive versus non-aversive outcomes; X_1); the next level on the outcome probability (cue-outcome contingencies; X_2); and the highest level on the likelihood of environmental changes (stability versus volatility; X_3). Inference at each level i depends on that at a higher level, and involves the following five parameters: prior predictions on the outcome ($\hat{\mu}_i$); variance of prediction belief ($\hat{\sigma}_i$); learning rates (α_i); precision-weighted prediction errors (ε_i); and uncertainties (unc_i). In total, the HGF model has 13 parameters that are dynamically updated on a trial-to-trial basis (as the two are redundant; **Fig.7b**; see Methods for details). Each of these parameters provides an insight into distinct aspects of the Bayesian inference process at different levels of uncertainty associated with the sensory input.

To make a fair comparison between the model and neuronal outputs, we focused only on the trials with aversive outcomes because SC neurons were generally not responsive to the non-aversive stimulus in our paradigm (i.e., the omitted stimulus; see, e.g., **Supplementary Fig. 3**). Thus, after running the simulation with the entire trial sequence, we concatenated the output of the aversive outcome trials alone to generate the trajectories of the HGF model parameters (**Fig. 7b**). To generate corresponding trajectories of the neuronal signals ($\Delta z_k^{\mathrm{res}}$; **Fig. 7c,e**), we normalized the observed responses of individual neurons for each of the six different types of aversive stimuli, and further filtered out slow non-stationary components (see Methods for details).

We then conducted pairwise Pearson's correlation tests with the false discovery rate (FDR) correction between the neuronal and model-parameter trajectories to characterize neuronal representations of the HGF model in the SC of WT and Scn2a$^{+/-}$ mice. The presence of correlation here does not necessarily mean that these neurons implement the model or directly contribute to Bayesian computation. Nevertheless, it offers an interpretable representation of our experimental data from the hierarchical learning viewpoint. The cue stimulus responses were not considered here because they did not show context-dependent modulation (**Supplementary Fig. 2**).

We obtained three conclusions from this analysis. First, most of the HGF parameters were represented by some subsets of SC neurons, evidenced by statistically significant correlations between the two trajectories (**Fig. 7d and Supplementary Table 2**). This was the case with both WT and Scn2a$^{+/-}$ animals, regardless of the chemogenetic inhibition of the corticotectal projection. Nevertheless, SC neurons in WT mice had a distinct representation of the HGF parameters, compared to those in Scn2a$^{+/-}$ animals or when corticotectal projections were chemogenetically blocked. Specifically, WT mice showed a significant polarity bias in the HGF parameter representation, especially at level 3, whereas Scn2a$^{+/-}$ mice showed a tendency towards a positive correlation bias at lower level parameters. Blocking V1 inputs to SC generally attenuated the bias of the parameter representations in both WT and Scn2a$^{+/-}$ animals. This trend was also clear when we computed the cumulative sum of the number of cells correlated with the HGF model parameters at each level. Neuronal trajectories of the SC in WT mice were on average more correlated with level 3 parameter dynamics than those in Scn2a$^{+/-}$ mice or after blocking V1-to-SC projections (q=0.03 and 0.003, respectively; Chi-squared test with FDR correction), indicating a misrepresentation of environmental volatility in the SC of Scn2a$^{+/-}$ mice.

Second, V1-to-SC projections in WT mice showed significant correlations with a broad range of HGF parameters across hierarchical levels, while those in Scn2a$^{+/-}$ mice overall showed little correlations (**Fig. 7e,f and Supplementary Table 2**). In particular, the V1 inputs to SC in WT

366 mice exhibited a significant negative bias for level 3 uncertainties, much as local neurons in SC
367 did. Unlike SC, however, neuronal trajectories of those V1 boutons were more correlated with
368 lower-level parameters (q<0.001, Chi-squared test with FDR correction). In contrast, the
369 corticotectal signals in Scn2a$^{+/-}$ mice showed little correlation with any HGF parameter, except for
370 a significant negative bias for level 3 uncertainties. This absence of significant correlations further
371 supports that the V1-to-SC pathway in this ASD mouse model fails to convey contextual
372 information about the visual environment.

373 Third, hardly any HGF parameter was well represented by the incoming signals from the
374 retina (**Fig. 7e,f and Supplementary Table 2**). This suggests that retinal inputs have little to do
375 with the contextual learning process, but provide a faithful representation of the visual stimuli.

376 Taken together, these results provide additional evidence that SC neurons in WT animals
377 process uncertainties about the visual environment by incorporating contextual information from
378 V1. In contrast, SC neurons in Scn2a$^{+/-}$ mice focus primarily on lower levels of uncertainty from
379 recent sensory experiences due to anomalous V1 feedback signaling. This likely causes an
380 oversensitivity to immediate outcomes in ASD (Gomot et al., 2011), highlighting the importance
381 of feedback signaling from V1 in properly processing visual environment context.

382 Discussion

383 Many species have evolved innate behaviors such as flight or freeze responses as essential
384 survival strategies to evade predators (Tinbergen, 1951; LeDoux, 2012; Branco & Redgrave,
385 2020). Mice, for example, respond rapidly and robustly to overhead shadows as they often imply
386 a threat, such as an aerial predator (Yilmaz and Meister, 2013; De Franceschi et al., 2016). This
387 instinctive response is controlled by the superior colliculus (SC; Sahibzada et al., 1986; Evans et
388 al., 2018; Shang et al., 2018), a midbrain structure that directly transforms visual signals from the
389 retina into motor commands, hence bypassing higher cognitive processing in the neocortex (Cang

390 et al., 2018). This helps animals react fast to potential threats, prioritizing survival over conscious
391 vision. This can, however, do more harm than good to animals if their response flexibility is too
392 limited, much as in autism spectrum disorder (ASD; Marco et al., 2011; Johnson et al., 2016; Jure,
393 2022). In autumn, for instance, looming shadows arise frequently from falling leaves; however,
394 they are harmless and thus animals do not need to respond. Adaptive learning allows animals to
395 tailor their behavior according to the specific context, determining properly when to react to or
396 ignore potential threats, such as looming shadows, to ensure survival without unnecessary
397 reactions.

398 In this study, we dissected the neurophysiological basis of such contextual learning by
399 examining the responses of SC responses as well as the incoming signals from the retina and the
400 primary visual cortex (V1) in animals engaged in a visual implicit learning task (**Fig. 8**). In
401 particular, wild-type (WT) mice revealed two key features of the contextual visual integration in
402 the early visual system. The first is a critical role of the top-down signals from V1 to SC. Similar
403 contextual modulation patterns were observed in the aversive stimulus responses of SC cells (**Fig.**
404 **2**) and V1 inputs (**Fig. 4**), but not retinal ganglion cell (RGC) axons (**Fig. 3**). Such modulation in
405 SC was eliminated by chemogenetically blocking V1-to-SC projections (**Fig. 5**). These data
406 support that V1, but not the retina, provides SC with contextual information via the corticotectal
407 projections from layer 5 (L5) pyramidal cells (Tang and Higley, 2020). Furthermore, the pupil
408 dynamics of WT mice also demonstrated associated contextual modulation in a way dependent
409 on the V1 inputs to SC (**Fig. 6**). This suggests that the cortical feedback is central to the adaptive
410 responses at both neurophysiological and behavioral levels.

411 Another feature of the contextual visual processing is the selectivity of SC to biologically
412 relevant stimuli. While V1 inputs showed adaptation over trials in response to the cue and omitted
413 stimuli, no such change in the response strength was observed in SC or pupil dynamics
414 (**Supplementary Figs. 2 and 3**). Adaptation in SC cells was observed only to predictable aversive

415 stimuli in a stable environment, likely to prevent an animal's excess reaction; and these SC
416 neurons recovered sensitivity in a volatile environment to regain proper behavioral responses
417 (**Figs. 2 and 6**). This selectivity indicates that the visual processing in SC is geared to those stimuli
418 that trigger innate defensive behaviors (Dean et al., 1989; Wei et al., 2015; Yilmaz and Meister,
419 2013; De Franceschi et al., 2016).

420 As an ASD model animal to investigate, here we chose Scn2a-haploinsufficient mice for
421 two reasons. First, SCN2A is one of the genes most strongly linked to ASD in humans (Satterstrom
422 et al., 2020). Children with mutations in SCN2A have also demonstrated anomalous acoustic
423 responses at the level of electroencephalography (Ben-Shalom et al., 2017). This suggests
424 anomalies in sensory perception, though no direct tests on contextual learning (as in Lawson et
425 al., 2017) or visual processing have been reported thus far. Second, Scn2a heterozygous
426 knockout is arguably one of the best-established ASD mouse models available (Spratt et a.,
427 2019). It encodes a $Na_V1.2$ subunit that has a clear relevance to neuronal function. Furthermore,
428 $Scn2a^{+/-}$ mice have defects in L5 corticofugal neurons (Spratt et al., 2019). These phenotypes
429 made it suitable for addressing how SC cells integrate bottom-up retinal inputs and top-down V1
430 inputs during contextual visual learning.

431 As expected from previous studies on ASD subjects (Lawson et al., 2017), $Scn2a^{+/-}$
432 animals exhibited anomalous responses during the implicit learning task at both
433 neurophysiological and behavioral levels. In particular, SC neurons demonstrated stable, non-
434 adaptive responses to the aversive stimuli in a stable environment, and an over-enhanced activity
435 in a volatile environment (**Fig. 2**), while no contextual modulation was observed in the pupil
436 dynamics (**Fig. 6**). Notably, the V1 inputs exhibited no substantial modulation, either (**Fig. 4**); and
437 yet, blocking the V1-to-SC projections in $Scn2a^{+/-}$ mice abolished the atypical modulation pattern
438 of the SC populations (**Fig. 5**). Taken together with the results from WT animals, we then suggest
439 that the feedback signals from V1 have two counteracting roles in modulating SC responses (Tang

440 and Higley, 2020) – suppression linked with stability and facilitation with volatility – and their
441 imbalance underlies the anomaly in Scn2a$^{+/-}$. Specifically, the suppressive function of V1 feedback
442 seems primarily impaired, leading to an imbalance towards facilitation in SC (**Fig. 2**). It has been
443 shown that reduced expression of $Na_V1.2$ channels in Scn2a$^{+/-}$ affects signal integration properties
444 of these L5 corticofugal pyramidal cells (Spratt et al., 2019). They may then fail to properly
445 integrate top-down cortico-cortical signals at the apical dendrites that carry contextual information,
446 especially about environmental stability.

447 Following the Bayesian inference principles (Pellicano and Burr, 2012; Friston, 2012), we
448 designed our learning paradigm so that defects in the prior formation will lead to anomalies in the
449 posterior dynamics during the stable environment, whereas those in likelihood will alter the
450 responses in the volatile environment (**Fig. 1**). Because we do not know exactly how the brain
451 represents those Bayesian probability distributions (but see, e.g., Walker et al., 2020; Sohn and
452 Narain, 2021), we cannot directly relate them to the observed neuronal activity. Nevertheless,
453 Scn2a$^{+/-}$ mice failed to show adaptive responses in both SC and V1 inputs during the LS epoch
454 (**Figs. 2 and 4**), while facilitation of SC during the VL epoch remained intact (**Fig. 2**). This implies
455 that these animals may suffer from a too-weak prior in the learning process, likely taking place in
456 higher-order brain structures, such as the anterior cingulate cortex whose activity has been
457 suggested to be sensitive to the environmental volatility (Behrens et al., 2007); and that the
458 information about the posterior is likely transmitted to SC via the top-down modulatory signals.

459 The above-described scenario was further supported by our hierarchical Gaussian filter
460 (HGF) model analysis (**Fig. 7**). While the presence of correlations between neuronal and model
461 parameter dynamics does not necessarily mean that these neurons implement the model, we
462 found in both WT and Scn2a$^{+/-}$ animals that HGF parameters were widely represented across all
463 levels by subpopulations of SC neurons. There was, however, a representation bias at different
464 hierarchical levels: level 3 for WT mice, while level 1 for Scn2a$^{+/-}$ animals (**Supplementary Table**

465 **2**). This may explain why ASD individuals tend to focus more on immediate sensory inputs and
466 associated outcomes (Lawson et al., 2017; Baron-Cohen et al. 2009; Bertone et al., 2005), as
467 they both belong to lower-level HGF inference processes (Mathys, 2011). Notably, WT mice
468 showed a transformation of the HGF parameter representations from lo-level ones in V1 inputs to
469 high-level ones in SC populations, whereas V1 inputs in Scn2a$^{+/-}$ animals showed little correlations
470 to the HGF parameters (**Fig. 7**). If this is indeed the case, resulting sensory overload and/or
471 oversensitivity to unpredictability in ASD (Gomot et al., 2011) may be alleviated by adjusting
472 cortical feedback to bias neural computation towards higher-level HGF inferences in SC, hence
473 redirecting the focus from immediate details to broader contexts. Here we focused on a specific
474 brain region and a specific computation involved; however, our findings may extend as a more
475 generic computational pattern of ASD within complex, recursive feedforward-feedback
476 interactions of sensory and associative systems. In future studies, it would be interesting to
477 investigate other brain areas in a similar Bayesian framework, and also test different ASD mouse
478 models for better understanding neurophysiological underpinnings of ASD.

479 Methods

480 No statistical method was used to predetermine the sample size. We performed all animal
481 experiments under license 233/2017-PR from the Italian Ministry of Health, following the protocol
482 approved by the Institutional Animal Care and Use Committee at European Molecular Biology
483 Laboratory. We conducted data analyses using Python, Matlab (Mathworks) and R, and reported
484 all summary statistics as mean ± standard error of the mean (s.e.m.) unless specified otherwise.
485 We considered *p*- and *q*-values significant at $\alpha < 0.05$ *, 0.01 **, 0.001 ***, unless stated otherwise.

486 Animals

487 We used C57BL/6J mice (*Mus musculus*, both sex), aged approximately 5-10 weeks at the time
488 of the initial surgery. We housed the mice under a 12-hour light/12-hour dark cycle and gave them

489 *ad libitum* access to water and food. After the surgical procedures, we grouped the animals based
490 on their operation day. We conducted the imaging experiments when the mice were between 8-
491 16 weeks old.

492 We acquired SCN2A haploinsufficient mice as frozen sperm (Scn2a^{-}) from Riken
493 Bioresource Research Center (RBRC10243), and revitalized the line in a C57BL/6J background.
494 The homozygous mutants (Scn2a$^{-/-}$) were lethal. We used the heterozygous mice (Scn2a$^{+/-}$) as a
495 model of autism spectrum disorder (ASD; Sandlers 2012; Ben-Shalom et al., 2017), and
496 considered their homozygous littermates (Scn2a$^{+/+}$) as wild-type (WT).

497 Intracranial viral injections

498 We locally injected pseudotyped adeno-associated viruses (AAVs) into the mouse superior
499 colliculus (SC). These AAVs consisted of AAV2 *rep* and AAV9, AAV5, or AAV2-retro *cap* genes,
500 or a hybrid of AAV1 and AAV2 *cap* genes; and were used to express genetically-encoded calcium
501 indicators (GECIs; jGCaMP7f, jGCaMP8m, or jRGECO1a; **Fig. 2a**) or Cre-recombinase under the
502 pan-neuronal human synapsin (hSyn) promoter. We also injected pseudotyped AAVs, composed
503 of AAV2 *rep* and AAV8 *cap*, into the primary visual cortex (V1) to express jGCaMP7b (**Fig. 4a**) or
504 the inhibitory designer receptor exclusively activated by designer drugs (DREADD), hM4Di-
505 mCherry, under the hSyn promoter (Dana et al., 2019; **Fig. 5a,b**).

506 We performed the intracranial viral injection simultaneously with the cranial implantation,
507 as described below. We created a craniotomy over the right SC or V1 and inserted an injection
508 pipette (with a tip diameter of approximately 30 μm; WPI 1B120F-3 borosilicate glass capillary
509 pulled with Zeitz DMZ puller) filled with a virus solution (ranging from approximately $5x10^{12}$ to
510 $4x10^{14}$ vg/mL in phosphate-buffered saline, PBS) through the dura. For SC, we positioned the
511 pipette using coordinates relative to Bregma, approximately [-4]-[-4.5] mm anterior-posterior (AP),
512 0.5-0.7 mm medial-lateral (ML), and slowly advanced it to a depth of about 1.25 mm. For V1, we

513 used the following coordinates: -3.5 mm AP, 1.8-2.0 mm ML, and 0.6 mm dorsal-ventral (DV). We
514 injected the virus solution at a rate of 1-2 nL/s for SC (0.4-0.6 µL in total) and 0.5-1 nL/s for V1
515 (0.25-0.40 µL in total) using a microinjection pump (either Neurostar NanoW or WPI NanoLiter
516 2010). After completing the injection, we waited at least 10 minutes before slowly withdrawing the
517 pipette, and then proceeded with the cranial window implantation procedure.

518 Intravitreal viral injections

519 We performed Intravitreal injections of AAV2, pseudotyped with a hybrid of AAV1 and AAV2
520 capsids, to deliver hSyn-axon-GCaMP6s expression cassette to the mouse retinal ganglion cells
521 (RGCs; Molotkov, et al., 2023; **Fig. 3a**). We exposed the scleral surface on the left eye and made
522 a small piercing with a sterile 28-30G needle in between the sclera and the cornea. We then
523 inserted an injection pipette (~50 µm tip diameter with 30-40° bevel) prefilled with a virus solution
524 (~1.5×10^{14} vg/mL in phosphate-buffered saline with 0.001% Pluronic F68 and 0.001% FastGreen)
525 into the vitreous chamber, approximately 1 mm deep. After a good sealing of the pipette was
526 formed, we injected 1.2 µL of the virus solution at a rate of 10 nL/s, using a microinjection pump.
527 We slowly withdrew the pipette at least 5 minutes after the completion of the injection, and covered
528 the treated eye with the eye ointment. We carried on the rest of the procedure and the head
529 implant as described in the previous paragraph on intracranial injections.

530 Cranial implantations

531 We performed the cranial window implantation over the mouse SC as described previously
532 (Molotkov et al., 2023). Before the surgery, we prepared a cranial window assembly by activating
533 the surface of a circular glass coverslip (5 mm diameter, 0.13-0.15 mm thickness; Assistent Karl
534 Hecht) with a laboratory corona treater (BD-20ACV Electro-Technic Products). We then fused the
535 activated glass coverslip to a cylindrical silicone plug (1.5 mm diameter, 0.75-1.00 mm height;
536 Kwik-Sil, WPI) by baking it at 70-80 °C for 24 hours.

537 During the implantation, we anesthetized the mice (induction: 4% isoflurane in oxygen;
538 maintenance: 1.5-2.0%) and positioned them in a stereotaxic apparatus (Stoelting 51625). We
539 used a heated plate (Supertech Physiological Temperature Controller) to maintain body
540 temperature and prevent hypothermia, and applied eye ointment (VitA-POS, Ursapharm) to
541 protect the eyes. We disinfected and removed the scalp (Betadine 10%, Meda Pharma), then
542 scratched and cleaned the skull surface to ensure the cement adhered well.

543 Using a high-speed surgical drill (OmniDrill35, WPI) with a 0.4 mm ball-tip carbide bur
544 (Meisinger), we made a craniotomy of approximately 3.0 mm AP by 2.5 mm ML over the right SC.
545 To prevent bleeding, we applied haemostatic sponges (Cutanplast, Mascia Brunelli) soaked in
546 sterile cortex buffer (NaCl 125 mM, KCl 5 mM, Glucose 10 mM, HEPES 10 mM, $CaCl_2$ 2 mM,
547 $MgSO_4$ 2 mM, pH 7.4) to the craniotomy site. We carefully positioned the cranial window
548 assembly, pushing the transversal sinus and posterior cortex about 0.5 mm forward, with the
549 silicone plug covering the medial-caudal region of the right SC (see, e.g., **Fig. 2a**). We fixed and
550 sealed the implant with tissue adhesive (Vetbond, 3M). We then cemented a custom-made
551 titanium headplate (0.8 mm thick) to the skull using a blend of acrylic cement powder (Paladur,
552 Kulzer) and cyanoacrylate adhesive (Loctite 401, Henkel).

553 After the surgery, the animal was allowed to recover from anesthesia in a warmed chamber
554 and then returned to its home cage. For postoperative care, we administered intraperitoneal
555 injections of 5 mg/kg Rimadyl (Zoetis) and 5 mg/kg Baytril (Bayer) to the mice daily for 3-5 days.
556 We then waited an extra 20-30 days to ensure the cranial window fully recovered before starting
557 *in vivo* two-photon imaging sessions.

558 Habituation

559 One week before the experiment began, we habituated the mice to both the experimenter and the
560 head-fixation procedure in a dark setting. We positioned the animals on a stable platform (non-

561 rotating disc; about 20 cm diameter) and let them stay still. We mounted a plexiglass tunnel over
562 the rear part of their bodies to increase comfort. We conducted these habituation sessions twice
563 a day, each lasting for 2 hours, over five to ten days. This process accustomed the mice to the
564 experimental setup and aimed to reduce any stress during the actual data collection.

565 Experimental setup

566 Two to seven days after training, we placed the animal on a stable platform with its head fixed
567 under a two-photon microscope, in front of a spherical projection screen. We used a non-rotating
568 disc (about 20 cm diameter) as the platform to minimize the effect of motor activities on neuronal
569 firing during the recordings (Schröder et al., 2020). The platform, however, allowed for vertical
570 excursions, which was compensated by a spring and a counter-weight adjustable for each animal
571 and monitored by a Hall sensor. This helped minimize head movements by preventing the mouse
572 from exerting excessive force on the head-holder with its paws. We excluded from the analysis
573 any trials if the average vertical excursion of the platform during the trial exceeded 0.25 mm. We
574 used an infrared (IR) camera (Imaging Source DMK23UX174) with IR illumination to monitor the
575 movement of the animal's left eye during recordings (**Fig. 2a**; 30 frames/s; 640-by-480 pixels).

576 Chemogenetics and histology

577 We had a subset of mice with hM4Di expressed in V1 (7 WT and 7 Scn2a$^{+/-}$; **Fig. 5 and**
578 **Supplementary Fig. 5**). These animals underwent the experiment twice, with at least a two-week
579 interval. For both experiment repetitions, mice received intraperitoneal injections of either a saline
580 solution (0.9% NaCl) containing 0.2 mg/ml clozapine N-oxide (CNO), leading to a final dose of 3
581 mg/kg (administered in a volume of 0.1-0.3 ml, adjusted for mouse weight), or the saline alone.
582 We alternated the solution allocation: mice receiving saline alone in the first iteration were given
583 the CNO solution in the second, and vice versa. Starting 45-70 minutes post-injection, the mice
584 underwent the experiment for approximately one hour.

585 Once all the recording sessions were done, we histologically examined the expression of
586 hM4Di-mCherry. We anesthetized the mice terminally with a ketamine/xylazine overdose, and
587 perfused them intracardially with 4% paraformaldehyde (PFA) in PBS. We then harvested the
588 brains, incubated them for a day in 4% PFA, and for at least another day in PBS. We sliced the
589 fixed brains into coronal sections (100 μm thickness) with a cryostat (Leica CM3050S). We then
590 mounted the slices on a glass coverslip using 4′,6-diamidino-2-phenylindole (DAPI)-mount. We
591 performed fluorescent imaging with a slide scanner (VS200 automated slide scanner, Olympus)
592 with a 10X dry objective (UPLAPO 10x, NA: 0.4, X-line) across three channels: DAPI, GCaMP,
593 and mCherry (**Fig. 2b**). Only those animals that displayed clear mCherry expression (hence
594 hM4Di expression) in V1 layer 5 (L5) somata (e.g., **Fig. 5b**) were included in the analysis (**Fig.**
595 **5c,f; Supplementary Fig. 5a,b**; V1-block groups).

596 Visual stimulation

597 We presented visual stimuli to the mice as described previously (Boissonnet et al., 2023). We
598 used a custom gamma-corrected digital light processing device to project images onto a spherical
599 screen placed approximately 20 cm from the left eye of the animal. The projected images covered
600 the left monocular visual field within ±22° in elevation and ±36.5° in azimuth (**Fig. 2a**). The images
601 were displayed at a resolution of 1280-by-720 pixels and a frame rate of 60 Hz.

602 One day before the experiment, mice underwent a training process. We first exposed to
603 them a gray screen for at least 10 minutes to adapt to the average luminance in a head-fixed
604 condition under the microscope. Subsequently, we presented the following visual stimuli:

605 1. Still gratings: A random sequence of black and white square wave gratings in four
606 orientations (0°, 45°, 90°, 135°) with a spatial frequency of 0.1 cycles/°. Each stimulus
607 lasted for 1 second, followed by a 0.5-second interstimulus interval (ISI). This sequence
608 was repeated 46 times in 5 minutes.

2. Cue-outcome stable 100%: Horizontal and vertical bars (abovementioned still gratings with 0° and 90° orientations, respectively) served as cue stimuli. Horizontal bars were followed by an omitted stimulus (prolonged ISI with a gray screen for 2 seconds), while vertical bars were followed by one out of four possible aversive stimuli (2 s): Sweep-left (a black disk 8° in diameter, appearing at 10° from the top edge and moving horizontally from the left edge to the right one at 42 °/s), Sweep-right (moving from right to left), Sweep-left-loom (a Sweep-left for the first second to bring the disk at the upper center of the screen, followed by the disk increasing in size for the next second at a speed of 57°/s while moving its center downward until reaching the horizontal center and a final size of 64° in diameter), and Sweep-right-loom (Sweep-right for the first half, followed by the loom for the second half). To avoid pupillary light reflex, the average luminance of the screen was kept constant at any time. For the sweep and loom stimuli, the background luminance was adjusted for the relative area of the black disk at every frame. We presented three blocks of 72 cue-outcome trials (approximately 5 min), where the stimulus sequence of each block was pseudo-randomized with different seeds.
3. Cue-outcome stable 83%: The cue-outcome probabilities were changed so that horizontal-omitted and vertical-aversive trials occurred with a probability of 83.4%, while horizontal-aversive and vertical-omitted trials occurred with a probability of 16.6%. We presented three blocks of 72 cue-outcome trials (approximately 5 min), where the stimulus sequence of each block was pseudo-randomized with different seeds.

On the day of the experiment, mice were again habituated to the average luminance for at least 10 minutes. We then presented the following stimuli:

1. Still gratings.
2. Cue-outcome stable 83%, but with six different types of aversive stimuli. Besides the four aversive stimuli described above, we included: Slow-loom (a black 3° disk appeared in the

vertical center at the top edge and increased its size at a speed of 28.5°/s while moving its center downward until reaching the horizontal center and a final size of 64° in diameter) and Fast-loom (a black 3° disk appeared in the vertical center at the top edge and stayed still for 0.5 seconds; it then increased its size at a speed of 57°/s while moving its center downward until reaching the horizontal center and a final size of 64° in diameter; it then stayed still again for 0.5 seconds). In total, 216 trials were presented in three blocks (72 trials each; **Fig. 1 and Supplementary Fig. 1**). Animals typically moved a lot at the beginning and stayed still after a while. Thus, the first block was excluded from the analysis to avoid any movement artifact. The second and third ones were used for the analysis as early-stable (ES) and late-stable (LS) epochs, respectively (**Fig. 1b**).

3. Cue-outcome volatile: The cue-outcome stimulus sequences were presented with three different contingencies: a given cue-outcome (horizontal-omitted and vertical-aversive) probability of 16.6% for the first 24 trials; 50% for the next 24 trials; and 83.4% for the last 24 trials (**Fig. 1c**). These data (72 trials in total) were referred to as the volatile (VL) epoch in the analysis. Collectively, the cue-outcome stimuli #2 and #3 constituted an implicit visual learning task.
4. Random water-wave stimuli (3 minutes): These stimuli were used to generate binary masks for signal source extraction in calcium image analysis (see below “Calcium image analysis” section).
5. Randomly flickering black-and-white checkerboard stimuli (10 minutes): These stimuli were used for receptive field mapping (**Supplementary Fig. 6**). The rectangular fields had a width of 3.7° and a height of 2.9°. Each field was independently modulated by white noise at a frequency of 4 Hz.

In the cue-outcome stimuli (#2 and #3), the trials were evenly balanced within each analysis epoch (72 trials each), with 36 horizontal and 36 vertical cues presented before 36 omitted and 36

659 aversive outcomes. Each of the six different aversive stimuli was presented six times within the
660 36 aversive trials. The sequences of stimuli in each epoch were pseudorandomized with different
661 seeds, but we used the same set of seeds for all mice. This ensures that each mouse was
662 presented with a new random sequence of stimuli in each epoch, but all mice were presented with
663 the same sets of sequences and cue-outcome contingencies (**Fig. 1**).

664 For the subset of mice subject to the chemogenetics experiments (see "Chemogenetics
665 and histology" section above), we repeated the stimulus paradigm on the same experimental day
666 for CNO or saline treatment.

667 *In vivo* two-photon imaging

668 During the experiment, we kept the animal with its head fixed inside the rig for a maximum duration
669 of 2 hours (typically 2 to 5 imaging sessions per animal). We performed two-photon calcium
670 imaging using a galvo-resonant scanner (Scientifica HyperScope) and SciScan image acquisition
671 software. The microscope setup included a mode-locked Ti:sapphire tunable laser (InSight DS+,
672 Spectra-Physics) and a plan fluorite objective (CFI75 LWD 16X W, Nikon).

673 In each imaging session, we conducted single-plane time-lapse recordings with a field of
674 view approximately ranging from 0.65-by-0.65 mm to 0.43-by-0.43 mm at a depth of 115-130 μm
675 from the surface of SC. We used the excitation wavelength of 920 nm for GCaMPs, and 1,040 nm
676 for jRGECO1a. The average laser power under the objective ranged between 40-80 mW. We
677 bandpass-filtered the emitted fluorescence signal (BP 527/70 or BP 650/100 after beam-splitter
678 FF580-FDi01, Semrock) and detected it using a non-descanned gallium arsenide phosphide
679 photomultiplier tube (Hamamatsu GaAsP PMT). We acquired each frame with either 512-by-512
680 pixels (16-bit depth) at a rate of 30.9 Hz, or 1,024-by-1,024 pixels at a rate of 15.5 Hz (e.g., **Figs.
681 2b, 3a, and 4a**).

682 ## Calcium image analysis

683 To simplify manual alignment across different recordings, we first cropped the recorded images
684 from 512-by-512 pixels to 480-by-480 pixels (1.3-0.85 μm/pixel), or from 1,024-by-1,024 to 960-
685 by-960 (0.65-0.43 μm/pixel) using ImageJ software. We then applied a series of rigid and non-
686 rigid motion corrections in CaImAn (version 8 or 9; Pnevmatikakis et al., 2019).

687 To generate masks for cellular detection in CaImAn, we used either a segment of the
688 recording containing 3,000 frames from the random water-wave stimulus presentation period or
689 the entire recording frames. We divided the images into 6-by-6 patches for the constrained non-
690 negative matrix factorization (CNMF) deconvolution, resulting in a total of 36 patches. We set the
691 expected number of neurons per patch (*params.K*) to 5 for SC cell bodies, and set the size of the
692 neurons (*params.gSig*) to 5-by-5 pixels (*half size*) for 512-by-512 acquisitions, or 10-by-10 pixels
693 for 1,024-by-1,024 acquisitions. For RGC boutons, we set 8 for *params.K* and 2.5-by-2.5 pixels
694 for *params.gSig*. We manually selected putative cells that had a uniformly-filled round shape of
695 approximately 10-20 μm in size, and putative boutons to be below 2-3 μm. We then converted
696 these selected components into binary spatial masks and performed two iterations of masked
697 CNMF in CaImAn on the entire time-lapse recordings. We used the resulting spatial components
698 (*estimates.A*; e.g., **Figs. 2c,f, 3a,d, and 4a,d**) and deconvolved neural activities, represented as
699 spike estimations based on the df/f calcium transients (*estimates.S*; e.g., **Supplementary Fig. 1**),
700 for subsequent analyses.

701 ## Receptive field analysis

702 The receptive fields (RFs) of SC somata or putative boutons of RGCs or V1 L5 projection neurons
703 were estimated by reverse-correlation methods using the random checkerboard stimuli (as in
704 Molotkov et al., 2023). Specifically, we calculated the response-weighted average of the stimulus
705 waveform (0.5 s window; 16.7 ms bin width) and characterized its spatial profile by the two-

706 dimensional Gaussian curve fit at the peak latency (e.g., **Supplementary Fig. 6a**). The RF size
707 was then defined as twice the mean standard deviation of the long and short axes.

708 Two-way ANOVA (Matlab: anovan) was used to make a comparison across "Experimental Group"
709 = {WT SC, Scn2a$^{+/-}$ SC, WT SC with V1-block, Scn2a$^{+/-}$ SC V1-block, WT V1, Scn2a$^{+/-}$ V1, WT
710 RGC, Scn2a$^{+/-}$ RGC } and "Mouse" (41 mice), with the "Mouse" nested within the "Experimental
711 Group." For pairwise comparisons across epochs within each experimental group, we used the
712 Tukey-Kramer tests (R: pairs(emmeans) on aov; **Supplementary Fig. 6b**)

713 Analysis of neuronal activity contextual modulation

714 First, the estimated spiking activity of each putative neuronal or axonal component in each
715 analysis epoch was extracted from the entire time lapse and converted into z-scores (Matlab:
716 zscore): $Z(t)$. To compare the aversive stimulus response across analysis epochs, we then
717 calculated the average response across trials, $z(t) = \langle Z(t) \rangle_{\text{trials}}$, for each cell for each epoch (e.g.,
718 **Fig. 2c,d,f,g**), excluding those trials with excessive vertical platform excursion (>0.25 mm on
719 average). We then selected responsive cells for subsequent analyses that had $z(t)$ exceeding a
720 predetermined threshold (1 for SC somata; 0.75 for axon terminals of RGCs or V1 cells) in at least
721 one analysis epoch.

722 For each responsive neuron, we took an average of $z(t)$ over time during the aversive
723 stimulus presentation period (2 s) for each analysis epoch (e.g., **Fig. 2d,g**): $\bar{z} = \langle z(t) \rangle_t$. We then
724 subtracted the average of $\bar{z}$ across the three epochs for normalization: $\Delta z = \bar{z} - \langle \bar{z} \rangle_{\text{epochs}}$, and used
725 this quantity Δz for the population analysis (e.g., **Fig. 2e,h**). For comparisons across animals, we
726 performed a 3-way analysis of variance (ANOVA; Matlab: anovan) on $\bar{z}$ with a full interaction
727 model. The three variable categories are: Mouse (41 mice), Experimental Group (WT SC, Scn2a$^{+/-}$
728 SC, WT SC with V1-block, Scn2a$^{+/-}$ SC with V1-block, WT V1, Scn2a$^{+/-}$ V1, WT RGC, Scn2a$^{+/-}$
729 RGC), and Epochs (ES, LS, VL) with the "Mouse" nested within the "Experimental Group." This

approach ensures the preservation of the variance structure across groups for assessing the overall effects. For pairwise comparisons across epochs within each experimental group, we used the Tukey-Kramer tests on Δz (R: pairs(emmeans) on aov; e.g., **Fig. 2e,h**). This maintains the nested ANOVA structure for assessing within-group differences.

Pupillometry

Pupil detection and eye-tracking were performed with a region-based convolutional neural network (R-CNN; He et al. 2017). The network was trained by a set of 260 mouse eye video frames with manually segmented pupil outlines (Boissonnet et al, 2023). The detected pupil was parameterized by a 2-dimensional Gaussian curve fit to the pixel map. We then calculated the area of the pupil, $A(t)$, for each frame based on the estimated circular area from the long-axis diameter (6×10^3 pixel^2 = 1 mm^2; e.g., **Fig. 6a,b**). Note that, assuming a circular pupil on a spherical eyeball, the pupil position affects its short-axis diameter, making a circular pupil ellipsoidal, but the long-axis diameter is invariant. Mann-Whitney U-test was used to compare the median pupil area during the implicit visual task between WT and Scn2a$^{+/-}$ mice (**Fig. 6c**).

To analyze pupil size dynamics across analysis epochs, we first calculated the proportional change in pupil size upon each trial of the aversive stimulus presentation, and then took the average over trials for each epoch: $a(t) = \langle A(t)/A_0 \rangle_{\text{trials}}$, where A_0 is the average pupil size over the 10 frames before the stimulus onset for each trial (**Fig. 6d-g**). We then took the average over the stimulation presentation period (2 s) for each analysis epoch: $\overline{a} = \langle a(t) \rangle_t$, and subtracted the mean across epochs for normalization: $\Delta a = \overline{a} - \langle \overline{a} \rangle_{\text{epochs}}$ (**Fig. 6h-k**). We performed 2-way ANOVA (Matlab: anovan) to make a comparison across "Experimental Group" (WT, Scn2a$^{+/-}$, WT with V1-block, Scn2a$^{+/-}$ with V1-block) and "Epoch" (ES, LS, VL). For pairwise comparisons of Δa across epochs at the population level, we used the Tukey-Kramer tests (Matlab: multcompare).

Hierarchical Gaussian filter model

To simulate how an ideal Bayesian observer would behave in response to the implicit visual learning task, we implemented a hierarchical Gaussian filter (HGF) model (Mathys, 2011; **Fig. 7a**) and calculated its parameter trajectories based on the 216 cue-outcome contingencies (**Fig. 7b**). The stimuli presented during the experiments followed a pseudo-random sequence that was identical for all animals as described in the "visual stimuli" section (**Fig. 1**). Thus, the sequence of contingencies remained the same across the subjects. For each trial k, we assigned a binary value to the contingency: 0 for an expected outcome (vertical-aversive or horizontal-omitted) and 1 for an unexpected outcome (vertical-omitted or horizontal-aversive). These values were represented as a binary vector u that served as the sensory input for the task. From u, the posterior expectations at three levels (μ_i with $i = 1,2,3$) were calculated as follows.

- At level 1, the posterior is equal to the sensory input: $\mu_1 = u$.
- At level 2, the posterior expectation about $u = 1$ from the previous trial $(k-1)$ is: $\mu_2 = \mu_2^{(k-1)} + \sigma_2^k \delta_1^k$, where $\delta_1^k = \mu_1^k - s\left(\mu_2^{(k-1)}\right)$ is the prediction error for level 1, with $s(\cdot)$ being a sigmoid function; and $\sigma_2^k = \frac{1}{1/\hat{\sigma}_2^k + \hat{\sigma}_1^k}$ is the standard deviation or variance at level 2.
- At level 3, the posterior expectation regarding environmental volatility is: $\mu_3 = \mu_3^{(k-1)} + \sigma_3^{k^{\frac{k}{2}}} w_2^k \delta_2^k$, where the weight is given as $w_2^k = \frac{\exp\left(\kappa\mu_3^{(k-1)} + \omega_2\right)}{\sigma_2^{(k-1)}\exp\left(\kappa\mu_3^{(k-1)} + \omega_2\right)}$ and the prediction error as $\delta_2^k = \frac{\sigma_2^k + \left(\mu_2^k - \mu_2^{(k-1)}\right)^2}{\sigma_2^{(k-1)} + \exp\left(k\mu_3^{(k-1)} + \omega_2\right)} - 1$.

From these posterior expectations μ_i, we derived the prior expectations at each level $\hat{\mu}_i$ that represent the prior prediction about the probability of having $u = 1$ before the sensory input is presented.

- At level 1, the prior prediction is the sigmoid of the posterior expectation at level 2: $\hat{\mu}_1 = s(\mu_2) = \frac{1}{1+\exp(-\mu_2)}$.
- At level 2, the prior prediction about the probability that $\mu_1 = 1$ is: $\hat{\mu}_2 = \mu_2^{(k-1)}$.
- At level 3, the prior prediction about environment volatility, or probability in change of μ_2, is: $\hat{\mu}_3 = \mu_3^{(k-1)}$.

From here, the variance of the belief at every level $\hat{\sigma}_i$, or the inverse of the predicted precision $\frac{1}{\hat{\pi}_i}$, is given as follows:

- At level 1, the belief variance or imprecision of prediction that $u = 1$ is: $\hat{\sigma}_1^k = \hat{\mu}_1(1 - \hat{\mu}_1)$.
- At level 2, the belief variance about the probability that $\mu_1 = 1$ is: $\hat{\sigma}_2^k = \sigma_2^{(k-1)} + \exp\left(\kappa_2 \mu_3^{(k-1)} + \omega_2\right)$.
- At level 3, the belief variance about environment volatility, or probability in change of μ_2, is: $\hat{\sigma}_3^k = \sigma_3^{(k-1)} + \exp(\omega_3)$.

To calculate the precision-weighted prediction errors ϵ_i, we first calculated prediction errors at each level:

- level 1: $\delta_1^k = \mu_1^k - \hat{\mu}_1^k$,
- level 2: $\delta_2^k = \left(\left(\sigma_2^k + \left(\mu_2^k - \hat{\mu}_2^k\right)^2\right)\hat{\pi}_2^k\right) - 1$, and
- level 3: $\delta_3^k = \left(\left(\sigma_3^k + \left(\mu_3^k - \hat{\mu}_3^k\right)^2\right)\hat{\pi}_3^k\right) - 1$, where $\pi_i = 1/\sigma_i$.

Because precision weights on prediction errors are $\psi_2 = \sigma_2$ at level 2; and $\psi_3 = \pi_3/\hat{\pi}_2$ at level 3, we have: $\epsilon_2 = \psi_2 \delta_1$ and $\epsilon_3 = \psi_3 \delta_2$.

The uncertainties unc_i are calculated as outcome uncertainty.

- At level 1, this corresponds to the predicted variance: $unc_1^k = \hat{\sigma}_1^k$.

- At level 2, the informational uncertainty is: $unc_2^k = s(\mu_2^k)\left(1 - s(\mu_2^k)\right)\sigma_2^k$.
- At level 3, the environmental uncertainty is: $unc_3^k = \exp\left(\kappa_2 \mu_3^{(k-1)} + \omega_2\right)$.

Finally, the learning rates α_i are calculated as follows:

- At level 1, $\alpha_1 = \left(s(\mu_2^k) - s\left(\mu_2^{(k-1)}\right)\right) / (u - s(\mu_2))$.
- At level 2, learning rate is equivalent to the precision weight, or variance at level 2: $\alpha_2 = \psi_2 = \sigma_2$.
- At level 3, $\alpha_3 = \frac{\exp\left(\kappa_2 \mu_3^{(k-1)} + \omega_2\right)}{2\,\psi_3}$.

In sum, the HGF model consisted of a set of these 13 parameters in five classes across three hierarchical levels **(Fig. 7b)**:

1. Learning rates: $\alpha_1, \alpha_2, \alpha_3$
2. Prior predictions: $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$
3. Variances of Belief: $\hat{\sigma}_2, \hat{\sigma}_3$
4. Uncertainties: $unc_1 (= \hat{\sigma}_1), unc_2, unc_3$
5. Precision-weighted prediction errors: ϵ_2, ϵ_3

After generating the model parameter trajectories from the cue-outcome contingency vector u, we eliminated those trials with non-aversive outcomes to make a fair comparison with the neuronal response trajectories that included the responses to aversive outcomes only (see below; **Fig. 7c,e**).

Neuronal trajectory estimation

Because we used six different types of aversive stimuli, and because the cue-outcome stimulus patterns were presented in blocks, the observed neuronal responses had a certain systematic bias that needed to be corrected before comparing them to the HGF model parameter trajectories.

To this end, we normalized the neuronal responses in two steps: 1) stimulus-specific normalization and 2) non-stationarity elimination. Here we focused only on the aversive stimulus responses, but not the cue stimulus responses, because the latter did not demonstrate context-dependent modulation in our experimental paradigm (**Supplementary Fig. 2**).

First, for each neuron, we calculated the average response in z-score for each aversive stimulus trial: $\overline{z}_k = \langle Z(t) \rangle_t$, and subtracted the average per stimulus type across three epochs (i.e., 18 trials in total per type): $\Delta z_{k|\text{type}} = \overline{z}_{k|\text{type}} - \langle \overline{z}_{k|\text{type}} \rangle_{\text{trials}|\text{type}}$. The resulting concatenated response Δz_k, termed as "neuronal trajectory" here, represented trial-to-trial differences of the neuronal activity, independent of the stimulus type.

These neuronal trajectories typically showed a slow decay over trials in each analysis epoch, likely due to photobleaching. To eliminate such non-stationary components, we next fitted a power function, $f(k) = a \cdot k^b + c$ (Matlab: fit "power2"), to the population average of all the neuronal trajectories in the dataset, averaged across epochs. We then subtracted a scaled decay curve, $d \cdot f(k)$, from the neuronal trajectory of each neuron. The same scalar, d, was used across epochs to avoid altering the context-dependent neuronal activity. The residual trajectory ($\Delta z_k^{\text{res}} = \Delta z_k - d \cdot f(k)$; **Fig.7c**) had a Pearson's correlation with the decay within ±0.05, and was used for the subsequent correlation analyses with the HGF model parameter trajectories.

Correlation analysis between neuronal trajectories and HGF trajectories

We used Pearson's correlation (Matlab: corr) to examine the relationship between the trajectories of neurons and the HGF model parameters. Specifically, we correlated each of the 13 HGF parameters (**Fig. 7b**) with every individual neuronal trajectory. The associated *p*-values were then converted to *q*-values with Storey's FDR methods (Matlab: mafdr) to determine the statistical significance. We then counted the fraction of SC neurons significantly correlated with each HGF parameter for each of the four experimental groups (WT SC, Scn2a$^{+/-}$ SC, WT SC with V1-block,

841 Scn2a$^{+/-}$ SC with V1-block; **Fig. 7d**, left panels), which we considered as the neuronal
842 representation of the HGF model. A bias of the correlation polarity was tested with a shuffling test
843 (10,000 repetitions, with FDR correction; **Supplementary Table 2**). The same procedure was
844 applied to the axonal responses of RGCs and V1 cells (**Fig. 7e,f**).

845 At the population level, we also calculated Pearson's correlation, ρ, between the
846 population average of the neuronal trajectories $\langle \Delta z_k^{\mathrm{res}} \rangle_{\mathrm{cells}}$ (**Fig. 7c**) and each of the HGF parameter
847 trajectories (**Fig. 7b**). We then took the cumulative sum of the neurons significantly correlated with
848 any HGF parameter at each level to compare how each experimental group represents the HGF
849 parameters across different levels (**Fig. 7d,f,** right panels).

850 # References

851 Ahmadlou, M., Tafreshiha, A., Heimel, J.A. (2017). Visual Cortex Limits Pop-Out in the Superior
852 Colliculus of Awake Mice. *Cereb. Cortex* **27**, 5772–5783.

853 Armbruster, B.N., Li, X., Pausch, M.H., Herlitze, S., Roth, B.L. (2007). Evolving the lock to fit the
854 key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc.*
855 *Natl. Acad. Sci. U.S.A.* **104**, 5163–5168.

856 Baron-Cohen, S., Ashwin, E., Ashwin, C., Tavassoli, T., Chakrabarti, B. (2009). Talent in autism:
857 hyper-systemizing, hyper-attention to detail and sensory hypersensitivity. *Phil. Trans. R. Soc. B*
858 **364**, 1377–1383.

859 Basso, M.A., May, P.J. (2017). Circuits for Action and Cognition: A View from the Superior
860 Colliculus. *Annu. Rev. Vis. Sci.* **3**, 197–226.

861 Benavidez, N.L., Bienkowski, M.S., Zhu, M., Garcia, L.H., Fayzullina, M., Gao, L., Bowman, I.,
862 Gou, L., Khanjani, N., Cotter, K.R., Korobkova, L., Becerra, M., Cao, C., Song, M.Y., Zhang, B.,
863 Yamashita, S., Tugangui, A.J., Zingg, B., Rose, K., Lo, D., Foster, N.N., Boesen, T., Mun, H.-S.,
864 Aquino, S., Wickersham, I.R., Ascoli, G.A., Hintiryan, H., Dong, H.-W. (2021). Organization of the
865 inputs and outputs of the mouse superior colliculus. *Nat. Commun.* **12**, 4004.

Ben-Shalom, R., Keeshen, C.M., Berrios, K.N., An, J.Y., Sanders, S.J., Bender, K.J. (2017). Opposing Effects on Na V 1.2 Function Underlie Differences Between SCN2A Variants Observed in Individuals With Autism Spectrum Disorder or Infantile Seizures. *Biol. Psychiatry* **82**, 224–232.

Behrens, T.E.J., Woolrich, M.W., Walton, M.E., Rushworth, M.F.S. (2007). Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221.

Bertone, A., Mottron, L., Jelenic, P., Faubert, J. (2005). Enhanced and diminished visuo-spatial information processing in autism depends on stimulus complexity. *Brain* **128**, 2430–2441.

Boissonnet, T., Tripodi, M., Asari, H. (2023). Awake responses suggest inefficient dense coding in the mouse retina. *eLife* **12**, e78005.

Branco, T., Redgrave, P. (2020). The Neural Basis of Escape Behavior in Vertebrates. *Annu. Rev. Neurosci.* **43**, 417–439.

Cannon, J., O'Brien, A.M., Bungert, L., Sinha, P. (2021). Prediction in Autism Spectrum Disorder: A Systematic Review of Empirical Evidence. *Autism Res*. **14**, 604–630

Cang, J., Savier, E., Barchini, J., Liu, X. (2018). Visual Function, Organization, and Development of the Mouse Superior Colliculus. *Annu. Rev. Vis. Sci.* **4**, 239–262.

Chrysaitis, A.N., Seriès, P. (2023). 10 years of Bayesian theories of autism: A comprehensive review. *Neurosci. Biobehav. Rev.* **145**, 105022.

Dana, H., Sun, Y., Mohar, B., Hulse, B.K., Kerlin, A.M., Hasseman, J.P., Tsegaye, G., Tsang, A., Wong, A., Patel, R., Macklin, J.J., Chen, Y., Konnerth, A., Jayaraman, V., Looger, L.L., Schreiter, E.R., Svoboda, K., Kim, D.S. (2019). High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat. Methods* **16**, 649–657.

Dean, P., Redgrave, P., Westby, G.W.M. (1989). Event or emergency? Two response systems in the mammalian superior colliculus. *Trends Neurosci.* **12**, 137–147.

De Franceschi, G., Vivattanasarn, T., Saleem, A.B., Solomon, S.G. (2016). Vision Guides Selection of Freeze or Flight Defense Strategies in Mice. *Curr. Biol.* **26**, 2150–2154.

Dobbins, I.G. (2021). Pupil dilation signals recognition salience. *Psychon. Bull. Rev.* **28**, 565–573.

892 Dombeck, D.A., Khabbaz, A.N., Collman, F., Adelman, T.L., Tank, D.W. (2007). Imaging large-
893 scale neural activity with cellular resolution in awake, mobile mice. *Neuron* **56**, 43–57.

894 Evans, D.A., Stempel, A.V., Vale, R., Ruehle, S., Lefler, Y., Branco, T. (2018). A synaptic
895 threshold mechanism for computing escape decisions. *Nature* **558**, 590–594.

896 Feinberg, E.H., Meister, M. (2015). Orientation columns in the mouse superior colliculus. *Nature*
897 **519**, 229–232.

898 Friston, K. (2009). The free-energy principle: a rough guide to the brain? *Trends Cogn. Sci.* **13**,
899 293–301.

900 Friston, K. (2012). Prediction, perception and agency. *Int. J. Psychophysiol.* **83**, 248–252.

901 Ganea, D.A., Bexter, A., Günther, M., Gardères, P.-M., Kampa, B.M., Haiss, F. (2020). Pupillary
902 Dilations of Mice Performing a Vibrotactile Discrimination Task Reflect Task Engagement and
903 Response Confidence. *Front. Behav. Neurosci.* **14**, 159.

904 Gomot, M., Blanc, R., Clery, H., Roux, S., Barthelemy, C., Bruneau, N. (2011). Candidate
905 Electrophysiological Endophenotypes of Hyper-Reactivity to Change in Autism. *J. Autism Dev.*
906 *Disord.* **41**, 705–714.

907 Haker, H., Schneebeli, M., Stephan, K.E. (2016). Can Bayesian Theories of Autism Spectrum
908 Disorder Help Improve Clinical Practice? *Front Psychiatry* **7**, 107.

909 He, K., Gkioxari, G., Dollár, P., Girshick, R. (2018). Mask R-CNN. *arXiv*, 1703.06870.

910 Hoffmann, K.P., Straschill, M. (1971). Influences of cortico-tectal and intertectal connections on
911 visual responses in the cat's superior colliculus. *Exp. Brain. Res.* **12**, 120–131.

912 Inayat, S., Barchini, J., Chen, H., Feng, L., Liu, X., Cang, J. (2015). Neurons in the Most Superficial
913 Lamina of the Mouse Superior Colliculus Are Highly Selective for Stimulus Direction. *J. Neurosci.*
914 **35**, 7992–8003.

915 Johnson, B.P., Lum, J.A.G., Rinehart, N.J., Fielding, J. (2016). Ocular motor disturbances in
916 autism spectrum disorders: Systematic review and comprehensive meta-analysis. *Neurosci.*
917 *Biobehav. Rev.* **69**, 260–279.

Joshi, S., Li, Y., Kalwani, R.M., Gold, J.I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron* **89**, 221–234.

Jure, R. (2022). The "Primitive Brain Dysfunction" Theory of Autism: The Superior Colliculus Role. *Front. Integr. Neurosci.* **16**, 797391.

Kato, R., Takaura, K., Ikeda, T., Yoshida, M., Isa, T. (2011). Contribution of the retino-tectal pathway to visually guided saccades after lesion of the primary visual cortex in monkeys. *Eur. J. Neurosci.* **33**, 1952–1960.

Larkum, M. (2013). A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends Neurosci.* **36**, 141–151.

Lawson, R.P., Mathys, C., Rees, G. (2017). Adults with autism overestimate the volatility of the sensory environment. *Nat Neurosci.* **20**, 1293–1299.

LeDoux, J.E. (2012). Evolution of human emotion: a view through fear. *Prog. Brain. Res.* **195**, 431–442.

Loewenfeld, I.E. (1958). Mechanisms of reflex dilatation of the pupil; historical review and experimental analysis. *Doc. Ophthalmol.***12**, 185–448.

Lomber, S.G. (2002). Learning to see the trees before the forest: Reversible deactivation of the superior colliculus during learning of local and global visual features. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 4049–4054.

Lomber, S.G., Payne, B.R., Cornwell, P. (2001). Role of the superior colliculus in analyses of space: superficial and intermediate layer contributions to visual orienting, auditory orienting, and visuospatial discriminations during unilateral and bilateral deactivations. *J. Comp. Neurol.* **441**, 44–57.

Marco, E.J., Hinkley, L.B.N., Hill, S.S., Nagarajan, S.S. (2011). Sensory Processing in Autism: A Review of Neurophysiologic Finding. *Pediatr. Res.* **69**, 48R-54R.

Mathys, C. (2011). A Bayesian foundation for individual learning under uncertainty. *Front. Hum. Neurosci.* **5**, 39.

Mathys, C.D., Lomakina, E.I., Daunizeau, J., Iglesias, S., Brodersen, K.H., Friston, K.J., Stephan, K.E. (2014). Uncertainty in perception and the Hierarchical Gaussian Filter. *Front. Hum. Neurosci.* **8**, 825.

May, P.J. (2006). The mammalian superior colliculus: laminar structure and connections. *Prog. Brain Res.* **151**, 321–378.

Molotkov, D., Ferrarese, L., Boissonnet, T., Asari, H. (2023). Topographic axonal projection at single-cell precision supports local retinotopy in the mouse superior colliculus. *Nat Commun* **14**, 7418.

Ogasawara, K., McHaffie, J.G., Stein, B.E. (1984). Two visual corticotectal systems in cat. *J. Neurophysiol.* **52**, 1226–1245.

Pellicano, E., Burr, D. (2012). When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends Cogn. Sci.* **16**, 504–510.

Planells-Cases, R., Caprini, M., Zhang, J., Rockenstein, E.M., Rivera, R.R., Murre, C., Masliah, E., Montal, M. (2000). Neuronal death and perinatal lethality in voltage-gated sodium channel alpha(II)-deficient mice. *Biophys. J.* **78**, 2878–2891.

Pnevmatikakis, E.A., Soudry, D., Gao, Y., Machado, T.A., Merel, J., Pfau, D., Reardon, T., Mu, Y., Lacefield, C., Yang, W., Ahrens, M., Bruno, R., Jessell, T.M., Peterka, D.S., Yuste, R., Paninski, L. (2016). Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. *Neuron* **89**, 285–299.

Ruediger, S., Scanziani, M., (2020). Learning speed and detection sensitivity controlled by distinct cortico-fugal neurons in visual cortex *eLife* **9**, e59247.

Sahibzada, N., Dean, P., Redgrave, P. (1986). Movements resembling orientation or avoidance elicited by electrical stimulation of the superior colliculus in rats. *J. Neurosci.* **6**, 723–733.

Sanders, S.J., Murtha, M.T., Gupta, A.R., Murdoch, J.D., Raubeson, M.J., Willsey, A.J., Ercan-Sencicek, A.G., DiLullo, N.M., Parikshak, N.N., Stein, J.L., Walker, M.F., Ober, G.T., Teran, N.A., Song, Y., El-Fishawy, P., Murtha, R.C., Choi, M., Overton, J.D., Bjornson, R.D., Carriero, N.J., Meyer, K.A., Bilguvar, K., Mane, S.M., Šestan, N., Lifton, R.P., Günel, M., Roeder, K., Geschwind,

971 D.H., Devlin, B., State, M.W. (2012). De novo mutations revealed by whole-exome sequencing
972 are strongly associated with autism. *Nature* **485**, 237–241.

973 Satterstrom, F.K., Kosmicki, J.A., Wang, J., Breen, M.S., De Rubeis, S., An, J.-Y., Peng, M.,
974 Collins, R., Grove, J., Klei, L., Stevens, C., Reichert, J., Mulhern, M.S., Artomov, M., Gerges, S.,
975 Sheppard, B., Xu, X., Bhaduri, A., Norman, U., Brand, H., Schwartz, G., Nguyen, R., Guerrero,
976 E.E., Dias, C., Autism Sequencing Consortium, iPSYCH-Broad Consortium, Betancur, C., Cook,
977 E.H., Gallagher, L., Gill, M., Sutcliffe, J.S., Thurm, A., Zwick, M.E., Børglum, A.D., State, M.W.,
978 Cicek, A.E., Talkowski, M.E., Cutler, D.J., Devlin, B., Sanders, S.J., Roeder, K., Daly, M.J.,
979 Buxbaum, J.D. (2020). Large-scale exome sequencing study implicates both developmental and
980 functional changes in the neurobiology of autism. *Cell* **180**, 568–584.e23.

981 Schröder, S., Steinmetz, N.A., Krumin, M., Pachitariu, M., Rizzi, M., Lagnado, L., Harris, K.D.,
982 Carandini, M. (2020). Arousal Modulates Retinal Output. *Neuron* **107**, 487-495.e9.

983 Shang, C., Chen, Z., Liu, A., Li, Y., Zhang, J., Qu, B., Yan, F., Zhang, Y., Liu, W., Liu, Z., Guo, X.,
984 Li, D., Wang, Y., Cao, P. (2018). Divergent midbrain circuits orchestrate escape and freezing
985 responses to looming stimuli in mice. *Nat. Commun.* **9**, 1232.

986 Sohn H, Narain D. (2021). Neural implementations of Bayesian inference. *Curr Opin Neurobiol*
987 **70** 121–129.

988 Spratt, P.W.E., Ben-Shalom, R., Keeshen, C.M., Burke, K.J., Clarkson, R.L., Sanders, S.J.,
989 Bender, K.J. (2019). The Autism-Associated Gene Scn2a Contributes to Dendritic Excitability and
990 Synaptic Function in the Prefrontal Cortex. *Neuron* **103**, 673–685.e5.

991 Suzuki, Y., Minami, T., Nakauchi, S. (2018). Association between pupil dilation and implicit
992 processing prior to object recognition via insight. *Sci. Rep.* **8**, 6874.

993 Tang, L., Higley, M.J. (2020). Layer 5 Circuits in V1 Differentially Control Visuomotor Behavior.
994 *Neuron* **105**, 346-354.e5.

995 Tinbergen, N. (1951). The study of instinct, The study of instinct. Clarendon Press/Oxford
996 University Press, New York, NY, US.

997 Turi, M., Burr, D.C., Igliozzi, R., Aagten-Murphy, D., Muratori, F., Pellicano, E. (2015). Children
998 with autism spectrum disorder show reduced adaptation to number. *Proc Natl Acad Sci U S A.*
999 **112**, 7868–7872.

1000 Turi, M., Karaminis, T., Pellicano, E., Burr, D. (2016). No rapid audiovisual recalibration in adults
1001 on the autism spectrum. *Sci Rep* **6**, 21756.

1002 van Boxtel, J.J., Dapretto, M., Lu, H. (2016). Intact recognition, but attenuated adaptation, for
1003 biological motion in youth with autism spectrum disorder. *Autism Res* **9**, 1103–1113.

1004 Walker, E.Y., Cotton, R.J., Ma, W.J., Tolias, A.S. (2020). A neural basis of probabilistic
1005 computation in visual cortex. *Nat Neurosci* **23**, 122–129.

1006 Wang, L., Sarnaik, R., Rangarajan, K., Liu, X., Cang, J. (2010). Visual Receptive Field Properties
1007 of Neurons in the Superficial Superior Colliculus of the Mouse. *J. Neurosci.* **30**, 16573–16584.

1008 Wang, C.-A., Boehnke, S.E., White, B.J., Munoz, D.P. (2012). Microstimulation of the Monkey
1009 Superior Colliculus Induces Pupil Dilation Without Evoking Saccades. *J. Neurosci.* **32**, 3629–
1010 3636.

1011 Wang, Q., Burkhalter, A. (2013). Stream-Related Preferences of Inputs to the Superior Colliculus
1012 from Areas of Dorsal and Ventral Streams of Mouse Visual Cortex. *J. Neurosci.* **33**, 1696–1705.

1013 Wei, P., Liu, N., Zhang, Z., Liu, X., Tang, Y., He, X., Wu, B., Zhou, Z., Liu, Y., Li, J., Zhang, Y.,
1014 Zhou, X., Xu, L., Chen, L., Bi, G., Hu, X., Xu, F., Wang, L. (2015). Processing of visually evoked
1015 innate fear by a non-canonical thalamic pathway. *Nat. Commun.* **6**, 6756.

1016 Yilmaz, M., Meister, M. (2013). Rapid Innate Defensive Responses of Mice to Looming Visual
1017 Stimuli. *Curr. Biol.* **23**, 2011–2015.

1018 Zhao, X., Liu, M., Cang, J. (2014). Visual Cortex Modulates the Magnitude but Not the Selectivity
1019 of Looming-Evoked Responses in the Superior Colliculus of Awake Mice. *Neuron* **84**, 202–213.

1020 Zingg, B., Chou, X., Zhang, Z., Mesik, L., Liang, F., Tao, H.W., Zhang, L.I. (2017). AAV-Mediated
1021 Anterograde Transsynaptic Tagging: Mapping Corticocollicular Input-Defined Neural Pathways
1022 for Defense Behaviors. *Neuron* **93**, 33–47.

Figures

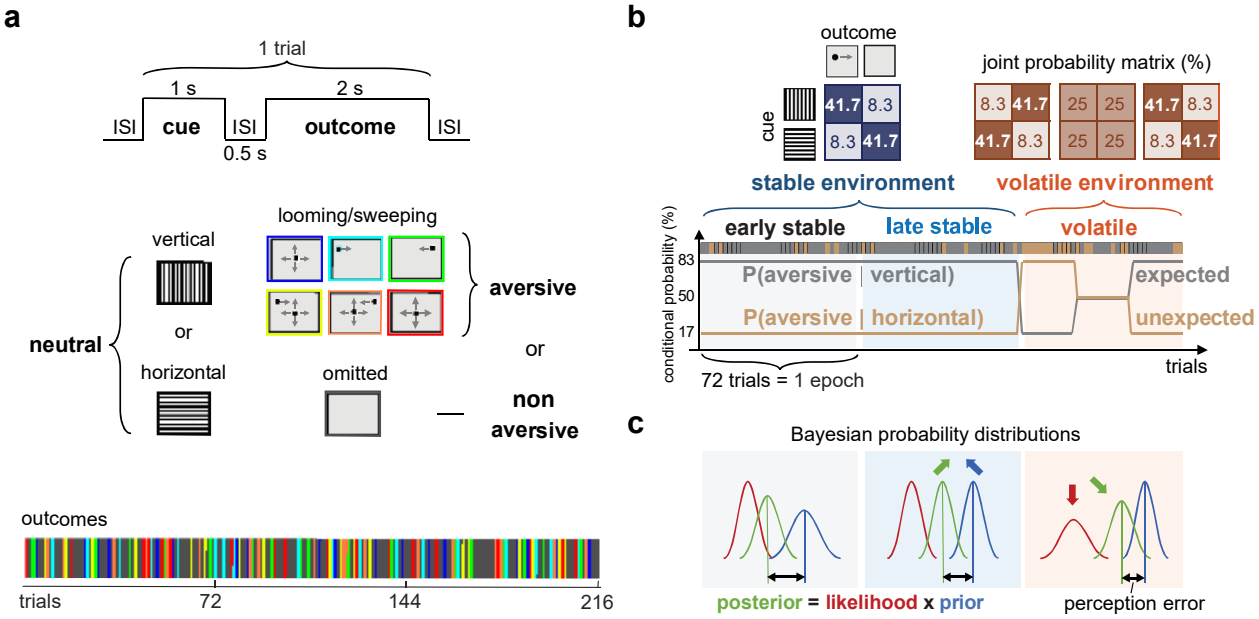


Figure 1: Implicit visual learning task to study adaptive sensory processing in mice. **a** Schematic of the implicit visual learning task. In each trial, a neutral cue (vertical or horizontal gratings; 1 s) is followed by an aversive outcome (looming and/or sweeping dark spot, randomly chosen from six different types; 2 s) or non-aversive one (omitted; 2 s) with an inter-stimulus interval (ISI) of 0.5 s. In total, the task consisted of 216 trials, including 108 aversive outcome trials (color-coded) and 108 non-aversive ones (dark gray). **b** The conditional probability of cue-outcome contingencies (expected, gray; unexpected, light brown) during stable and volatile environments. The stable environment is twice as long as the volatile one (orange), and divided into early- and late-stable epochs (72 trials each; light gray and light blue, respectively). **c** Animals are expected to form cue-outcome associations during the stable environment, leading to increased prior and posterior beliefs in the late-stable epoch. In the volatile epoch, however, such associations are violated, and the precision of the posterior decreases as the likelihood reduces.

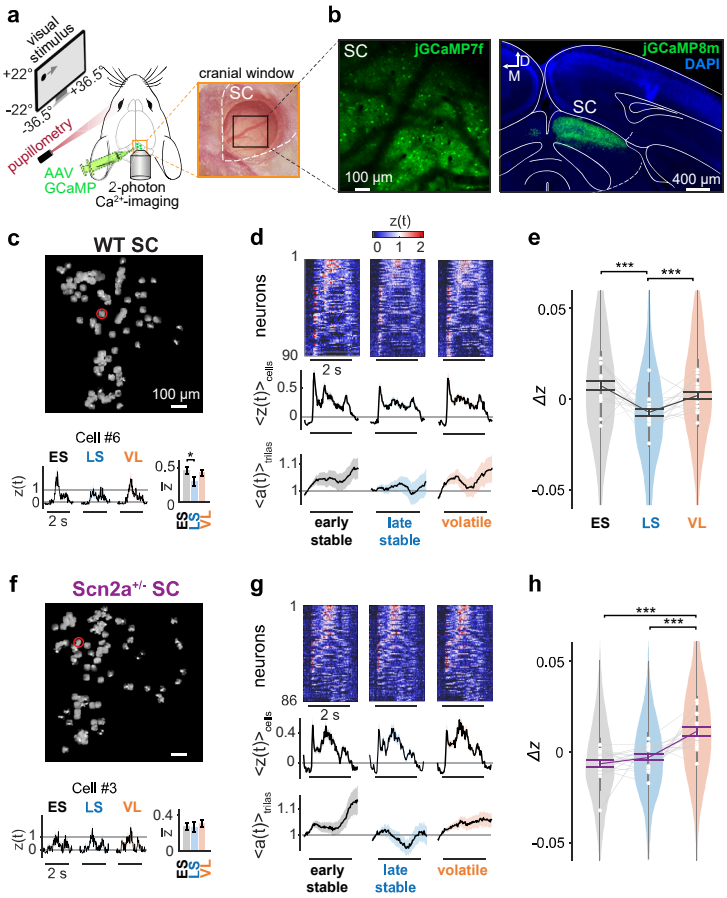


Figure 2: Neurons in the superior colliculus show contextual adaptation to a stable environment in wild-type mice but not in Scn2a$^{+/-}$ mice.

a Experimental setup. *Left*: The schematic illustrates a head-fixed mouse pre-injected with adeno-associated viruses (AAVs) carrying genetically encoded calcium indicators (e.g., jGCaMP7f). Visual stimulus was presented to the left eye whose pupil size was monitored with an infrared (IR) camera. *Right*: Two-photon calcium imaging was performed through a glass window implanted with a silicone plug (diameter, 1.5 mm) over the right superior colliculus (SC, outlined with white dashed line; black square, field of view). **b** *Left:* An example frame of motion-corrected two-photon calcium imaging in the SC of a wild-type (WT) mouse. *Right*: A coronal slice image taken two months after AAV injections, demonstrating a high expression of jGCaMP8m (green) in the right SC, along with 4',6-diamidino-2-phenylindole stain (DAPI, blue). **c** *Top*: Deconvoluted spatial components of SC neurons extracted from WT mouse data with CaImAn. *Bottom*: Representative aversive stimulus responses of an SC neuron (circled in red on top), averaged over 36 aversive outcome trials for each analysis epoch ($z(t) = \langle Z(t) \rangle_{\text{trials}}$; shaded area, standard error of the mean, s.e.m.). The average over the stimulus period ($\bar{z} = \langle z(t) \rangle_t$) and s.e.m. is shown on the right for each epoch (from left to right: early stable (ES), gray; late stable (LS), light blue; volatile (VL), orange). **d** Heatmap ($z(t)$; top) and the population average ($\langle z(t) \rangle_{\text{cells}}$; shaded area, s.e.m; middle) of the aversive stimulus responses in each epoch for the responsive SC neurons (sorted with $\max[z(t)] > 1$), recorded from a representative WT mouse (90 cells from **c**). Corresponding pupil dynamics are shown at the bottom ($\langle A(t) \rangle_{\text{trials}}$, average proportional change in pupil area; shaded area, s.e.m.). **e** Distributions of the normalized SC responses across epochs ($\Delta z = \bar{z} - \langle \bar{z} \rangle_{\text{epochs}}$) from all WT animals in the dataset (1,252 cells from 17 mice): violin plots, distribution of all cells; gray lines with open circles, average for each mouse; vertical bars, interquartile range; thick line graph with error bars, mean ± s.e.m.; ***, p<0.001; **, p<0.01; *, p<0.05 with Tukey-Kramer test for pairwise comparisons. See Methods and **Supplementary Table 1** for details. **f-h** Corresponding data for Scn2a$^{+/-}$ mice (**f**, simultaneously recorded SC cell location in the image plane and a representative cell's responses; **g**, representative data from all simultaneously recorded responsive cells (86 cells); **h**, population data, 1201 cells in total from 16 Scn2a$^{+/-}$ mice).

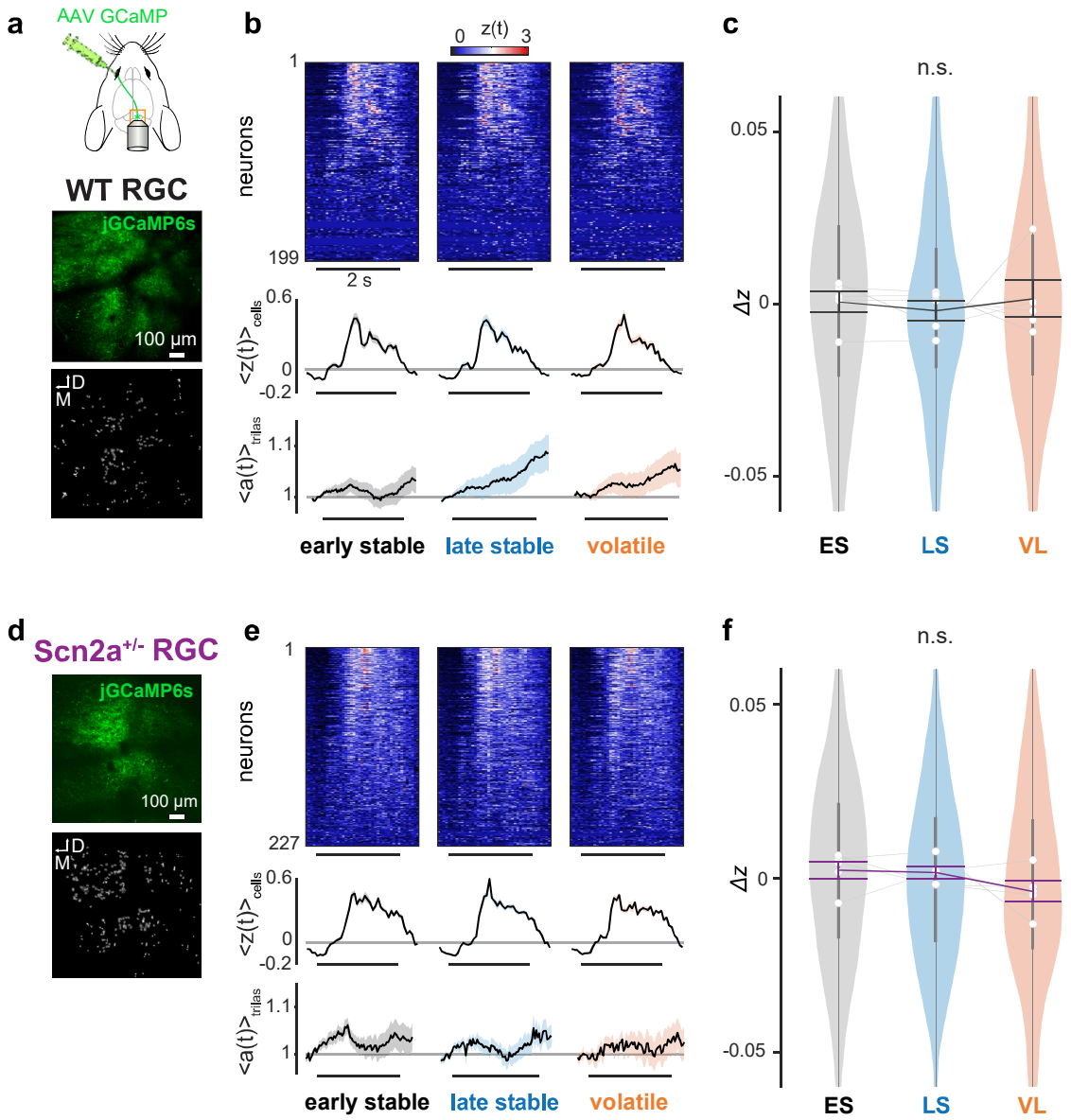


Figure 3: Axon terminals of retinal ganglion cells in the superior colliculus do not show any context-dependent modulation during the implicit learning task.
a Experimental setup. Animals were intraocularly injected with adeno-associated viruses (AAVs) carrying jGCaMP6s and underwent cranial window surgery over the right superior colliculus (SC) to monitor the activity of retinal ganglion cell (RGC) axons *in vivo* (top). An example frame of RGC axonal imaging in SC (middle) and deconvoluted spatial components by CaImAn (bottom). **b** Aversive stimulus responses of simultaneously recorded RGC putative boutons (top: heatmap of $z(t)$, sorted by $\max[z(t)]$), the population average (middle, $\langle z(t)\rangle_{\text{boutons}}$; 199 boutons in total from **a**; shaded area, s.e.m.), and pupil size dynamics (bottom, $\langle a(t)\rangle_{\text{trials}}$) from a representative wild-type (WT) mouse in each analysis epoch (shown in the same format as in **Fig. 2d,g**). **c**: Distributions of the normalized RGC population responses across epochs (Δz; violin plot, 877 boutons from 5 WT mice), shown in the same format as in **Fig. 2e,h**. No significant difference was found by pairwise comparisons (Tukey-Kramer test). **d-f**: Corresponding data for Scn2a$^{+/-}$ mice (**d**, example image frame of simultaneously recorded RGC axons and putative bouton locations; **e**, representative data of simultaneously recorded responses of RGC axons to the aversive stimuli, 227 putative boutons; **f**, population data, 1208 responsive boutons in total from 5 animals).

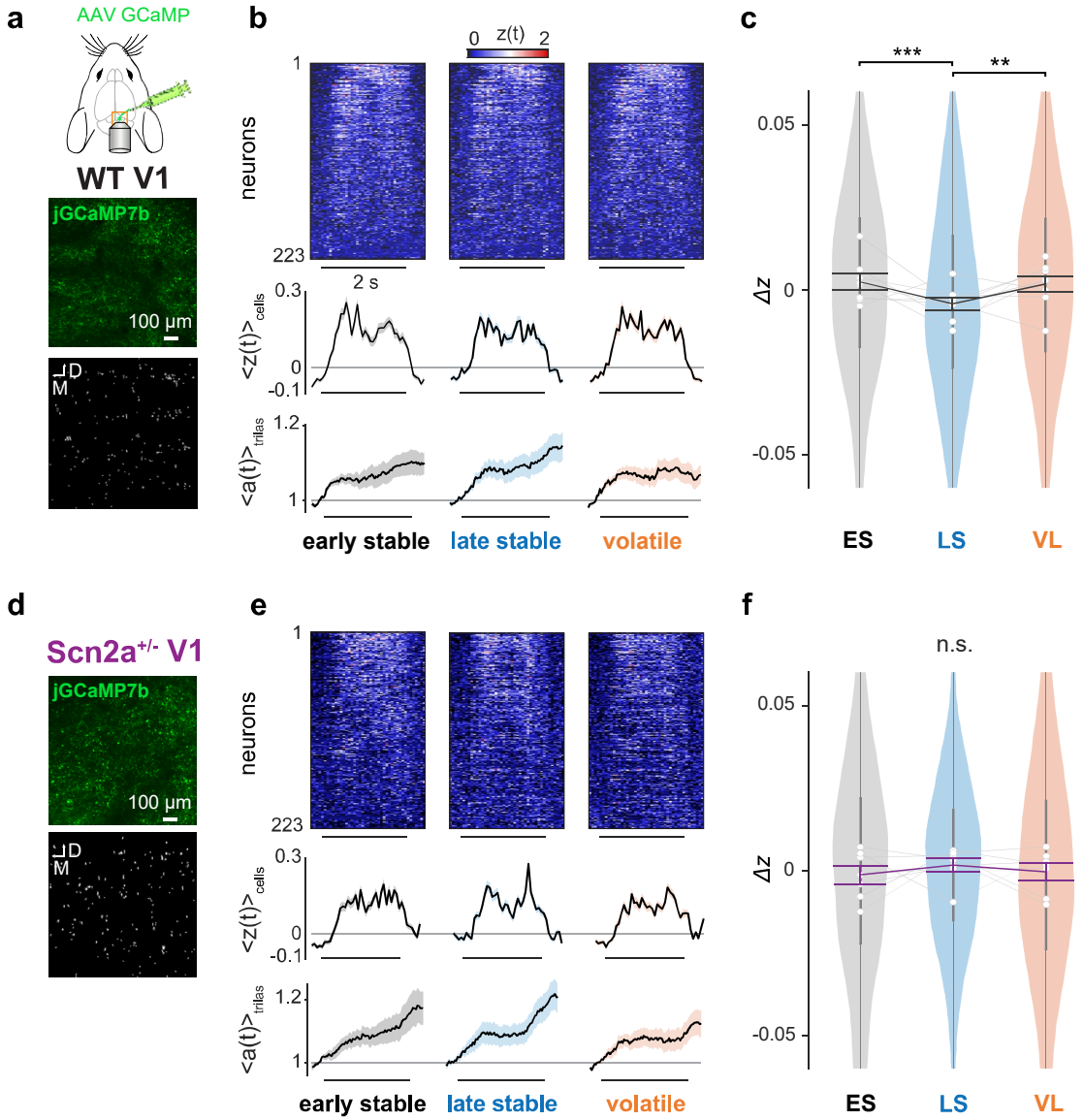


Figure 4: Axon terminals of primary visual cortical cells in the superior colliculus show similar context-dependent modulation to the collicular neurons in wild-type mice, but not in Scn2a$^{+/-}$ animals.
a Experimental setup. Animals were injected with adeno-associated viruses (AAVs) carrying GCaMP7b into the primary visual cortex (V1) and underwent cranial window surgery over the superior colliculus (SC) to monitor the axonal activity of V1 cells in vivo (middle, an example frame of V1 axonal imaging in SC; bottom, deconvoluted spatial components by CaImAn). **b** Aversive stimulus responses of simultaneously recorded putative boutons of V1 axons (top: heatmap of $z(t)$, sorted by $\max[z(t)]$), and the population average (middle, $\langle z(t)\rangle_{\text{boutons}}$; 223 boutons in total from **a**) and pupil size dynamics (bottom, $\langle a(t)\rangle_{\text{trials}}$) from a representative WT mouse in each analysis epoch (shown in the same format as in **Fig. 2d,g**). **c** Distributions of the normalized axonal responses across epochs (Δz; violin plot, 1376 boutons from 8 WT mice), shown in the same format as in **Fig. 2e,h**: ***, p<0.001; **, p<0.01 with Tukey-Kramer test for pairwise comparisons. **d-f**: Corresponding data for Scn2a$^{+/-}$ mice (**d**, example image frame of simultaneously recorded V1 axons projected to SC and putative bouton locations; **e**, representative data of simultaneously recorded responses of V1 axons to the aversive stimuli, 223 putative boutons; **f**, population data, 1254 responsive boutons in total from 7 mice).

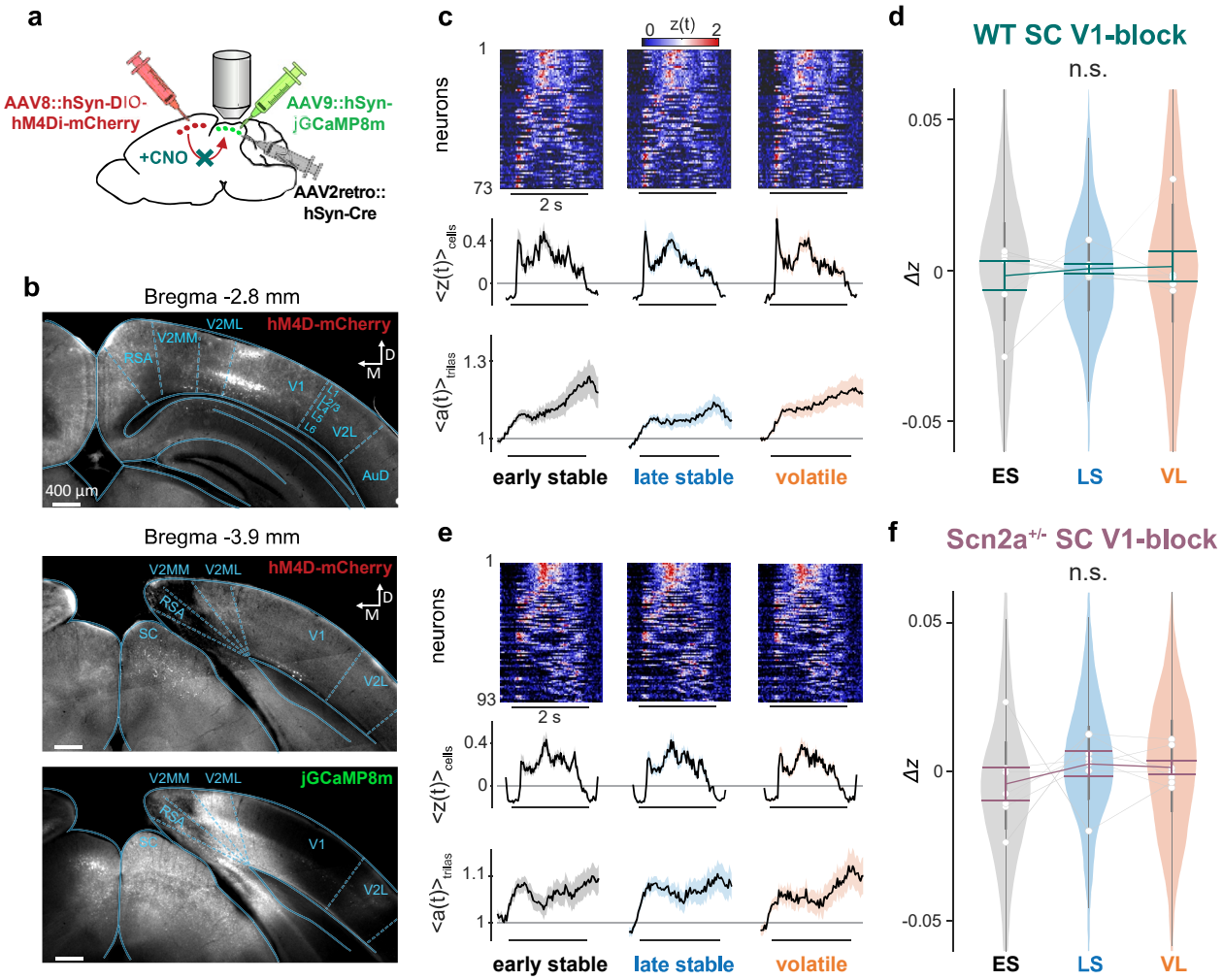


Figure 5: Inhibition of the corticotectal projection abolishes context-dependent modulation in the superior colliculus.
a Schematic of the viral vector strategy for chemogenetic inhibition of the projection from the primary visual cortex (V1) to the superior colliculus (SC). AAV8::hSyn-DIO-hM4Di-mCherry was injected into V1, while AAV2retro::hSyn-Cre and AAV9::hSyn-jGCaMP8m were co-injected into SC. Systemic administration of clozapine N-oxide (CNO) activated the hM4Di receptor, hence inhibiting the V1-to-SC projection. **b** Histological verification of the hM4Di receptor expression (with mCherry) in the V1 layer 5 (L5) cells (top) and their axons in SC (middle). Expression of hM4Di receptors was also observed in the V1 layer 2/3 cells (top), likely because they project to the secondary visual cortical area (V2MM) where we had a leakage of the AAV2retro virus from the injection needle during insertion and/or retraction process, as indicated by jGCaMP8 signals via co-injected virus (bottom). **c** Heatmap (top, $z(t)$, sorted by $\max[z(t)]$; 73 cells) and the population average (middle, $\langle z(t)\rangle_{\text{cells}}$) of the aversive stimulus responses in each epoch for the responsive SC neurons recorded from a representative wild-type (WT) mouse after chemogenetically blocking V1-to-SC projection. Corresponding pupil dynamics ($\langle a(t)\rangle_{\text{trials}}$) were shown in the bottom panels (as in **Fig. 2d,g**). **d** Distributions of the normalized SC responses across epochs, Δz, from all WT animals after chemogenetically blocking V1-to-SC projection (423 cells from 7 mice), shown in the same format as in **Fig. 2e,h**. No significant difference was found by pairwise comparisons (Tukey-Kramer test). **e,f** Corresponding data for Scn2a$^{+/-}$ mice with chemogenetic inhibition of V1-to-SC projections (**e**, representative data, 93 cells; **f**, population data, 550 cells in total from 7 mice).

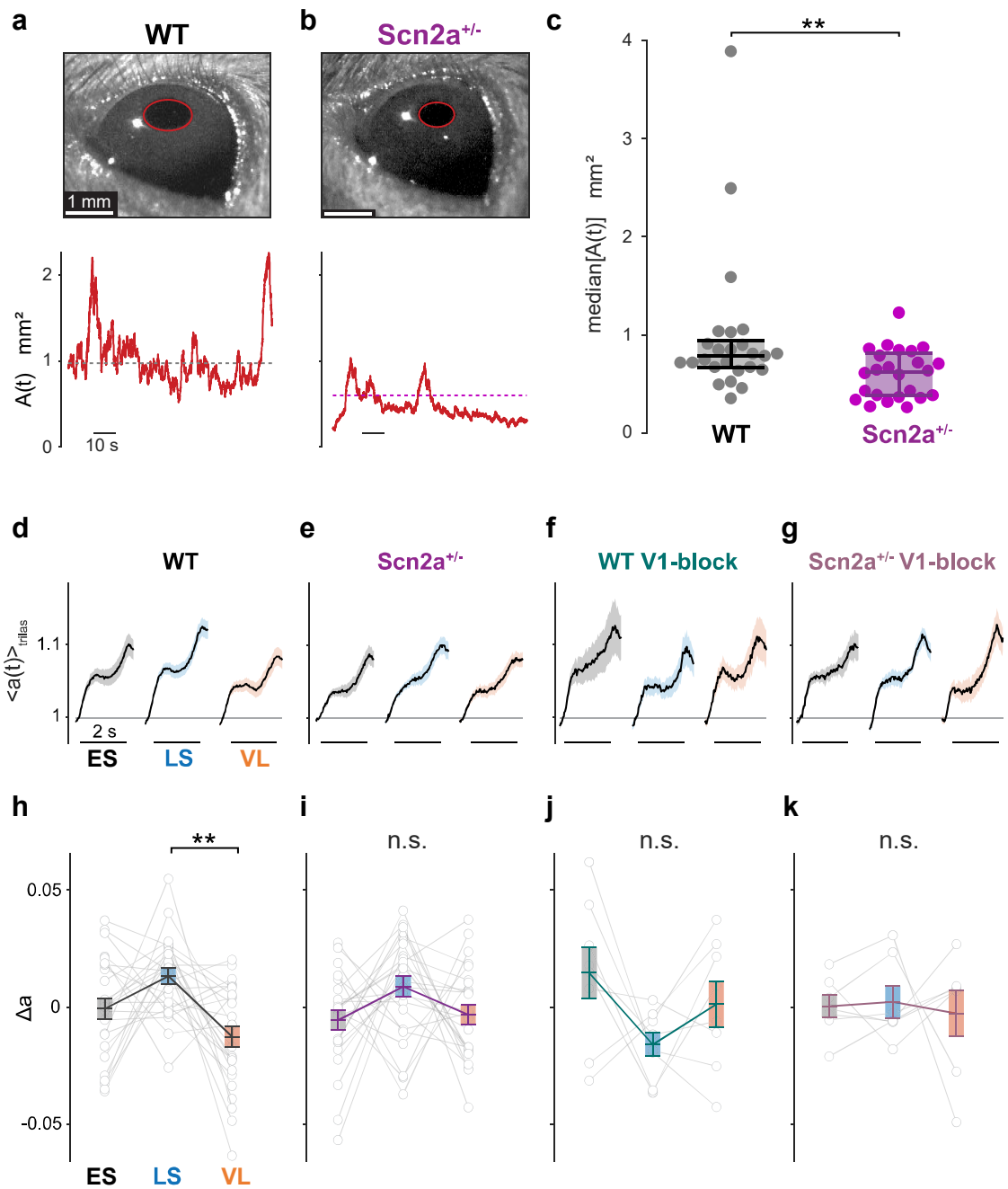


Figure 6: Volatility-dependent pupillometric response patterns in wild-type but not in Scn2a$^{+/-}$ mice. **a,b** Representative image frame of the left eye (top; red, pupil outline) of wild-type (WT, **a**) and Scn2a$^{+/-}$ mice (**b**) and the pupil area dynamics $A(t)$ during visual stimulation (bottom; dashed line, median). See also **Supplementary Fig. 1** for example data traces. **c** Median pupil size of 25 WT and 24 Scn2a$^{+/-}$ mice during the implicit visual task (error bar, median and interquartile range): **, p<0.01, Mann-Whitney U-test. **d-g** Pupil size dynamics of mice (average proportional change in pupil area, $\langle a(t)\rangle_{\text{mice}}$; shaded area, s.e.m.) in response to the aversive stimulus in each analysis epoch (from left to right: early stable, ES; late stable, LS; volatile, VL): **d**, 25 WT mice; **e**, 24 Scn2a$^{+/-}$ mice; **f**, 8 WT mice with chemogenetic inhibition of V1-to-SC projection (WT V1-block); and **g**, 7 Scn2a$^{+/-}$ mice with chemogenetic inhibition of V1-to-SC projection (Scn2a$^{+/-}$ V1-block). **h-k** Normalized change in pupil size upon aversive stimulus presentation in each analysis epoch ($\Delta a = \overline{a} - \langle\overline{a}\rangle_{\text{epochs}}$): thin line, individual animal; error bars, population average and s.e.m.; **h-k**, from corresponding groups in **d-k**, respectively. **, p<0.01 with Tukey-Kramer test for pairwise comparisons.

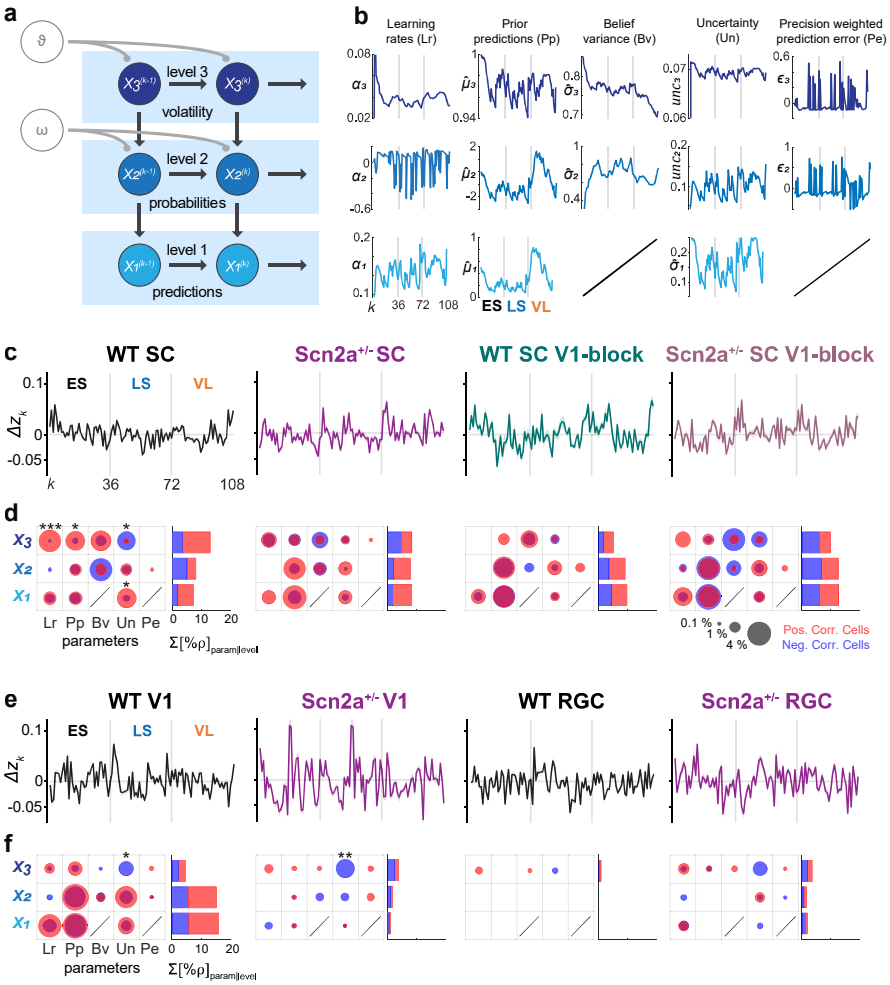


Figure 7: Inputs from the primary visual cortex carry much less information in Scn2a$^{+/-}$ mice than in wild-type mice, leading to misrepresentations of environmental volatility in the superior colliculus. **a** Schematic representation of the HGF model as a Bayesian inference network. The perceptual model tracks an animal's learning of the task structure across three levels. Level 1 (X_1) represents trial-wise stimulus transitions from one stimulus to the next; X_2 the probability of the transitions of contingencies; and X_3 the environment volatility, where k is the current trial number. Mice hold and update beliefs about the true quantities at each level, with a mean μ and a variance σ; ϑ and ω are mouse-specific parameters that couple the levels and determine the respective speed of belief updating about phasic volatility and transition contingencies. **b** HGF model parameter trajectories, generated with the cue-outcome contingencies of the implicit visual learning task. Non-aversive outcome trials were excluded to match neuronal trajectories (in **c**). Learning rates (Lr) at X_1-X_3; Prior predictions (Pp) at X_1-X_3; Believe variances (Bv) at X_2 and X_3; Outcome uncertainty (Un) at X_1; Information uncertainty at X_2; Volatility uncertainty at X_3; Precision-weighted prediction errors (Pe) at X_2 and X_3. See Methods for details. **c** Population average of the normalized neuronal trajectories for each experimental group ($\langle \Delta z_k^{\mathrm{res}} \rangle_{\mathrm{cells}}$). From left to right: WT SC, Scn2a$^{+/-}$ SC, WT SC with V1-block, and Scn2a$^{+/-}$ SC with V1-block. **d** Neural representation of HGF parameters for each experimental group (as in **c**), given as the percentage of the cells (left) whose response trajectory was significantly correlated with each HGF model parameter trajectory (from **b**), and the cumulative sum of those cell percentages at each level (right): circle area, percentage of cells with q<0.05 from Storey's FDR; red, positive correlation; blue, negative correlation. A bias of the correlation polarity for each HDR parameter was examined with a shuffling test with FDR correction: *, q<0.05; **, q<0.01. See also **Supplementary Table 2** for details. **e,f** Corresponding data for V1 and RGC axonal responses (**e**, population average of the normalized neuronal trajectories; **f**, neural representation of HGF parameters). From left to right, WT V1, Scn2a$^{+/-}$ V1, WT RGC, and Scn2a$^{+/-}$ RGC.

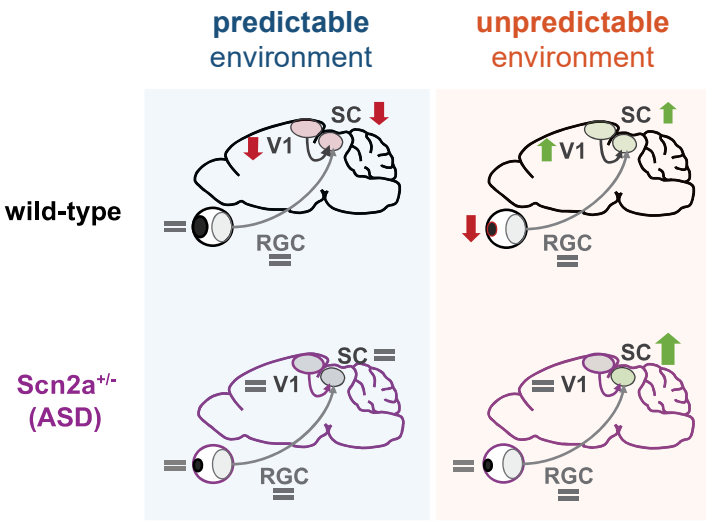


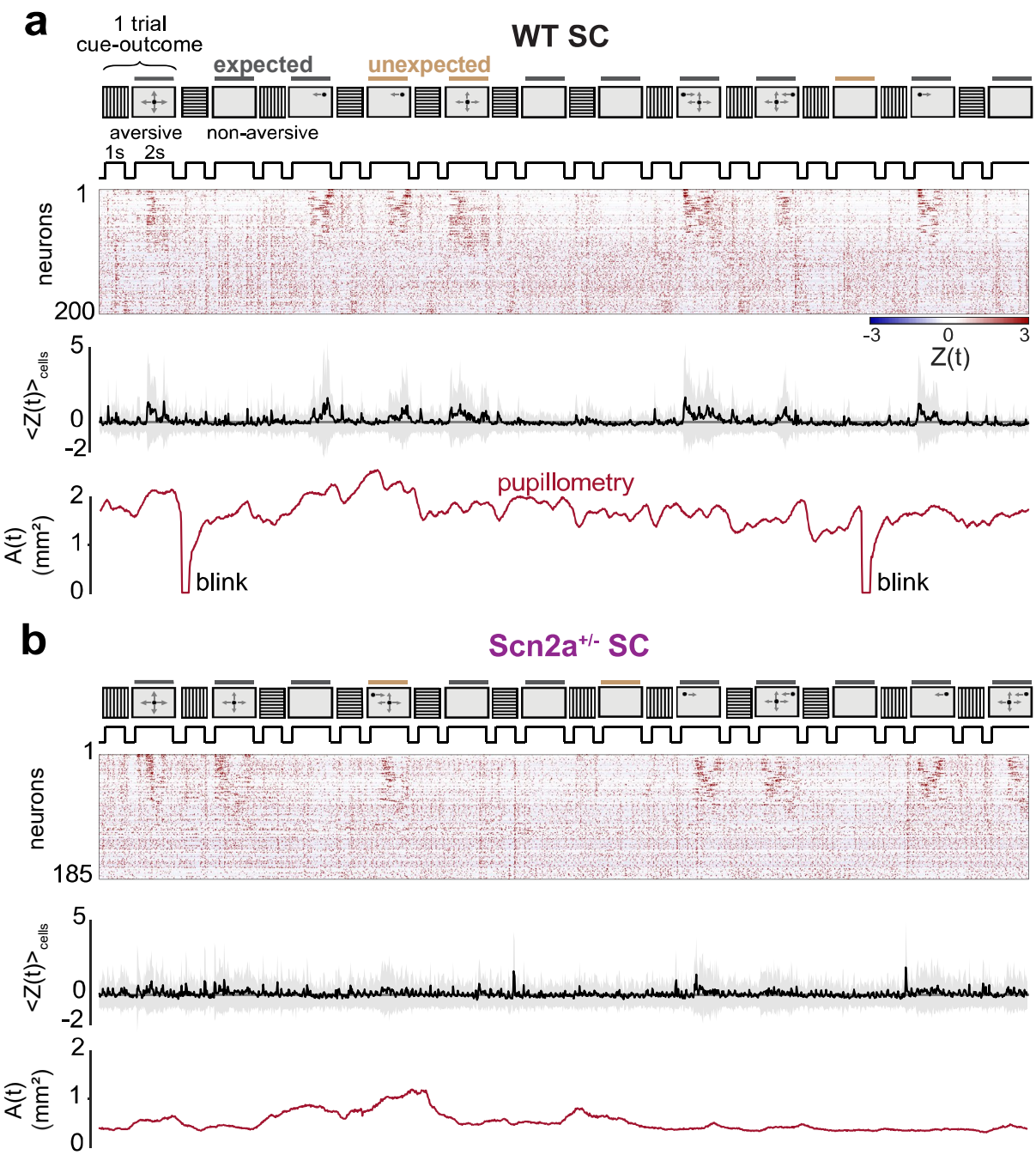
Figure 8: Summary diagram on anomalies in the early visual system of Scn2a$^{+/-}$ mice during implicit visual learning. In WT mice, inputs from V1 provide critical signals for stimulus-specific context-dependent modulation of SC dynamics during implicit visual learning (down arrow, suppression; up arrow, enhancement). A milder behavioral response, such as associated changes in pupil dynamics, will follow that reflect an animal's stimulus pattern recognition. In Scn2a$^{+/-}$ mice, defective V1 inputs lead to anomalous modulation patterns in SC.

Atypical cortical feedback underlies failure to process contextual information in the superior colliculus of Scn2a$^{+/-}$ autism model mice

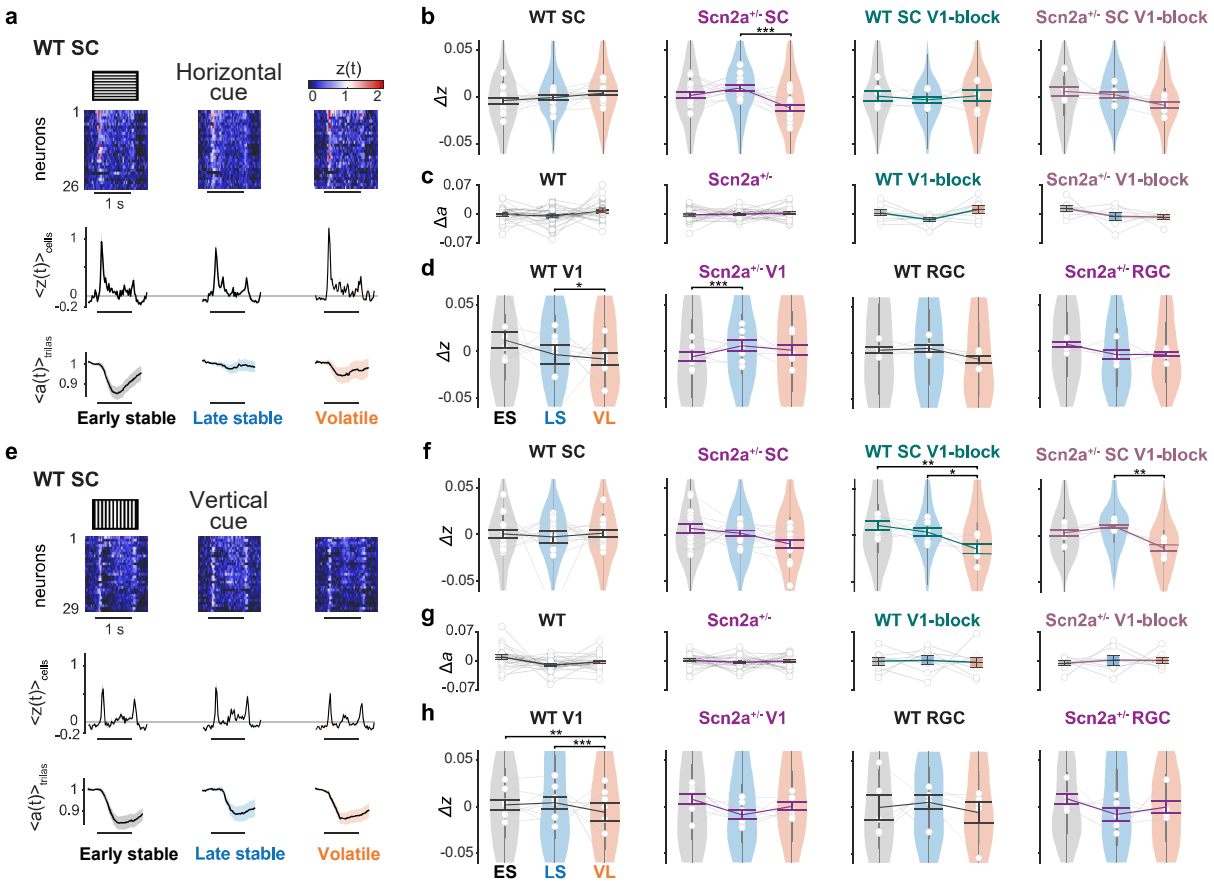
Leiron Ferrarese, Hiroki Asari

Supplementary Information

- Supplementary Figure 1
- Supplementary Figure 2
- Supplementary Figure 3
- Supplementary Figure 4
- Supplementary Figure 5
- Supplementary Figure 6
- Supplementary Table 1
- Supplementary Table 2

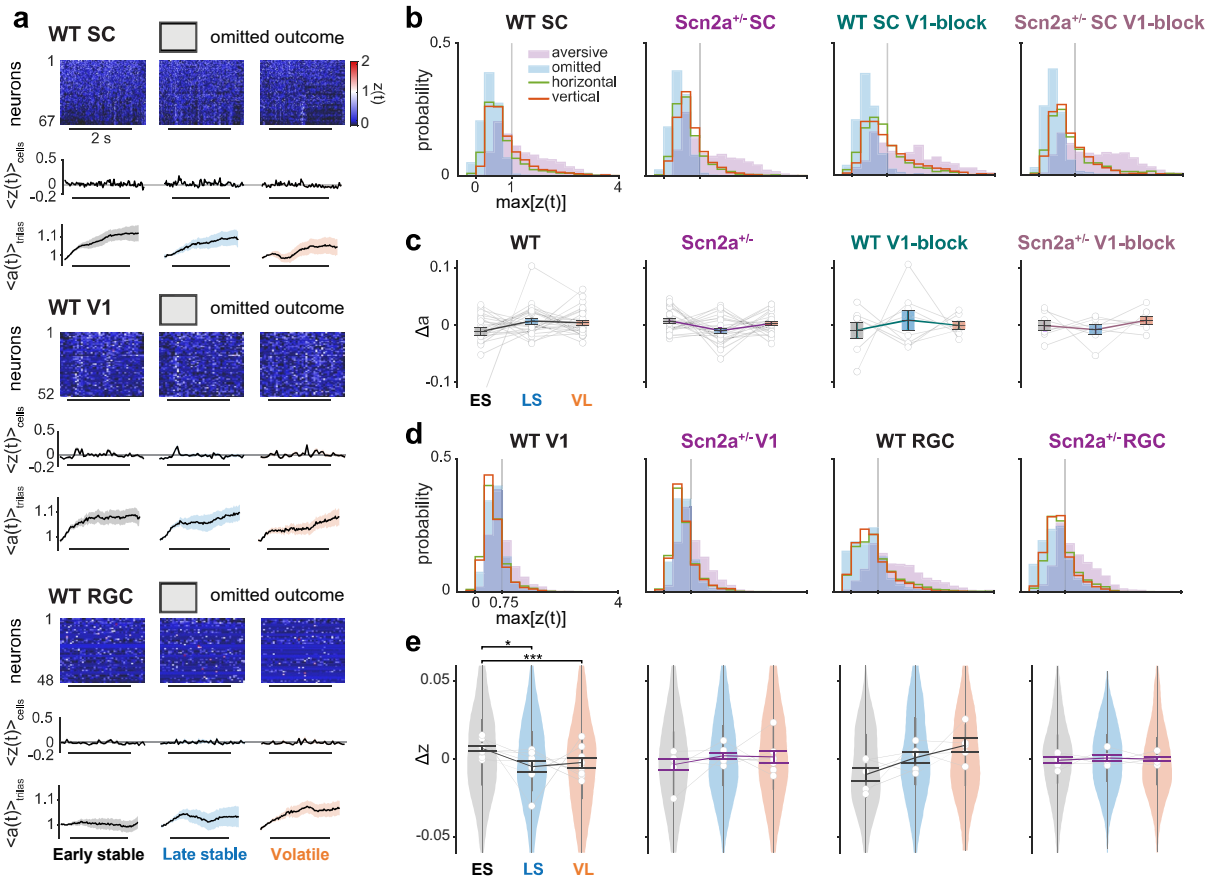


Supplementary Figure 1: Representative neuronal activity of the mouse superior colliculus and pupil dynamics in response to the cue-outcome stimulus sequences. **a** Trial-by-trial response example obtained from a wild-type (WT) mouse during the early stable (ES) epoch, with 83% expected and 17% unexpected outcomes. *From top to bottom*: The normalized spiking activity of the superior colliculus (SC) neurons $Z(t)$ (z-score), estimated from recorded calcium dynamics; corresponding population average response $\langle Z(t) \rangle_{\text{cells}}$ (gray shade, ±1 standard deviation); and concurrent dynamics of the estimated pupil size $A(t)$ in mm^2. **b** Corresponding data from a representative Scn2a$^{+/-}$ mouse.



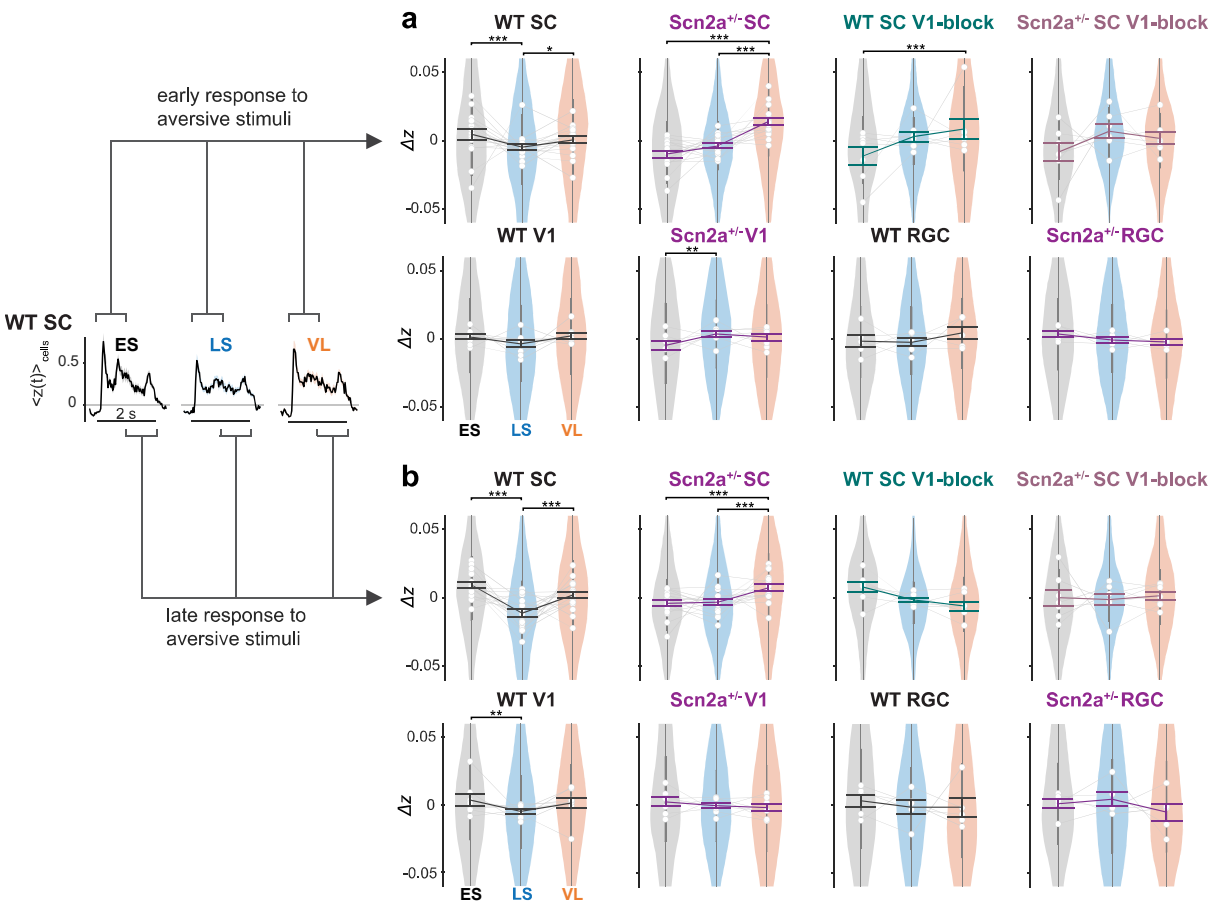
Supplementary Figure 2: Contextual modulation patterns of the cue stimulus responses in the mouse early visual system.

Here and thereafter, ***, p<0.001; **, p<0.01; *, p<0.05 with Tukey-Kramer test for pairwise comparisons (see Supplementary Table 1 and Methods for details). **a** Heatmap ($z(t)$; sorted with $\max[z(t)]$; top) and the population average ($\langle z(t)\rangle_{\text{cells}}$; shaded area, standard error of the mean (s.e.m.); middle) of the responses to one of the cue stimuli (horizontal gratings) in each epoch for SC neurons of a representative WT mouse. Corresponding pupil dynamics are shown at the bottom ($\langle A(t)\rangle_{\text{trials}}$, average proportional change in pupil area; shaded area, s.e.m.). **b** Distributions of the normalized horizontal grating responses (Δz) of SC cells in WT or Scn2a$^{+/-}$ mice, with or without chemogenetically blocking the projections from the primary visual cortex (V1), across epochs. A decrease of the population response strength was observed in the SC of Scn2a$^{+/-}$ mice in the volatile (VL) epoch. **c** Normalized change in pupil size (Δa) upon horizontal grating stimulus presentations (from left to right, 25 WT mice, 24 Scn2a$^{+/-}$ mice, 8 WT mice with V1-block, 7 Scn2a$^{+/-}$ mice with V1-block). No contextual modulation was observed across epochs. **d** Distributions of the normalized horizontal grating responses of V1 or retinal ganglion cell (RGC) axons in WT or Scn2a$^{+/-}$ mice. No modulation was found in RGC responses in both genotypes, while reduction of the response strength across epochs was observed for V1 axons. **e-h** corresponding figure panels in response to the other cue stimulus (vertical gratings). In WT mice, no modulation was observed in RGC or pupil dynamics, while V1 axons showed adaptation, consistent with the responses to the vertical grating stimuli. In Scn2a$^{+/-}$ mice, no contextual modulation was identified.

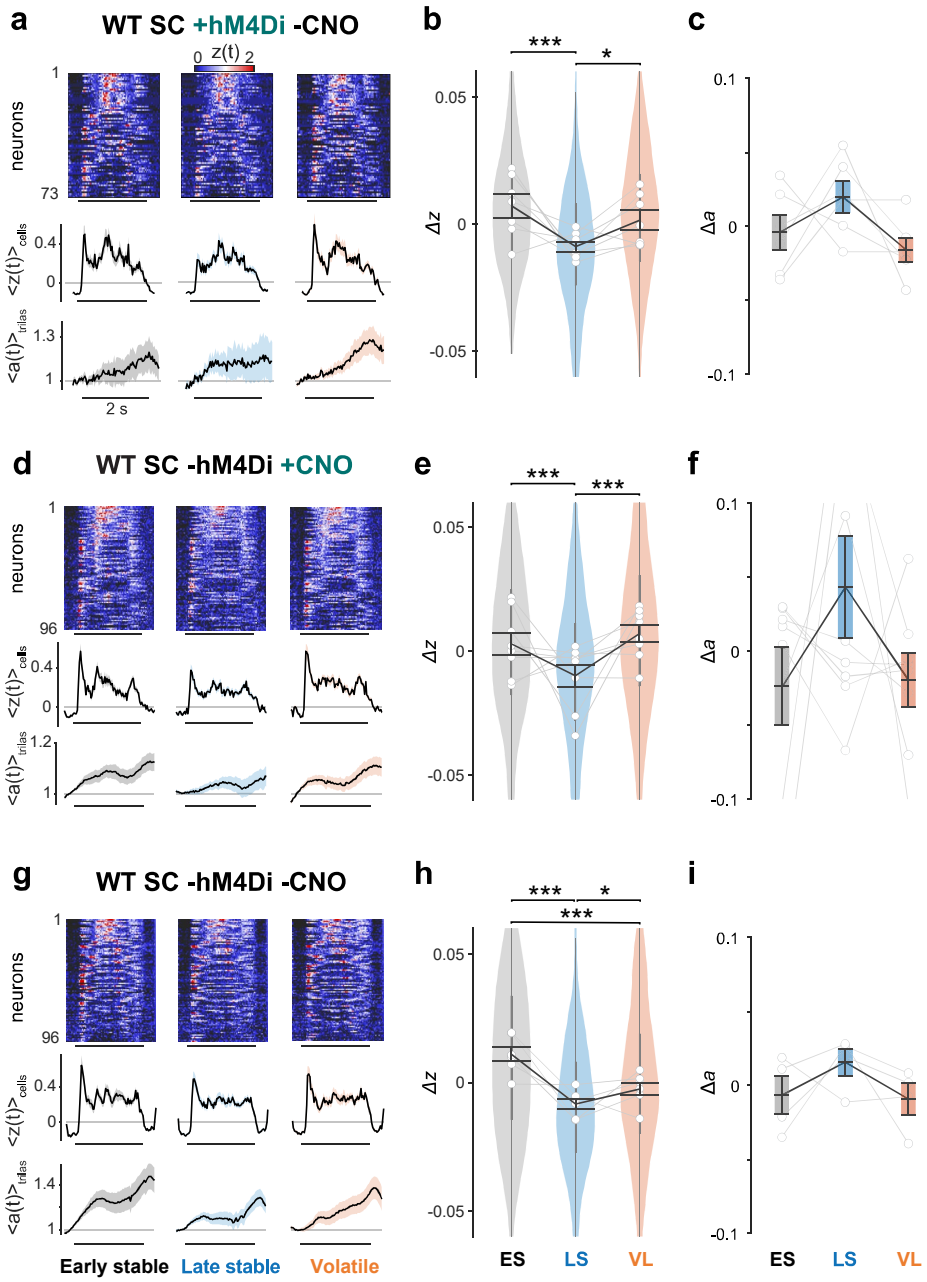


Supplementary Figure 3: no contextual modulation was observed in response to the omitted stimulus.

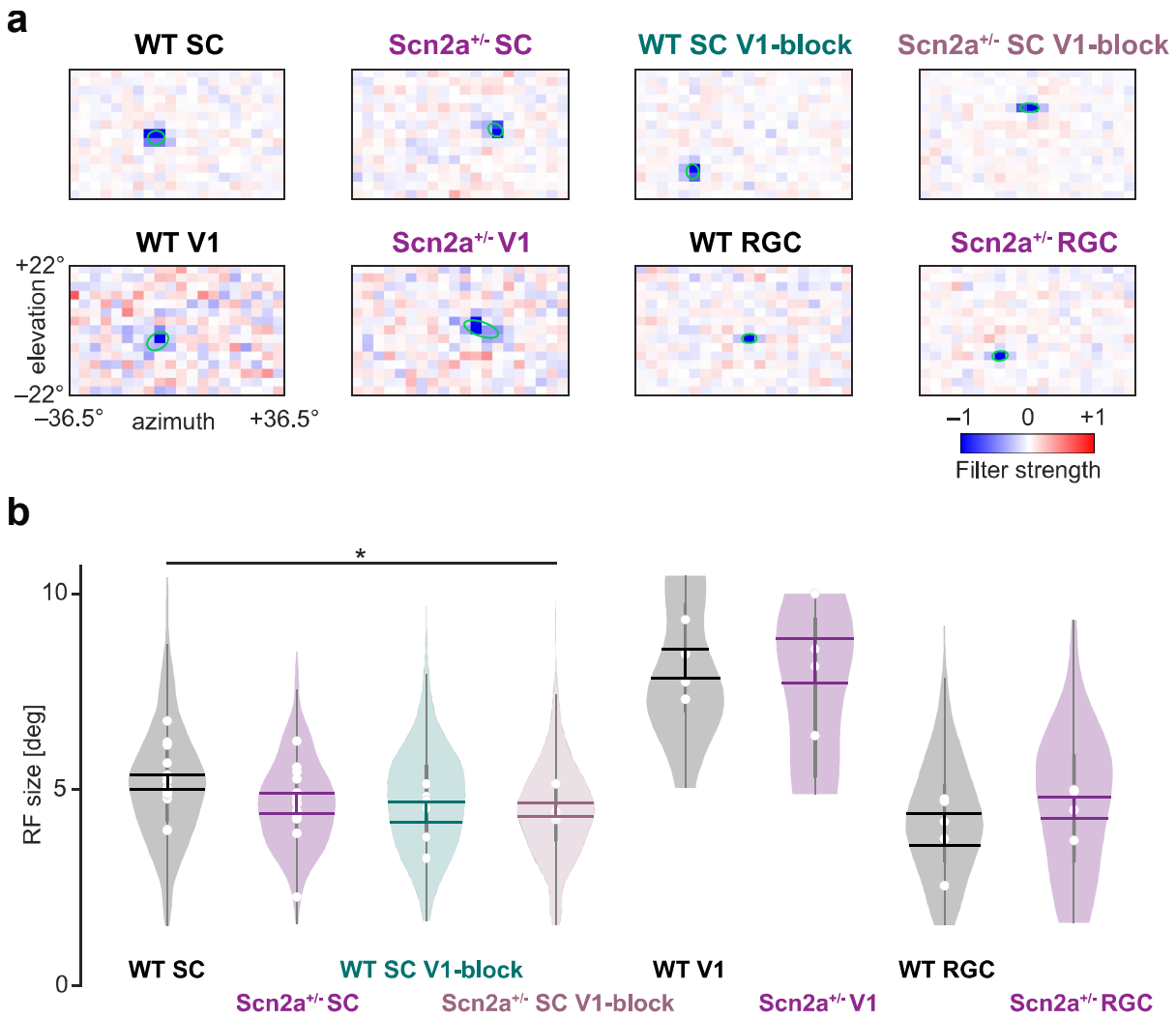
a Heatmap ($z(t)$; sorted with $\max[z(t)]$; top) and the population average ($\langle z(t)\rangle_{\text{cells}}$; shaded area, s.e.m.; middle) of the omitted stimulus responses in each epoch for SC neurons (top), V1 axons (middle), and RGC axons (bottom), each along with the pupil dynamics ($\langle\langle A(t)\rangle_{\text{trials}}$; shaded areas, s.e.m.) from the representative WT mice. **b** Probability distributions of SC responses ($\max[z(t)]$) to each stimulus type in WT or Scn2a$^{+/-}$ mice, with or without chemogenetically blocking V1 input. SC cells showed strong responses to the aversive stimuli (purple), modest responses to the cue stimuli (horizontal gratings, olive; vertical gratings, tangerine), and virtually no response to the non-aversive, omitted stimulus (blue) during the implicit visual learning paradigm. Hence, no comparison was made for the omitted response strength across epochs. **c** Normalized change in pupil size ($\Delta a = \overline{a} - \langle\overline{a}\rangle_{\text{epochs}}$) upon non-aversive omitted stimulus presentation in each epoch (from left to right: 25 WT mice, 24 Scn2a$^{+/-}$ mice, 8 WT mice with chemogenetic inhibition of V1-to-SC projections (V1-block), and 7 Scn2a$^{+/-}$ mice with V1-block): thin line, individual animal; error bars, population average and s.e.m.. **d** Probability distribution of V1 and RGC axonal responses to each stimulus type presented during the implicit learning task in WT and Scn2a$^{+/-}$ mice. Unlike SC cells, V1 and RGC axons demonstrated modest responses to the omitted stimulus. **e** Distributions of the normalized responses of V1 or RGC axons in WT or Scn2a$^{+/-}$ mice across epochs ($\Delta z = \overline{z} - \langle\overline{z}\rangle_{\text{epochs}}$). No contextual modulation was identified (Tukey-Kramer test for pairwise comparisons; see Supplementary Table 1 for details).



Supplementary Figure 4: Contextual-modulation was stronger for later response components in WT mice, but for earlier response components in Scn2a$^{+/-}$ mice.
We divided the aversive stimulus presentation period into the first (**a**; early responses) and second (**b**; late responses) halves in time (1 s each), and respectively analyzed the changes in the response strength across epochs. The response patterns were largely consistent between the two time windows, but the modulation was relatively stronger for the later components of the population SC responses and V1 inputs in WT mice. In contrast, the modulation was rather stronger for the earlier response components in Scn2a$^{+/-}$ mice, or after chemogenetically blocking V1 inputs to SC in WT mice. No significant modulation was found in RGC axonal responses (Tukey-Kramer test for pairwise comparisons; see Supplementary Table 1 for details).



Supplementary Figure 5: hM4Di expression in V1 cells projecting to SC or CNO injection alone had no effects on context-dependent modulation of SC cells in WT mice.
a Aversive stimulus responses of simultaneously recorded SC neurons in a representative WT mouse expressing hM4Di receptors in V1 cells projecting to SC (top, heatmap $z(t)$ of 73 cells, sorted with $max[z(t)] > 1$; middle, population average $\langle z(t)\rangle_{\text{cells}}$; bottom, corresponding pupil dynamics $\langle a(t)\rangle_{\text{trials}}$; shown in the same format as Fig. 2d,g). **b** Distributions of the normalized SC responses across epochs (Δz) from WT animals expressing hM4Di in V1-to-SC projection (405 cells from 7 mice; shown in the same format as Fig. 2e,h). The data from the same animals after CNO systemic injection are shown in Fig. 5c,d. **c** Normalized change in pupil size ($\Delta a = \overline{a} - \langle\overline{a}\rangle_{\text{epochs}}$) from the same animals as in **b** (6 mice; thin line, individual animal; error bars, population average with s.e.m.). The data from the same animals after CNO systemic injection are shown in Fig. 6j**. d-f** Corresponding data for WT mice without hM4Di expression but with CNO injection (**c**, representative data from simultaneously recorded SC cells, 96 cells; **d**, population data for WT mice with CNO injection, 573 cells in total from 8 mice. **f**, population data for changes in pupil size as in **c**, 8 mice). **g-i** Corresponding data for a subset of mice shown in **d-f**, assessed without the systemic injection of CNO (**g**, representative data from simultaneously recorded SC cells, 96 cells; **h**, population data, 486 cells in total from 6 mice; **i**, population data for change in pupil size as in **c** (4 mice). This is a subset of data shown in **Fig. 2e**.



Supplementary Figure 6: The early visual systems of WT and Scn2a$^{+/-}$ mice have similar receptive field sizes.

a Representative spatial receptive field (RF, estimated by a statistical ensemble approach; see Methods for details) of SC cells (top; from left to right, WT, Scn2a$^{+/-}$, WT with V1-block, and Scn2a$^{+/-}$ with V1-block), V1 cells projecting to SC (bottom left) and RGCs (bottom right), respectively. The green ellipse shows a contour of Two-dimensional Gaussian curve fit at 1 standard deviation. **b** Distribution of RF sizes. No substantial difference was found between genotypes (except for between SC cells in WT mice and SC cells in Scn2a$^{+/-}$ mice with V1-block): open circle, average for each mouse; vertical thick gray bar, interquartile range; horizontal lines, mean ± s.e.m.; *, $p<0.05$ with Tukey-Kramer test for pairwise comparisons. WT SC, 5.2±0.2; Scn2a$^{+/-}$ SC, 4.7±0.3; WT SC with V1-block, 4.4±0.3; Scn2a$^{+/-}$ SC with V1-block, 4.5±0.2; WT V1, 8.2±0.4; Scn2a$^{+/-}$ V1, 8.3±0.6; WT RGC, 4.0±0.4; Scn2a$^{+/-}$ RGC, 4.5±0.3 degrees.

Supplementary Table 1: FDR-adjusted p-values for contextual modulation analysis (Tukey-Kramer test). Texts in bold show those below a significance threshold (0.05).

genotype	chemogenetics	data type	stimulus type	p(ES-LS)	p(ES-VL)	p(LS-VL)	# cells	# animals	relevant figure #
WT	-	RGC	horizontal	1	0.65	0.057	304	5	S2d
WT	-	RGC	vertical	0.26	0.99	0.64	234	5	S2h
WT	-	RGC	omitted	0.9	0.99	1	138	5	S3e
WT	-	RGC	aversive early late	0.99 0.99 0.75	0.99 0.38 0.95	0.94 0.81 0.99	877	5	3c S4a S4b
WT	-	SC	horizontal	0.99	0.6	1	449	15	S2b
WT	-	SC	vertical	1	1	1	572	15	S2f
WT	-	SC	aversive early late	**0** **4.2e-06** **0**	0.27 0.99 0.79	**8.8e-17** **0.04** **3.2e-15**	1252	17	2e S4a S4b
WT	-	V1	horizontal	1	1	**0.02**	654	8	S2d
WT	-	V1	vertical	0.99	**5.6e-3**	**2.7e-4**	279	8	S2h
WT	-	V1	omitted	**0.02**	**1.8e-9**	0.3	249	8	S3e
WT	-	V1	aversive early late	**2.4e-5** 0.5 **2e-3**	1 0.99 0.99	**5.1e-3** 0.66 0.35	1376	8	4c S4a S4b
WT	-	pupil	horizontal	0.99	0.91	0.5	n/a	25	S2c
WT	-	pupil	vertical	0.065	0.76	0.98	n/a	25	S2g
WT	-	pupil	omitted	0.4	0.7	1	n/a	25	S3c
WT	-	pupil	aversive	0.5	0.7	**0.001**	n/a	25	6h
WT	hM4Di+CNO	SC	horizontal	1	1	1	168	6	S2b
WT	hM4Di+CNO	SC	vertical	0.99	**1e-4**	**0.032**	215	6	S2f
WT	hM4Di+CNO	SC	aversive early late	1 0.2 0.93	0.6 **3.3e-10** 0.4	0.98 0.13 0.99	423	7	5d S4a S4b
WT	hM4Di	SC	aversive	**2.9e-6**	0.99	**0.02**	405	7	S5b
WT	CNO saline	SC	aversive	**3.9e-9** **0**	0.055 **5.6e-5**	**0** **0.01**	573 486	8 6	S5e S5h
WT	hM4Di+CNO	pupil	horizontal	0.9	0.99	0.35	n/a	8	S2c
WT	hM4Di+CNO	pupil	vertical	1	1	1	n/a	8	S2g
WT	hM4Di+CNO	pupil	omitted	0.98	1	0.99	n/a	8	S3c
WT	hM4Di+CNO	pupil	aversive	0.15	1	0.9	n/a	8	6j
Scn2a$^{+/-}$	-	RGC	horizontal	**3.3e-4**	0.8	0.6	335	5	S2d
Scn2a$^{+/-}$	-	RGC	vertical	0.12	0.73	0.99	254	5	S2h
Scn2a$^{+/-}$	-	RGC	omitted	0.2	1	0.99	444	5	S3e
Scn2a$^{+/-}$	-	RGC	aversive early late	0.99 0.99 0.99	0.75 0.15 0.99	0.99 0.96 0.99	1208	5	3f S4a S4b
Scn2a$^{+/-}$	-	SC	horizontal	0.6	1	**4.8e-7**	424	15	S2b
Scn2a$^{+/-}$	-	SC	vertical	1	0.6	0.7	464	15	S2f
Scn2a$^{+/-}$	-	SC	aversive early late	0.29 0.17 0.99	**0** **0** **1.7e-11**	**0** **0** **3.2e-8**	1201	16	2h S4a S4b
Scn2a$^{+/-}$	-	V1	horizontal	**3.3e-4**	0.8	0.6	257	7	S2d
Scn2a$^{+/-}$	-	V1	vertical	0.12	0.73	0.99	252	7	S2h
Scn2a$^{+/-}$	-	V1	omitted	0.2	1	0.99	494	7	S3e
Scn2a$^{+/-}$	-	V1	aversive early late	0.99 **0.0048** 0.99	0.99 0.93 0.19	0.85 0.99 0.99	1254	7	4f S4a S4b
Scn2a$^{+/-}$	-	pupil	horizontal	0.99	0.99	1	n/a	24	S2c
Scn2a$^{+/-}$	-	pupil	vertical	0.99	1	0.99	n/a	24	S2g
Scn2a$^{+/-}$	-	pupil	omitted	0.5	1	0.9	n/a	24	S3c
Scn2a$^{+/-}$	-	pupil	aversive	0.4	1	0.7	n/a	24	6i
Scn2a$^{+/-}$	hM4Di+CNO	SC	horizontal	1	0.98	0.9	208	6	S2b
Scn2a$^{+/-}$	hM4Di+CNO	SC	vertical	0.99	0.65	**0.0088**	192	6	S2f
Scn2a$^{+/-}$	hM4Di+CNO	SC	aversive early late	0.99 0.09 0.99	0.99 0.98 0.99	1 0.99 0.99	550	7	5f S4a S4b
Scn2a$^{+/-}$	hM4Di+CNO	pupil	horizontal	0.85	0.74	1	n/a	7	S2c
Scn2a$^{+/-}$	hM4Di+CNO	pupil	vertical	1	1	1	n/a	7	S2g
Scn2a$^{+/-}$	hM4Di+CNO	pupil	omitted	1	1	0.99	n/a	7	S3c
Scn2a$^{+/-}$	hM4Di+CNO	pupil	aversive	1	1	1	n/a	7	6k

Supplementary Table 2: **Correlation analysis results between neuronal and HGF model parameter dynamics, related to Figure 7.** The number of cells significantly correlated with each of the 13 model parameters (in columns) is shown in each experimental group (in rows): top, positive correlation; middle, negative correlation; bottom below dotted line, FDR-adjusted p-value from a shuffling test against a polarity bias: blue, significantly more negatively correlated cells; red, significantly more positively correlated cells.

genotype	Area (Total *n*) q-value	Lr1	Lr2	Lr3	Pp1	Pp2	Pp3	Bv2	Bv3	Un1	Un2	Un3	Pe2	Pe3	Subtotal: *n*(R>>0) *n*(R<<0)	% total corr. cells
WT	SC	13	0	37	14	11	29	10	31	29	10	2	1	0	187	28%
	(1080)	7	1	1	9	8	3	40	9	4	6	26	0	0	114	
	15 mice	0.6	0.6	7e-4	0.6	0.6	0.02	0.6	0.2	0.03	0.6	0.01	0.6	0.6		
	SC V1-block	5	0	0	11	12	2	1	7	2	5	1	1	0	48	23%
	(343)	2	0	0	12	9	1	1	4	1	1	2	1	0	33	
	6 mice	1	1	1	1	1	1	1	1	1	1	1	1	1		
	V1	51	0	10	61	77	19	7	0	27	46	0	1	2	301	38%
	(1254)	19	3	2	47	45	5	8	1	10	17	21	1	0	179	
	8 mice	0.1	0.6	0.6	0.6	0.4	0.3	0.6	0.6	0.4	0.2	0.01	0.6	0.6		
	RGC	0	0	3	0	0	0	0	1	0	0	0	0	0	4	0.6%
	(998)	0	0	0	0	0	0	0	0	0	0	2	0	0	2	
	5 mice	1	1	1	1	1	1	1	1	1	1	1	1	1		
Scn2a$^{+/-}$	SC	14	0	17	42	37	9	10	7	14	14	6	0	1	171	24%
	(1060)	0	0	13	14	8	11	13	20	3	1	4	0	0	87	
	15 mice	0.07	0.6	0.6	0.2	0.07	0.6	0.6	0.6	0.5	0.2	0.7	0.7	0.7		
	SC V1-block	11	1	1	15	14	4	1	3	4	8	2	1	0	72	34%
	(450)	3	2	7	20	18	2	7	16	3	3	8	0	0	82	
	6 imce	1	1	1	1	1	1	1	1	1	1	1	1	1		
	V1	0	0	7	2	2	2	0	2	1	0	0	5	3	24	5%
	(1376)	5	0	0	1	1	0	6	0	1	6	31	0	0	51	
	7 mice	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.003	0.9	0.9		
	RGC	9	0	10	0	0	3	0	3	0	8	0	0	1	34	6%
	(1208)	8	3	3	0	0	2	0	0	3	3	18	1	0	41	
	5 mice	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.9	0.1	0.9	0.9		