Integration of Audiovisual Danger Stimuli in the Zebrafish Tectum is Linked with Premotor and Behavioral Enhancements

Nicolás Martorell[1] and Violeta Medan[1, 2]

[1] Instituto de Fisiología, Biología Molecular y Neurociencias (IFIBYNE) CONICET, [2] Departamento de Fisiología, Biología Molecular y Celular "Dr. Héctor Maldonado", Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, (C1428EGA), Ciudad Autónoma de Buenos Aires, Argentina.

Abstract

Multisensory integration allows animals to combine cues from different sensory modalities to improve threat detection and drive evasive behaviors. While the zebrafish Tectum is classically viewed as a visual processing center, its role in integrating auditory and visual danger signals remains unclear. Using in vivo calcium imaging in larval zebrafish, we find that brief, high-amplitude auditory stimuli recruit a significant population of tectal neurons, predominantly in deep tectal layers. While visual looms produce strong responses in the zebrafish Tectum, combined auditory and visual stimuli robustly increases both tectal neural activity and the probability of executing rapid escape maneuvers, especially for low salient stimuli. Interestingly, simultaneous auditory and visual cues engage more tectal neurons than either modality alone. Single neuron analysis reveals a population of neurons that respond to both auditory and visual cues, and importantly, shows that a proportion of those neurons only responds to multisensory stimulation. These results establish the zebrafish Tectum as a multisensory processing area. Notably, the heightened tectal response during multisensory stimulation is strongly correlated with increased activity in hindbrain neurons implicated in escape-like motion events. These results link tectal multisensory integration to premotor circuits that drive behavioral output. Overall, our findings reveal the zebrafish Tectum as a hub for integration of audiovisual threat signals and provide a mechanistic basis for tectal multisensory processing in enhancing behavioral escape responses.

Introduction

Multisensory integration is a neural process by which animals combine information from different sensory modalities to build a coherent representation of their environment (Stein & Stanford, 2008; Van Atteveldt et al., 2014). This capability is essential for reducing ambiguity of sensory inputs, such as visual, auditory, and tactile cues, enhancing an organism's ability to detect and react to critical stimuli, such as predators, prey, or environmental hazards (Martorell & Medan, 2022; Lukas et al., 2021; Ohyama et al., 2015; McIntyre & Preuss, 2019).

This adaptive process has been widely documented across various animal species and contexts (Macaluso & Driver, 2005; Olcese et al., 2013; Whitchurch & Takahashi, 2006; Ghosh et al., 2017). In the mammalian superior colliculus, neurons integrate multisensory stimuli due to synaptic convergence of afferents conveying information from different sensory modalities (Meredith & Stein, 1986). Early studies in cats established that the neural response to multisensory stimulation can be supralinear with respect to the sum of the unisensory responses, and that integration effects are larger when weak or ambiguous stimuli are combined (Wallace et al., 1993; Stanford et al., 2005).

In other vertebrates, the Tectum, a structure homologous to the superior colliculus, is known to receive inputs from multiple sensory modalities (Isa et al., 2021; Knudsen, 1982; Hiramoto & Cline, 2009; Luksch, 2009; Hartline et al., 1978). This brain region is organized into distinct layers, primarily the neuropil -which receives afferents from retinal and other structures- and the periventricular layer (PVL) composed of tectal somata (Sajovic & Levinthal, 1982; Knudsen, 1982; Luksch, 2009). Retinotopic inputs of the visual field are received in the superficial layers of the neuropil, while deeper layers receive inputs from additional sensory modalities, like auditory, somatosensory and lateral-line inputs (Luksch, 2009). Neurons in the PVL are known to project to downstream targets in the hindbrain and spinal cord, where premotor neurons are involved in selecting and executing motor responses (Scott & Baier, 2009; Nevin et al., 2010). Despite receiving inputs from multiple sensory modalities and projecting to downstream motor centers, it is still unknown if the Tectum performs multisensory integration and if that integration influences motor output.

The Tectum of zebrafish (*Danio rerio*) has been traditionally associated with visual processing (Baier & Scott, 2024; Del Bene et al., 2010). While a few studies showed that tectal neurons can respond to non-visual stimuli such as auditory and water flow cues, responses were weak and sparse (Thompson et al., 2016; Vanwalleghem et al., 2017; Poulsen et al., 2021). One study reported multisensory suppression in some tectal neurons (Del Bene et al., 2010). However, no reports of robust multisensory enhancement in the zebrafish Tectum have been reported. Furthermore, it is not known whether tectal integration, if present, plays a functional role in influencing behavioral outputs.

Zebrafish are vision and hearing specialists with a well documented behavioral repertoire, making it a valuable model for studying audiovisual integration (Baier & Scott, 2024; Vanwalleghem et al., 2017; Fleisch & Neuhauss, 2006) with a well-documented behavioral repertoire (Kalueff et al., 2013; Marques et al., 2018). Specifically, danger stimuli evoke rapid escape responses known as C-starts, which allow fish to evade potential threats with remarkable speed and precision (Foreman & Eaton, 1993). Two types of fast escapes have received considerable attention due to their distinctive motor pattern and identifiable neural basis. The fastest escapes are short-latency C-starts (SLC), which are mediated by a pair of hindbrain premotor neurons, the Mauthner cells. Long-latency C-starts (LLC) are slower than SLC and mediated by Mauthner cell homologs and other hindbrain circuits (Burgess & Granato, 2007; Liu & Fetcho, 1999; Kohashi & Oda, 2008; Bhattacharyya et al., 2017; Marquart et al., 2019). These decision-making neurons, located in the hindbrain, receive sensory inputs from different structures, including the Tectum (Medan & Preuss, 2014; Zottoli et al., 1987; Preuss et al., 2006). Furthermore, SLC and LLC responses are known to be triggered by visual (e.g., looming objects) and auditory cues (e.g. abrupt sounds) (Temizer et al., 2015; Dunn et al., 2016; Kimmel et al., 1974).

In this study, we investigate whether audiovisual integration occurs within the Tectum and characterize its effects on behavioral decision-making in larval zebrafish. Through behavioral essays and calcium imaging at single-cell resolution we describe a strong multisensory enhancement of neural activation in the Tectum and of C-start probability during behavior. We observe robust auditory responses in the deep layers of the Tectum, extending its role in sensory processing beyond visual computation. Moreover, we find that integration in the tectum is particularly effective when visual stimuli are harder to detect and that a fraction of tectal neurons show supralinear responses to multisensory stimulation. Crucially, we find that tectal multisensory enhancement is strongly correlated with the activation of escape premotor neurons in the hindbrain, providing evidence for the zebrafish Tectum is a key site for multisensory integration and behavior selection in the context of threat detection.

Methods

Animal husbandry

We used larval zebrafish (*Danio rerio*) for this study. All animals were group-housed in 5 l aquaria of recirculating water, in a 14-10 h light–dark cycle (lights on at 8 am) in a temperature controlled room at 24-26 °C. All experiments were conducted between 9 am and 5 pm. Embryos were raised in E3 medium (5mM NaCl, 0.17 mM KCl, 0.33 mM CaCl2, 0.33 mM MgSO4) buffered with 10 mM HEPES (pH 7.3) at 28 °C, with medium changes every day. Larvae were fed AP100 (Zeigler Larval AP100 Diet) daily from 7 days post-fertilization (dpf) onward. We utilized 68 larvae aged 4 to 7 dpf of the line Nacre(elav3:GCaMP6f) for calcium imaging experiments. For behavioral experiments, we employed 80 wild type AB larvae aged 4 to 15 dpf. All experimental procedures were conducted in compliance with current regulations and were approved by the Institutional Animal Care and Use Committee of the Exact and Natural Sciences School, University of Buenos Aires (protocol 152).

Behavioral Experiments Setup

Behavioral assays were conducted in an arena consisting of a translucent plastic weighing boat of flat bottom, measuring 40 x 40 x 7 mm. The arena was filled with buffered E3 medium and was transilluminated using an infrared LED grid (55 x 55 mm, 36 LED units) and passed through opal diffusion glass to achieve more uniform irradiance. An overhead high-speed monochrome camera (720 x 540 px, 437 fps, Blackfly BFS-U3-04S2M-CS, FLIR) equipped with a 35 mm objective (59872, Edmund Optics) recorded larval behavior. The behavioral setup was mounted on an anti-vibration table and enclosed with opaque curtains to avoid unintended stimulation (Fig. 1A).

Calcium Imaging Experiments Setup

For calcium imaging experiments zebrafish larvae were immobilized by complete embedding in 4% ultrapure low-melting agarose (A9414-25G, Sigma) and submersed in E3 medium. Calcium imaging was performed using a Zeiss LSM-900 confocal microscope equipped with a 20× water-immersion objective lens (Zeiss W N-ACHROPLAN 420957) driven by a 488 nm laser. Emission was collected through a 509 nm

114 dichroic mirror and 490-620 nm emission filter using a GaASP-PMT detector. For each larva, we acquired
115 a z-stack encompassing the whole volume of both Tecta and adjacent structures (317 x 317 x 300 μm^3,
116 101 planes, interplane interval of 3 µm) to align brain images with the Z-Brain Atlas (Randlett et al., 2015).
117 In each larva, a coronal plane of the Tectum (317 x 317 µm) was recorded at 4.3 Hz to capture the neural
118 activity in response to a complete set (see below) of sensory stimulation. We also obtained a higher-
119 resolution reference image of each plane used for ROI segmentation. In some experiments, we performed
120 two sequential acquisitions of coronal planes of the Tectum at depths located between 30-60% of the
121 total dorsoventral extension of the Tectum. If an acquired plane suffered substantial drift or gross motion
122 artifacts it was excluded from the analysis.

123

124 **Stimulus Presentation**

125 ***Visual Stimuli***

126 Visual stimuli were delivered using a 1.5-inch OLED display module (Waveshare) connected to an Arduino
127 Due microcontroller. The display presented a looming stimulus (Adafruit_SSD1327 and custom C++ code)
128 consisting of a dark circle that expanded over a white background. The looming stimulus had an l/v ratio
129 of 1, where l is the radius of the virtual object and v is the approach speed (Gabbiani et al., 2002). The
130 expansion phase lasted 25 s, followed by a 10-s period where the circle remained at its maximum size
131 (Supp. Fig. 1A).

132 As multisensory enhancement depends on stimulus salience (Martorell & Medan, 2022), we designed
133 stimuli of varying intensity. To change the intensity of the visual stimulus, we adjusted the contrast of the
134 expanding circle against a constant white background. The contrast is expressed as:

135 $$\textit{Michelson Contrast (MC)} = (I_{background} - I_{disk}) / (I_{background} + I_{disk})$$

136 where $I_{background}$ is the irradiance of the background (33 mW/m²) and I_{disk} is the irradiance of the dark disk.
137 The irradiance was measured using a Tektronix J17 photometer equipped with a J1812 sensor. Based on
138 preliminary experiments we defined three levels of visual contrast (Fig. 1B): V_{min} (MC = 0.13), V_{med} (MC =
139 0.29), V_{max} (MC = 0.59) which showed progressive increase in the probability of eliciting an escape
140 response (Fig. 1D).

141 ***Auditory Stimuli***

142 Auditory stimuli were produced by a speaker (Thonet & Vander HK096-03451). The speaker was placed
143 on foam insulation padding and positioned 10 cm from the arena (Fig. 1 and Supp. Fig. 1). The intensity of
144 the auditory stimuli was measured on air using a sound level meter (Yu Fung Model YF-20) and reported
145 in decibels (dB re. 20 µPa).

146 The speaker emitted a single-cycle sine wave at 220 Hz generated and transmitted directly from the sound
147 card of the stimulation computer. To modulate the intensity of the auditory stimuli, we varied the
148 amplitude of the sine wave, defining three levels (Fig. 1B): A_{min} (60 dB), A_{med} (69 dB), A_{max} (75 dB) which
149 had respectively low, medium and high salience in terms of the probability of eliciting an escape response

(Fig. 1D). Visual and auditory stimuli were triggered and synchronized using a custom script written in Processing (Processing Foundation).

Multisensory Stimuli

We combined the 3 visual and 3 auditory unisensory stimuli to create 9 multisensory stimuli (Fig. 1B). In multisensory stimuli, the beginning of the single-cycle auditory stimulus was timed to coincide with the end of the visual looming expansion, simulating the simultaneous arrival of visual and auditory cues indicative of an approaching threat (Supp. Fig. 1A). Previous work from our lab showed that this stimulus configuration effectively evokes escape responses in fish (Martorell & Medan, 2022). Visual and auditory stimuli were presented from sources located at 90° (behavioral experiments, Fig. 1A) or 135° (calcium imaging experiments, Supp. Fig. 1B) from each other. Preliminary experiments showed that the auditory stimulus does not have a strong directional component in our conditions. For imaging experiments, we used only tested responses to V_{min}, V_{max} and A_{max} and their combinations.

Experimental Procedures

Prior to the start of experiments, larvae were given at least 6 minutes to acclimate to the experimental setup. In behavioral experiments individual larvae were exposed to each of the uni- or multisensory stimuli once, and an additional control condition with no stimulation in random order totaling 16 trials with an intertrial interval (ITI) of 40 s (Fig. 1B, Supp. Fig. 1A). In calcium imaging experiments, larvae were exposed to 5 repetitions of a subset of these conditions: V_{max}, A_{max}, their multisensory combination and the control condition (random order, 20 trials in total). Some calcium imaging experiments (Fig. 4) also included the V_{min} in combination with A_{max} (30 trials in total, ITI = 40 s).

Behavioral Data Analysis

Pose Estimation Using DeepLabCut and Behavioral Event Classification

Larval motor behavior was tracked using DeepLabCut (Mathis et al., 2018). We trained a ResNet-50 convolutional neural network where the output layer was replaced with a pose estimation layer designed to predict the location of 12 anatomical landmarks as scoremap grids: one point for each eye, two points for the swim bladder, and eight points along the tail. Training was performed on a single NVIDIA GeForce RTX 3060 GPU with 12 GB of VRAM.

An initial training dataset was created by annotating 152 frames from a library of 1,280 behavioral recordings. Frames were chosen to maximize variability in body poses, positions within the arena, fish sizes, and lighting conditions. For a second training iteration, we annotated an additional 125 frames where the model's initial predictions had low likelihood scores (high uncertainty in model predictions).

The annotated dataset was split into training and validation sets, with 95% of the images used for training and the remaining 5% reserved for model evaluation. Using pose estimation data we calculated 36 kinematic parameters, such as linear head velocity or total body curvature (Supp. Table 1). Discrete activity events, or bouts, were identified by detecting peaks in the linear head velocity time series. Events exceeding a threshold of 6 cm/s were classified as fast responses, as visual inspection of video recordings showed that all events exceeding this threshold were C-starts. (Fig. 1C).

187

188 ***Quantification of Multisensory Integration***

189 I0n both behavioral and imaging experiments, multisensory integration was quantified by comparing the
190 activity elicited during multisensory stimulation with the activity observed during the most effective
191 unisensory stimulus (Stevenson et al., 2014). For behavioral experiments, activity was measured as the
192 probability of eliciting a fast response, while in imaging experiments, it was measured as the proportion
193 of active ROIs.

194 The multisensory integration coefficient (MI) was calculated using the formula:

195 $$MI = (P_{multi} - P_{max\ uni}) / (P_{multi} + P_{max\ uni})$$

196 where P_{multi} represents the activity during the multisensory stimulus, and $P_{max\ uni}$ represents the activity
197 during the unisensory stimulus that elicited the largest response. An MI value of zero indicates no effect
198 of multisensory integration beyond the maximum unisensory response. A negative MI value indicates
199 multisensory suppression, whereas a positive MI value indicates multisensory enhancement.

200

201 **Calcium Imaging Data Processing and Analysis**

202 ***Image Registration, Motion Correction and Cell Segmentation***

203 Calcium imaging recordings were processed to correct for motion artifacts and registered to the higher-
204 quality template using the CaImAn library (Giovannucci et al., 2019). We next used custom Python code
205 for semi-automated Region Of Interest (ROI) detection based on the anatomical features of the images.
206 Note that the calcium sensor expression is cytosolic and therefore cell nuclei are not fluorescent. The
207 Tectum is organized into the PVL, a deep layer containing most neuronal somas, and the neuropil, a
208 superficial layer where dendrites and afferent axons converge. We detected ROIs in three specific brain
209 regions: the tectal neuropil, the periventricular layer of the Tectum (PVL), and the hindbrain (Supp. Fig.
210 1C). To estimate the proportion of the visible neuropil area exhibiting significant activity we analyzed the
211 neuropil both as an entire undivided region and by subdividing it into a grid 5 µm x 5 µm to capture
212 regional changes in neurite activity.

213 ***Fluorescence Normalization and Time Series Analysis***

214 For each ROI we calculated fluorescence time series by averaging the pixel intensities within the ROI for
215 each recorded frame. These time series were then normalized to ΔF/F, defined as:

216 $$\Delta F/F = (F - F_0) / F_0$$

217 where F is the fluorescence intensity at a given time point, and F_0 is the baseline fluorescence (example
218 shown in Supp. Fig. 1Di) calculated as the 8th percentile of a 60-frame (approx. 14 s) window preceding
219 the current frame. We next binarized the ΔF/F time series for each ROI and set a threshold of 1.2 SD above
220 the mean ΔF/F value. Any ΔF/F value exceeding this threshold was considered an event of activation at
221 that frame (example shown in Supp. Fig. 1Dii). ROIs inactive during the entire recording session were

excluded from subsequent analysis. All experiments had ROIs responding to visual stimuli, however we only included experiments where at least one ROI responded in at least 3 out of 5 auditory trials (33 planes from 24 animals, out of 113 planes from 68 animals). Responses were quantified in a 2.8 s (12 frames) window after stimulus presentation (Supp. Fig. 1D, magenta shaded area). ROIs were classified as responsive if they exhibited activity greater than 1.2 SD above the mean ΔF/F value in at least one of the five trials of that stimulus during the stimulation window. The consistency of a ROI's response was quantified by counting the number of trials in which the ROI was active during the stimulation window (i.e. the consistency values ranged 0 to 5).

To examine population-level neural representations, we performed temporal principal component analysis (t-PCA) on ΔF/F time series from PVL neurons (Supp. Fig. 2). The analysis was restricted to a window following stimulus onset, which encompassed the peak of population responses. This dimensionality reduction allowed us to project N-dimensional neural activity patterns (where N is the number of PVL neurons recorded in each experiment) onto a two-dimensional space, facilitating the visualization and comparison of population trajectories across sensory conditions.

Identification of Motion-Correlated Neurons

Although embedded in agarose, larvae showed clear motor activity as displacements of the image. The activity often was time-locked to the frames immediately following sensory stimuli. The frame-by-frame displacement data obtained from motion correction analysis were used to construct a continuous estimate of the animal's movement over time (Supp. Fig. 1E). To detect motor events, we calculated the mean displacement for each experiment, averaged over all recorded frames. Any frame where the displacement exceeded 4 SD above the experiment's mean displacement was considered a motion event. This threshold captured large, abrupt displacements indicative of motor-related activity while minimizing the influence of minor shifts that could result from slight image drift.

Statistical Analysis

All statistical analyses were conducted using Python (v3.9) with the SciPy and Statsmodels libraries. Statistical significance was assessed at a threshold of $\alpha = 0.05$. Binomial Generalized Linear Models (GLMs) were used to evaluate the effects of visual and auditory stimuli on escape response probabilities. Mixed effect models were used in some cases to factor in the variance accounted for by individual differences. Model significance was tested using likelihood ratio tests, and parameter estimates are reported with their associated p-values. The significance of MI values across conditions was assessed using one-sample t-tests. Wilcoxon Signed-Rank Tests were used to compare paired conditions in imaging experiments (e.g., auditory vs. control, multisensory vs. visual), while non-paired tests were performed using Mann-Whitney U-tests. Z-Tests for Regression Slopes were used to compare slopes between different conditions. All analyses were corrected for multiple comparisons where applicable using the Holm-Bonferroni method. Error bars represent standard error unless otherwise specified. Graphs and visualizations were generated using Matplotlib and Seaborn libraries. The code is open-source and available in GitHub.

Results

Multisensory Stimulation Enhances Escape Probability in Freely-Swimming Fish

To determine if multisensory information influences escape behaviors of larval zebrafish, we conducted behavioral experiments in 4-7 dpf freely swimming zebrafish (n = 40, Fig. 1A). Visual and auditory stimuli of different intensities were presented individually and in combination, including three levels of visual contrast, three levels of auditory amplitude, and their nine combinations, along with a no-stimulus control (Fig. 1B).

The probability of eliciting a fast escape response (events where head velocity was higher than 6 cm/s, Fig. 1C) increased with the intensity of both visual and auditory stimuli (Fig. 1D; Mixed Effects Binomial GLM, fish identity as a random effect, auditory effect: 0.91, $p<0.0001$; visual effect: 0.46, $p<0.0001$). To quantify the effect of combining sensory modalities, we calculated a multisensory integration coefficient (MI) for each of the nine audiovisual conditions. The MI compares the escape probability during multisensory stimulation to the most effective unisensory stimulus (see Methods). This analysis revealed that combined audiovisual stimuli significantly increased escape probabilities beyond the maximum unisensory response (Fig. 1E; the mean of the MI values for all multisensory conditions is significantly greater than 0, Wilcoxon Signed-Rank Test, $p = 0.037$). Interestingly, the magnitude of this enhancement was inversely related to the intensity of the auditory stimulus (Gaussian GLM, auditory effect: -0.115, $p = 0.009$) while the intensity of the visual stimulus had no significant impact on the MI (visual effect: -0.006, $p = 0.90$). This was further supported by the negative correlation between the MI of each audiovisual condition and its maximum unisensory response probability (Fig. 1F, Linear Regression, slope = -0.50, $p = 0.013$, $R^2=0.61$).

These findings agree with the principle of inverse effectiveness, where weaker unisensory stimuli produce proportionally greater multisensory enhancements (Stein et al., 2009; Rowland et al., 2007). For instance, the weakest auditory stimulus (60 dB), elicited responses in 15% of trials while the lowest contrast visual stimulus (Michelson Contrast 0.13) elicited responses in 10% of the trials. When these two weak stimuli were combined (stim. 7 in Fig. 1B), the response probability rose to 33% (Fig. 1D), yielding a mean multisensory integration coefficient of 0.37 (Fig. 1E-F). In contrast, responses evoked by the loudest auditory stimulus (75 dB) did not show multisensory enhancement, showing almost no change compared to its unisensory response probability (Stims. 3, 6 and 9 in Fig. 1F).

Overall, these results confirm that larval zebrafish can combine audiovisual information to decide whether to execute a fast escape response (Martorell & Medan, 2022), and that the strongest multisensory enhancement is produced by exposure to the weakest stimuli. However, the mechanistic and neuronal bases of this multisensory integration are unknown. To this aim, we next decided to study the neuronal correlates in the Tectum.

Auditory Stimulation Elicits Neural Responses in the Tectum

296 Larval zebrafish are capable of integrating multisensory information during fast escapes (Fig. 1). While
297 visual processing in the fish Tectum has been described, accounts of auditory coding are scarce and
298 focused on the representation of pure tones (Thompson et al., 2016; Vanwalleghem, Heap, & Scott, 2017;
299 Privat et al., 2019; Poulsen et al., 2021). We thus investigated audiovisual representation within the
300 Tectum using a brief and abrupt sound which can evoke fear responses (Fig. 1) to ask whether this
301 classically visual brain region also responds to auditory danger stimuli.

302 To answer this question, we performed *in vivo* confocal calcium imaging in agarose embedded larval
303 zebrafish at 4-7 dpf expressing the cytoplasmic calcium indicator GCaMP6f (Nacre(elav3:GCaMP6f))
304 (Supp. Fig. 1B). We imaged the PVL and the tectal neuropil during the presentation of auditory and visual
305 stimuli (Supp. Fig. 1C) and quantified the tectal response during a trial as the proportion of active ROIs
306 ($\Delta F/F >$ 1.2 SD above the mean, see Methods) during a 2.8 s window after sound presentation (Fig. 2).

307 We found that the presentation of an auditory stimulus produced a robust response in the PVL (Fig. 2A)
308 whereas the number of spontaneously active neurons in this region was relatively low (control trials, Fig.
309 2B). Across 33 imaged planes (13851 ROIs, 24 larvae), the mean proportion of active ROIs activated by
310 auditory stimulation showed a sharp increase after onset of stimulation (Fig. 2A) while control trials
311 showed a steady and low proportion of spontaneously active ROIs. Figure 2C shows the time-average of
312 the 33 planes shown in Figs. 2A and B. Across experiments, auditory stimulation produced a mean
313 simultaneous activation of around 8% of PVL neurons (Fig. 2C), a proportion notably higher than the basal
314 3% observed in control trials. The mean proportion of active neurons for auditory trials was also
315 significantly higher than for control trials (Fig. 2D, median: Ctrl = 2.7%, Aud = 8.9%; Wilcoxon signed-rank
316 test, $p < 0.001$).

317 The auditory-evoked neural representation in the PVL displays a consistent pattern across trials of the
318 same animal. Figure 2E and 2F show representative examples of responsive neurons in the PVL during
319 auditory or control trials respectively. We hypothesize that auditory responsive neurons could be
320 randomly scattered in the tectum or show some kind of topographical assembly. To test this, we measured
321 the compactness of the auditory response as the mean nearest-neighbor distance among auditory-
322 responsive ROIs. Taking the ratio of compactness of auditory neurons and a random subset of ROIs reveals
323 that auditory-evoked representations are more compact than expected by chance (median: Aud: 0.12,
324 Ctrl: 0.09; compactness ratio is significantly greater than 1; Wilcoxon signed-rank test, $p < 0.001$). This
325 indicates a non-random clustering of responsive neurons which was, however, different from animal to
326 animal.

327 The tectal neuropil also showed robust auditory responses (Fig. 2G). To quantify those responses, we
328 subdivided the neuropil into a grid of neuropil ROIs given the absence of clearly identifiable somas.
329 Neuropil ROIs had approximately the same area as ROIs defined in the PVL, however, contrary to PVL ROIs,
330 neuropil ROIs do not correspond to somata but quantify the fluorescence emitted by the neurites located
331 within.

332 Responses in the tectal neuropil during auditory trials (Fig. 2G, K) were stronger and more synchronous,
333 i.e. reached a higher proportion of simultaneously active ROIs, than during control trials (Fig. 2H, L).
334 Auditory stimulation produced a bilateral and synchronized activity of neuropil ROIs which clustered

335 towards the medial edge of the tectal neuropil (Fig. 2K) while in control (no sound) trials fewer, scattered
336 grid ROIs were active (Fig. 2L). The neuropile activity showed dynamics that on average reached 13% of
337 its area for auditory trails, which was significantly higher than the 3% active for control conditions (Fig.
338 2I). Analysis of individual planes showed the same behavior, with a consistently larger area activated for
339 auditory than for control trials (Fig. 2J; median: Ctrl = 2.9%, Aud = 7.3%; Wilcoxon signed-rank test, $p <$
340 0.001).

341 The neuropil's response to auditory stimulation seemed to be concentrated in deeper (closer to the
342 medial edge) layers (Fig. 2K) while activity in the superficial layers was rarely observed. To characterize
343 the spatial distribution of the neuropil activity, we computed the mean distance of the activated neuropil
344 ROIs to the medial edge of the Tectum (green line in Fig. 2K) and found that this distance was significantly
345 smaller than that of a null model created by randomly selecting the same number of ROIs (mean +/- sd:
346 26.78 ± 10.18 µm vs. 30.56 ± 7.19 µm; Wilcoxon signed-rank test, $p = 0.002$), indicating a non-random
347 clustering of responsive neurons. Beyond clearly showing that there is auditory coding in the Tectum, our
348 results suggest auditory input reaches the Tectum at the deeper layers of the neuropile and contacts the
349 dendrites of neurons whose somata are in the PVL.

350 In addition, we consider the lack of regionalization of auditory responses along the anteroposterior axis
351 both in the neuropile nor in the PVL suggestive of a distributed auditory input to the Tectum (as opposed
352 to a regionalized area that computes auditory input). In turn this implies that visual input coming for any
353 part of the visual field could be integrated with auditory input.

354

355 **Auditory Inputs Enhance Visual Coding in the Tectum**

356 The Tectum participates in the processing of visual looming stimuli (Barker & Baier, 2015; Marquez-
357 Legorreta, Piber, & Scott, 2020). Here we investigated whether auditory cues presented concurrently with
358 a visual looming stimulus (Supp. Fig. 1A-B) could affect tectal activity (Fig. 3).

359 In the PVL, multisensory trials activated a larger number of neurons across experiments compared to
360 visual stimuli alone. Visual trials elicited clear responses that recruited a significant proportion of PVL
361 neurons (Fig. 3A), but multisensory trials induced larger and synchronic activation immediately following
362 the auditory stimulus onset (Fig. 3B-C). Across experiments, the number of active ROIs reached a higher
363 peak in multisensory trials, the mean proportion of active ROIs in the PVL (shaded area in Fig. 3C) was
364 significantly larger than in visual activation alone trials (Fig. 3D, median: Vis = 13.0%, Multi = 20.7%;
365 Wilcoxon signed-rank test, $p < 0.001$). These results indicate that combining an auditory and a visual cue
366 can significantly increase neuronal recruitment in the PVL. Note that as the visual stimulus was delivered
367 from the left side of the animal (Supp. Fig. 1B) it evoked a visual response principally in the contralateral
368 (right) hemitectum (Fig. 3E). Intriguingly, the same visual loom paired with an auditory stimulus recruited
369 a comparable proportion of ROIs in both hemispheres (Fig. 3F).

370 We next asked if multisensory enhancement could be detected in the neuropil as well, and if the increased
371 activity observed in the PVL during multisensory trials was matched with an increase in the proportion of
372 active neuropil ROIs. Also, we asked how the anatomical representation of the multisensory responses

would relate to the visual representation at input (neuropil) level. Since we observed that adding auditory input to the visual loom produced a more distributed activation across the contralateral PVL (Fig. 2K) and notably of the ipsilateral PVL, we wondered if multisensory responses in the neuropile would be also more distributed, suggesting that integration of visual and auditory afferences enhance Tectum activation at the expense of losing spatial acuity.

We observed a robust increase in the proportion of active neuropil area during multisensory trials compared to visual trials (Fig. 3G vs. 3H). The peak proportion of active grid ROIs during multisensory stimulation was 10% higher than in visual only trials (Fig. 3I) while the mean proportion of active neuropil area was 5% higher in multisensory trials compared to visual trials (Fig. 3J, Vis = 19.3%, Multi = 25.4%; Wilcoxon signed-rank test, $p < 0.001$). Notably, the addition of the auditory stimulus caused the activation of deeper layers of the visible part of the ipsilateral neuropil (ROIs closer to the medial edge, green curves, Fig. 3L), just like we observed with the unisensory auditory stimulus (Fig. 2K). Indeed, the mean distance of the activated neuropil regions to the medial edge was lower in multisensory trials compared to visual trials (Fig. 3K-L, visual, mean +/- sd: 29 +/- 7 um vs. multisensory: 27 +/- 8 um; Wilcoxon signed-rank test, $p = 0.0014$). These results imply that the neuropil's response to audiovisual stimulation is a combination of visually-evoked activity in superficial layers and auditory-evoked activity in deeper layers, serving as the multisensory inputs subsequently processed by the PVL. It also confirms that the combination of auditory and visual input not only enhances activity in the hemitectum contralateral to the visual input but produces a significant activation of the ipsilateral neuropil.

Multisensory Enhancement of Tectal Response Shows Inverse Effectiveness with Stimulus Salience

Numerous studies have reported that the multisensory enhancement of neural responses tends to be inversely correlated to the effectiveness of the inputs when received separately (Meredith & Stein, 1986; Stein et al., 2008; Gil-Guevara, Bernal, & Riveros, 2022; Sumby & Pollack, 1954, Martorell & Medan, 2022) and thus, particularly effective when individual signals are weak or ambiguous. To ask if multisensory integration in the Tectum also shows inverse effectiveness, we presented the auditory cue (A_{max}, Fig. 1B) in combination with a low-contrast visual loom (V_{min}. Fig. 4A) or a high contrast loom (V_{max}, Fig. 4C) and tested if decreasing the intensity of the visual stimulus increased multisensory enhancement in the zebrafish Tectum.

While low-contrast visual stimulus (V_{min}) evoked a tectal response in 1.5% of PVL neurons (Fig. 4B), V_{max} recruited 10.2 of PVL neurons (Fig. 4D), a significantly higher proportion (V_{min} vs. V_{max}, Wilcoxon signed-rank test, $p < 0.001$). However, we observed an increase in the mean percentage of active ROIs both with V_{min} combined with an auditory stimulus (Fig. 4B, V_{min}= 1.5% vs. AV_{min}= 4.7%; Wilcoxon signed-rank test, $p = 0.004$) and with V_{max} combined with an auditory stimulus (Fig. 4D, V_{max}= 10.2% vs. AV_{max}= 15.5%; Wilcoxon signed-rank test, V_{max} vs. AV_{max}, $p < 0.001$). The MI (Multisensory Index, see Methods) was MI_{min}= 0.25 for the low-contrast multisensory stimulus and MI_{max}= 0.14 for the high-contrast multisensory stimulus. This larger multisensory enhancement for the low-contrast condition than for the high-contrast condition demonstrates that in the tectal PVL the multisensory integration shows inverse effectiveness.

411 In the neuropil we observed a similar relationship between visual contrast and the proportion of active
412 neuropile area (Fig. 4E and G) or the mean percentage of active neuropile area (Figs. F and H, V_{min}= 6.5%
413 vs. V_{max}= 26.3%). The effect of combining a visual and an auditory stimulus was stronger for V_{min} (Fig. 4F;
414 V_{min}= 6.5% vs. AV_{min}= 13.1%; Wilcoxon signed-rank test, p = 0.001, MI_{min} = 0.22) but there was also a
415 significant increase for the high-contrast condition (Fig. 4H, V_{max}= 26.2% vs. AV_{max}= 32.1%; Wilcoxon
416 signed-rank test, $p < 0.001$, MI_{max} = 0.08).

417 The effective salience of a visual stimulus is not only a function of its contrast, but also depends on gaze
418 direction, attentional state and other factors that can vary between animals and trials. We thus analyzed,
419 for each individual experiment (imaged tectal plane) the relationship between the MI and activity during
420 the visual unisensory condition in the PVL (Fig. 4I) and in the neuropil (Fig. 4J). The inverse effectiveness
421 hypothesis predicts that in experiments where the visual stimulus evokes a weaker response there will be
422 a stronger multisensory enhancement. Indeed, MI and the maximal unisensory response are inversely
423 related for both conditions using a single exponential-decay model, both in the PVL (Fig. 4I; exponential
424 fit, b>0, two-sided p = 0.0013, R^2 = 0.571, n = 58) and in the neuropil (Fig. 4J; exponential fit, b>0, two-
425 sided p = 0.00070, R^2 = 0.589, n = 54). Low-contrast stimuli are usually perceived as less salient and
426 therefore tend to produce higher MI values compared to the high-contrast condition (purple vs. magenta
427 dots in Figs. 4I and J), although this difference is not significant in the PVL (Wilcoxon signed-rank test,
428 paired by experiment, p = 0.20), possibly due to variability across animals. In the neuropil, however, MI
429 values for the low-contrast condition are significantly higher than for the high-contrast condition
430 (Wilcoxon signed-rank test, paired by experiment, p = 0.019). Overall, this shows that the Tectum exhibits
431 stronger multisensory enhancement when the unisensory components of the multimodal stimulus have
432 lower salience, a hallmark of multisensory integration.

433

434 **Multisensory Stimulation Recruits Unresponsive PVL Neurons and Boosts Response Probability of**
435 **Audiovisual Neurons**

436 While the Tectum shows multisensory enhancement at the population level (Fig. 3) and this effect alone
437 could have important consequences on downstream circuits, actual multisensory integration in tectal
438 neurons requires that individual neurons change their response when auditory and visual cues are
439 combined. This is because an increase in the proportion of active neurons could be due to a sum of
440 independent unisensory responses occurring in distinct cells, and not to true multisensory integration at
441 neuronal level. We thus asked if this multisensory enhancement would also be observable at the individual
442 neuron level, which in our experiments, should be reflected in cells that are only active when auditory and
443 visual inputs are simultaneously present.

444 We analyzed the response profiles to unisensory and multisensory stimuli of individual PVL neurons.
445 Based only on their response to unisensory stimuli (i.e. excluding multisensory trials) we categorized PVL
446 neurons in four distinct classes: neurons that responded only to auditory stimulation (EA neurons),
447 neurons that responded only to visual stimulation (EV neurons), neurons that responded to both
448 modalities (i.e. to unisensory auditory or unisensory visual stimuli, AV neurons), and neurons that did not
449 respond to either unisensory auditory or unisensory visual stimuli. This last group was further divided into

two classes depending on whether neurons were (only) responsive to multisensory stimulation (EM neurons) or unresponsive to all stimulation (U neurons). It's important to note that, although auditory stimulation produced bilateral activation of the Tectum (see examples on Figs. 2E, K), visual stimulation produced preferential activation of the hemitectum contralateral to the screen (Fig. 3E, K). We thus restricted the analysis to the contralateral Tectum (left Tectum, Supp. Fig. 1B), where activation of both visual and auditory inputs was robust.

A representative example of this distribution is illustrated in Fig. 5A, where ROIs are color-coded based on the neuron response profile. The spatial distribution and proportions of these neuron types varied across experiments, and we did not find any indication of spatial clustering of the responses (other than retinotopy for stimuli with a visual component). On average, 21% of PVL neurons (24 larva, 13851 neurons) were unresponsive to all stimuli (Fig. 5B, U). On the other hand, 29% of PVL neurons were responsive to visual looms but never to auditory stimuli (Fig. 5B, EV) while 19% showed the opposite, being only responsive to auditory stimuli (Fig. 5B, EA). To confirm that these neurons were truly insensitive to their non-preferred modality in our stimulation conditions, we compared their activity during its preferred-modality stimulation with their activity during multisensory trials. Neither EA nor EV neurons showed an increased response to multisensory stimulation (EA, Wilcoxon signed-rank test, $p = 0.12$; EV, Wilcoxon signed-rank test, $p = 0.89$, not shown). This could be explained by the nature of our stimuli and/or the spatial configuration of our setup not being effective (e.g. retinotopic effects for the visual stimulus and possible dependence on sound features for the auditory stimulus), but it could also point to some PVL neurons truly receiving only one kind of sensory input.

Since on average 19% of PVL neurons responded to the visual loom but also to unisensory auditory trials (Fig. 5B, AV), a proportion much larger than the <1% previously reported for non-danger audiovisual cues (Thompson et al., 2016), we asked if these AV neurons that are activated by both unisensory auditory and unisensory visual stimuli would increase their responsiveness when presented with those stimuli simultaneously, or if they would show audiovisual suppression as previously shown for non-danger cues (Thompson et al., 2016). To answer this, we quantified the consistency of each AV neuron as the number of trials in which it was active for a particular stimulus. This value represented the number of trials of a given stimulus condition in which a neuron was active, ranging from 0 (for a consistently inactive ROI) to 5 (for a ROI that was activated on each of the 5 trials for a particular condition). We calculated the consistency of each AV neuron individually for each stimulation condition: control (no stimulation), auditory, visual and multisensory (Fig. 5C). Given that visual responses were stronger and more consistent than auditory responses across our experiments, we compared the multisensory consistency against the visual consistency and found that AV neurons increased their consistency (i.e. their individual probability of response) for multisensory vs. visual trials (Mixed Effects Linear Regression, fish identity as a random effect, effect size = 0.15, p = 0.04). This can also be seen when calculating the proportion of active audiovisual ROIs over time (Fig. 5Di), where the peak multisensory activity of these neurons is significantly higher than their peak activity for either of the two unisensory stimuli (Fig. 5Dii, Wilcoxon signed-rank test, $p = 0.012$). These results parallel those observed for the whole population of PVL neurons (Figs. 3D, 4B-F) and critically show that simultaneous auditory and visual input can enhance response probability at the individual neuron level.

490 Strikingly, we found that about 13% of PVL neurons were silent during all ten unisensory trials but
491 responded to (at least once) exclusively multisensory stimulation (Fig. 5B, EM). To ensure that these
492 responses were not due to spontaneous activity, we compared the activity of these neurons during
493 multisensory trials against their activity during control trials. EM neurons were more consistently active
494 during multisensory stimulation than in unisensory or control trials (Fig. 5E) and their peak response
495 probability was as well significantly higher when compared to control trials (Fig. 5F, Wilcoxon signed-rank
496 test, $p < 0.001$). This shows that multisensory stimulation can recruit cells which, although receive inputs
497 from both modalities, are silent unless those inputs are activated simultaneously.

498 We next asked if the activity of AV neurons would form distinct representations of visual and auditory
499 stimuli (i.e. different responses to auditory and visual unisensory stimuli), and, if so, how multisensory
500 representations would relate to unisensory representations (Supp. Fig. 2). To answer this question, we
501 performed temporal principal component analysis (t-PCA) on the $\Delta F/F$ data from individual AV PVL
502 neurons during visual, auditory, and multisensory trials within a 1.8 second window after stimulation time
503 (which encompasses the period where populational responses reaches their peak, Fig. 5Di). This allowed
504 us to transform N-dimensional activation patterns (where N is the number of PVL neurons in a particular
505 experiment) into trajectories in a low-dimensional space which encode information about the temporal
506 dynamic of the tectal representation for each trial.

507 Two representative examples from two different animals are shown in Supp. Fig. 2A and B. Plots of the
508 first two components of the PCA show trajectories for trials that showed a significant response, i.e. there
509 can be up to 5 trajectories for each type of stimulus (lines correspond to PVL trajectories in response to
510 visual, auditory and multisensory stimuli). All trajectories were normalized to start at the origin, which
511 represents the Tectum state before stimulation. Trajectories corresponding to visual and auditory trials
512 diverged in PCA space (median angle between visual and auditory trajectories was 60°), suggesting that
513 the tectal representation to visual and auditory stimuli are different.

514 To describe the variability of the multisensory patterns of activity and compare it to their unisensory
515 counterparts in all experiments, we calculated the median unisensory direction for each unisensory
516 stimulus Supp. Fig. 2A and 2B) and evaluated the angle of each multisensory trajectory with both
517 unisensory vectors (Supp. Fig. 2C). Most multisensory patterns of activity align with either the visual or
518 the auditory representation, i.e. each pattern of activity has a small angle with the visual vector and a
519 larger one with the auditory vector or vice versa, with the visual mode being more common (Supp. Fig.
520 2C, 73% visual, 27% auditory). To quantify the separation between unisensory representations, we treated
521 each complete trajectory (spanning 8 post-stimulus frames) as a single vector in a 16-dimensional space
522 (2 coordinates x 8 time points). We then computed the average Euclidean distance between all pairs of
523 visual-auditory trajectories (inter-modality distance) and compared it to the average distances within the
524 same modality (intra-modality distances, i.e., visual-visual and auditory-auditory). The inter-modality
525 distance was significantly greater than intra-modality distances (Supp. Fig. 2D, Mann-Whitney U test, $p <$
526 0.001) confirming that visual and auditory stimuli evoke distinct and separable population activity
527 patterns in the PVL.

528 To determine whether the multisensory response aligned with one of the unisensory inputs or had a
529 distinct activity pattern, we compared the multisensory trajectories against their unisensory counterparts.

530 First, we computed the average Euclidean distance between multisensory and auditory trajectories (inter-
531 modality) and compared it to their respective intra-modality distances. The multisensory-auditory
532 distance was significantly greater than both the intra-multisensory and intra-auditory distances (Supp. Fig.
533 2E, Mann-Whitney U test, $p < 0.001$). Conversely, when we performed the same comparison between
534 multisensory and visual trajectories, the inter-modality distance was not significantly different from the
535 intra-modality distances (Supp. Fig. 2F, Mann-Whitney U test, $p = 0.12$). This analysis confirms that the
536 population response to the multisensory stimulus is closer to the visual-only response, as opposed to the
537 auditory-only response.

538

539 **Multisensory Integration in the Tectum Is Associated with Enhanced Motor Responses in the Hindbrain**

540 We have shown that the zebrafish Tectum integrates visual and auditory information, enhancing neural
541 responses. Does this multisensory enhancement have a functional behavioral value? The Tectum is known
542 to project to premotor regions in the hindbrain, both in zebrafish and in other vertebrate species (Del
543 Bene et al., 2010; Baier & Scott, 2024). Some hindbrain neurons receive input from the Tectum to evoke
544 fear-related responses, such as the Mauthner cell-initiated C-start responses (Zottoli et al., 1987; Zottoli,
545 1977; Zottoli et al., 1999). It is therefore possible that audiovisual integration in the Tectum could affect
546 premotor neurons and, consequently, influence the execution of behavioral responses. For this reason,
547 we next explored the relationship between the enhancement of tectal activity during multisensory trials
548 and behavioral output.

549 Although larvae were agarose-embedded to restrict brain movement during image acquisition they still
550 exhibited small, sudden movements that were registered during the recording. These sudden movements
551 in the x-y plane are routinely corrected using a motion correction algorithm before ROI extraction
552 (Giovannucci et al., 2019). Interestingly, we noticed that these motion events were more frequent during
553 stimulus presentation. Frame-by-frame displacements of confocal images larger than 4 SD above the
554 mean displacement of each experiment (see Methods) were considered to be motion events. Across all
555 imaged planes (n=33) we found that these motion events were more common within the period following
556 visual (4%) or auditory (6%) stimulation than during spontaneous activity (1%). Crucially, the probability
557 of eliciting motion events during multisensory trials (12%) was significantly higher than the maximum
558 unisensory (auditory) condition (Fig. 6A, Paired t-test, Aud vs. Multi, $t= 2.26$, $p = 0.03$). These results
559 parallel the multisensory enhancement we observed for escape responses in behavioral experiments (Fig.
560 1D-E), although response probabilities here are much lower, likely because of the differences in the
561 conditions of stimulus administration.

562 If these events represent actual motor responses this should be reflected in activity of hindbrain neurons.
563 We focused on neurons primarily located in rhombomeres 1 and 2, which were imaged simultaneously
564 with tectal activity. To identify candidate premotor neurons, we first calculated the activation probability
565 of each hindbrain neuron considering only spontaneous motion events (i.e. events occurred in the
566 absence of sensory stimulation). We next produced a shuffled distribution of activations for each trial and
567 each neuron, where motion event times were randomized to eliminate real correlations and compared

568 the real and shuffled distributions. We defined motion-correlated neurons (MCN) as those neurons with
569 an activation probability higher than 99th percentile of the shuffled distribution (Fig. 6B).

570 We confirmed that this population of hindbrain neurons was active during both spontaneous and
571 stimulus-evoked motion events (Fig. 6C). Notably, the proportion of MCNs that became active in response
572 to sensory stimulation was higher for auditory than for visual trials and showed multisensory
573 enhancement over the auditory condition (Fig. 6D, Mann-Whitney U test, p = 0.037), reproducing the
574 pattern observed in the motion events (Fig. 6A).

575 The enhancement in the proportion of active MCN of the hindbrain after multisensory stimulation (Fig.
576 6D) was reminiscent of the multisensory enhancement we observed in the Tectum (Figs. 3B, 4B). We thus
577 asked if the level of activity of the MCNs correlated with MI observed in the Tectum. For each multisensory
578 trial, we calculated tectal MI, similarly to what we did in Fig. 4A-B to quantify multisensory enhancement
579 (see Methods). We observed that when tectal MI was zero or negative—indicating no enhancement or
580 suppression—the activation of MCN in the hindbrain was weak and not significantly associated with MI
581 (Fig. 6E, linear regression, $p = 0.716$). However, when tectal MI was positive, there was a strong positive
582 association between MI and MCN activity (Fig. 6E, linear regression, $p < 0.001$, coefficient = 0.65). This
583 shows that there is a positive relationship between multisensory enhancement in the Tectum and
584 hindbrain activity linked to behavioral output. Crucially, this relationship is rectified as a function of
585 integration valence: only positive multisensory effects in the Tectum are associated with an increased
586 probability of MCN activation in the hindbrain.

587 To determine if this association between tectal MI and hindbrain activity was driven by all or by a subset
588 of MCNs, we calculated the mean ΔF/F of the five multisensory trials for each MCN and computed the
589 correlation coefficient between its mean ΔF/F and the tectal MI value for the corresponding trials (Tectal
590 M - corr. coef. MCN). This metric indicates, for each experiment, how strongly each MCN was correlated
591 with tectal MI during multisensory stimulation. In Fig. 6F black bars represent the real distribution of
592 Tectal M - corr. coef. MCN for an example experiment while gray bars show the results when shuffling
593 trial order to remove all significant associations. All neurons with coefficients larger than the 95th
594 percentile of the shuffled distribution (green shaded area in Fig. 6F) were considered to be correlated with
595 tectal activity. Across experiments, the proportion of MCNs exceeding this threshold varied (Fig. 6G) but
596 on average 30% (+/- 22% SD) of MCNs were positively correlated with tectal multisensory integration.

597 The correlation between tectal multisensory integration and activity in hindbrain neurons driving fast
598 motor responses provides not only a functional link between these two structures but a mechanistic basis
599 for the multisensory enhancement observed at the behavioral level (Fig. 1). The Tectum neurons receive
600 visual and auditory afferences at different depths on the neuropil (Figs. 2, 3). On neurons insensitive to
601 unisensory stimulation, multisensory input could result in suprathreshold responses, resulting in a larger
602 number of active PVL neurons. In turn, this increased tectal input to premotor neurons such as the
603 Mauthner cell could increase the probability of a motor response, effectively lowering the threshold for
604 threat detection.

605

606 **Discussion**

607 This study provides the first comprehensive evidence of multisensory enhancement elicited by audiovisual
608 stimuli across neural and behavioral levels in zebrafish. Using visual and auditory danger cues, we
609 observed robust multisensory integration in the Tectum. We identified motion-correlated neurons in the
610 hindbrain which also show an increased response to multisensory stimulation. On a behavioral level,
611 multisensory stimulation significantly enhanced escape response probability. This paradigm enables an
612 examination of how multisensory processing at the neural level influences overt behavioral decisions that,
613 to date, has not been explored in zebrafish and remains relatively rare in the literature on other species.

614

615 *Auditory, visual and multisensory units in the zebrafish Tectum*

616 Work in zebrafish and other species suggests that deeper tectal layers receive non-visual sensory inputs
617 (Knudsen, 1982; Kardamakis et al., 2016; Vanwalleghem et al., 2017; Isa et al., 2021), supporting the role
618 of the Tectum as a multisensory integrative structure. Our findings build on this concept, demonstrating
619 that auditory information reaches the deeper layers of the zebrafish tectal neuropil. The auditory
620 response within these layers is robust and enhanced during multisensory stimulation (Fig. 4), suggesting
621 that integratory circuits that have been found in other vertebrates are anatomically conserved in
622 zebrafish.

623 We also identified a significant presence of neurons in the Tectum that respond to multiple sensory
624 modalities, a finding documented in other species but scarcely observed in zebrafish. Previous reports
625 using non-threatening cues estimated that less than 1% of tectal cells were responsive to both visual and
626 auditory stimulation (Thompson et al., 2016). Here we found a much larger proportion of audiovisual-
627 responsive cells which averaged 19% across experiments (Fig. 3C-D). These audiovisual neurons were
628 responsive to unisensory stimuli, but exhibited increased activation probability under multisensory
629 conditions (Fig. 3D), suggesting that neurons that receive audiovisual inputs have a higher firing
630 probability if those signals arrive simultaneously.

631 We also found a proportion of tectal neurons that were only recruited when the auditory stimulus was
632 simultaneous with the end of the visual expansion and never responded to either visual or auditory stimuli
633 independently (EM neurons, Fig. 5B-E-F). It is important to note that these neurons are likely not
634 exclusively multisensory for all possible visual and auditory stimuli (i.e. stronger unisensory stimulation
635 might evoke responses in these units), but under our stimulation regime they only show suprathreshold
636 activity when both inputs are present. This finding parallels what is known about the superior colliculus in
637 mammals, supporting a preserved functional role between these homologous brain structures (Stein &
638 Stanford, 2008; Meredith & Stein, 1986). Previous reports on the zebrafish Tectum from Thompson et al.
639 (2016) identified a suppression of visual responses when an auditory stimulus was presented
640 simultaneously. In contrast, we did not find suppression but enhancement of visual responses when the
641 auditory component was present (Fig. 4 and 5). One possible explanation for this discrepancy is the nature
642 of the stimuli used in each study. Unlike Thompson et al., which used long pure tones intended to cover
643 the larval hearing range, our study employed abrupt auditory stimuli specifically associated with aversive

644 contexts (Fig. 1). The Tectum is known to differentiate between cues of varying ecological significance,
645 such as prey and predator signals (Semmelhack et al., 2014; Temizer et al., 2015; Dunn et al., 2016). This
646 may specifically enhance integration for aversive or appetitive stimuli, while neutral or mismatched cues
647 may lead to multisensory suppression. Indeed, we found that when multisensory integration in the
648 Tectum was suppressive, activity in MCN was minimum (Fig. 6E) suggesting that no motor response would
649 be evoked. Further studies are needed to investigate the effect of stimulus valence in tectal multisensory
650 integration.

651 About 20% of tectal neurons did not respond to any of our stimuli (Fig. 5B). This is unsurprising given the
652 topographic organization of the Tectum and the spatially localized nature of our visual stimulus within the
653 animal's visual field. However, previous studies in zebrafish (Thompson et al., 2016) indicate that certain
654 neurons respond exclusively to non-audiovisual modalities. Specifically in zebrafish, tectal neurons known
655 to be selectively responsive to water flow stimuli receive inputs from the lateral line nerve (Fame et al.,
656 2006). Thus, lack of response in a subpopulation of neurons might be indicative of a modality not tested
657 in our study.

658 Finally, it is noteworthy that we observed a pronounced inverse effectiveness pattern, particularly in the
659 neuropil (Fig. 4E-F), where the dendrites of tectal neurons are located (Hiramoto & Cline, 2009; Baier &
660 Scott, 2024). This anatomical arrangement suggests that subthreshold effects of dendritic integration
661 could be most visible in this area compared to the somas in the PVL. The presence of inverse effectiveness
662 in the zebrafish Tectum mirrors findings in the superior colliculus of mammals and the Tectum in other
663 vertebrate models (Meredith & Stein, 1986; Stein et al., 2008; Gil-Guevara, Bernal, & Riveros, 2022; Sumby
664 & Pollack, 1954), suggesting an evolutionary conserved multisensory integration mechanism.

665

666 *A functional relationship between tectal integration and behavior*

667 In this work, we introduce a methodology for extracting motor event data when simultaneously recording
668 tail movements is not feasible. We demonstrated that these movement events are reliably evoked by both
669 visual and auditory stimuli and show multisensory integration consistent with our behavioral findings (Fig.
670 6A). Additionally, we showed that, even with this inferred movement data, it is possible to identify
671 neurons correlated with these motor events, including both spontaneous movements and stimulus-
672 evoked responses (Fig. 6B-C). In freely moving and in agarose embedded fish, auditory stimuli evoked
673 slightly higher response rates than visual stimuli (Fig. 1D and 6A). The fact that motion events of agarose-
674 embedded fish were associated with the presentation of danger stimuli and that they exhibit multisensory
675 enhancement qualitatively similar to freely-swimming fish suggests that these events represent escape-
676 like attempts.

677 We report for the first time a relationship between the multisensory enhancement observed in the
678 Tectum and the activity of motion-correlated neurons in the hindbrain, which receive projections from
679 the Tectum (Del Bene et al., 2010; Robles et al., 2011; Kardamakis et al., 2016; Baier & Scott, 2024).
680 Hindbrain neurons drive premotor activity—for instance, prepontine neurons in rhombomere 1 are
681 implicated in evoking long latency C-starts (Marquart et al., 2019) or the Mauthner cells, in rhombomere

682 4, are responsible of the short latency C-start. Here we show that multisensory enhancement in the
683 Tectum correlates with an increase in the likelihood of premotor activity (Fig. 6-G). This, in turn, may
684 contribute to the behavioral enhancement we observe, both in agarose-embedded and freely moving
685 animals (Fig. 1), implying a causal link between tectal and behavioral multisensory enhancement.

686 Our work shows that the zebrafish Tectum integrates multisensory stimuli, producing both sublinear and
687 supralinear enhancement at the single-neuron level and at the population level. This multisensory activity
688 is closely related to the activation of functionally identifiable neurons in the hindbrain. Multisensory
689 integration increases the activity of these neurons and in turn enhances escape probability, particularly
690 when unisensory components are of low salience providing critical information for threat detection.

691 **References**

692 Baier, H., & Scott, E. K. (2024). The visual systems of zebrafish. Annual Review of Neuroscience, 47.

693 Barker, A. J., & Baier, H. (2015). Sensorimotor decision making in the zebrafish Tectum. Current Biology,
694 25(21), 2804–2814.

695 Bhattacharyya, K., McLean, D. L., & MacIver, M. A. (2017). Visual threat assessment and reticulospinal
696 encoding of calibrated responses in larval zebrafish. Current Biology, 27(18), 2751–2762.

697 Burgess, H. A., & Granato, M. (2007). Sensorimotor gating in larval zebrafish. Journal of Neuroscience,
698 27(18), 4984–4994.

699 Del Bene, F., Wyart, C., Robles, E., Tran, A., Looger, L., Scott, E. K., ... & Baier, H. (2010). Filtering of visual
700 information in the Tectum by an identified neural circuit. Science, 330(6004), 669–673.

701 Dunn, T. W., Gebhardt, C., Naumann, E. A., Riegler, C., Ahrens, M. B., Engert, F., & Del Bene, F. (2016).
702 Neural circuits underlying visually evoked escapes in larval zebrafish. Neuron, 89(3), 613–628.

703 Fame, R. M., Brajon, C., & Ghysen, A. (2006). Second-order projection from the posterior lateral line in
704 the early zebrafish brain. Neural Development, 1, 4.

705 Fleisch, V. C., & Neuhauss, S. C. (2006). Visual behavior in zebrafish. Zebrafish, 3(2), 191–201.

706 Foreman, M. B., & Eaton, R. C. (1993). The direction change concept for reticulospinal control of goldfish
707 escape. Journal of Neuroscience, 13(10), 4101–4113.

708 Gabbiani, F., Krapp, H. G., Koch, C., & Laurent, G. (2002). Multiplicative computation in a visual neuron
709 sensitive to looming. Nature, 420(6913), 320–324.

710 Ghosh, D. D., Nitabach, M. N., Zhang, Y., & Harris, G. (2017). Multisensory integration in C. elegans.
711 Current Opinion in Neurobiology, 43, 110–118.

712 Gil-Guevara, O., Bernal, H. A., & Riveros, A. J. (2022). Honey bees respond to multimodal stimuli
713 following the principle of inverse effectiveness. Journal of Experimental Biology, 225.

714 Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B. L., Koay, S. A., ... & Pnevmatikakis, E. A.
715 (2019). CaImAn an open source tool for scalable calcium imaging data analysis. eLife, 8, e38173.

716 Hartline, P. H., Kass, L., & Loop, M. S. (1978). Merging of modalities in the optic Tectum: infrared and
717 visual integration in rattlesnakes. Science, 199(4334), 1225–1229.

718 Hiramoto, M., & Cline, H. T. (2009). Convergence of multisensory inputs in Xenopus tadpole Tectum.
719 Developmental Neurobiology, 69(14), 959–971.

720 Isa, T., Marquez-Legorreta, E., Grillner, S., & Scott, E. K. (2021). The Tectum/superior colliculus as the
721 vertebrate solution for spatial sensory integration and action. Current Biology, 31(11), R741–R762.

Kalueff, A. V., Gebhardt, M., Stewart, A. M., Cachat, J. M., Brimmer, M., Chawla, J. S., ... & Schneider, and the Zebrafish Neuroscience Research Consortium, H. (2013). Towards a comprehensive catalog of zebrafish behavior 1.0 and beyond. Zebrafish, 10(1), 70–86.

Kardamakis, A. A., Perez-Fernandez, J., & Grillner, S. (2016). Spatiotemporal interplay between multisensory excitation and recruited inhibition in the lamprey optic Tectum. eLife, 5, e16472.

Kimmel, C. B., Patterson, J., & Kimmel, R. O. (1974). The development and behavioral characteristics of the startle response in the zebra fish. Developmental Psychobiology, 7(1), 47–60.

Knudsen, E. I. (1982). Auditory and visual maps of space in the optic Tectum of the owl. Journal of Neuroscience, 2(9), 1177–1194.

Kohashi, T., & Oda, Y. (2008). Initiation of Mauthner-or non-Mauthner-mediated fast escape evoked by different modes of sensory input. Journal of Neuroscience, 28(42), 10641–10653.

Liu, K. S., & Fetcho, J. R. (1999). Laser ablations reveal functional relationships of segmental hindbrain neurons in zebrafish. Neuron, 23(2), 325–335.

Lukas, J., Romanczuk, P., Klenz, H, Klamser, P., Arias Rodriguez, L., Krause, J., & Bierbach, D. (2021). Acoustic and visual stimuli combined promote stronger responses to aerial predation in fish. Behavioral Ecology, 32(6), 1094–1102.

Luksch, H. (2009). Optic Tectum: Sensorimotor Integration. , 263–269.

Macaluso, E., & Driver, J. (2005). Multisensory spatial interactions: a window onto functional integration in the human brain. Trends in Neurosciences, 28(5), 264–271.

Marques, J. C., Lackner, S., Félix, R., & Orger, M. B. (2018). Structure of the zebrafish locomotor repertoire revealed with unsupervised behavioral clustering. Current Biology, 28(2), 181–195.

Marquart, G. D., Tabor, K. M., Bergeron, S. A., Briggman, K. L., & Burgess, H. A. (2019). Prepontine non-giant neurons drive flexible escape behavior in zebrafish. PLOS Biology, 17(10), e3000480.

Marquez-Legorreta, E., Piber, M., & Scott, E. K. (2020). Visual escape in larval zebrafish: stimuli, circuits, and behavior. In Behavioral and Neural Genetics of Zebrafish (pp. 49–71). Academic Press.

Martorell, N., & Medan, V. (2022). Audiovisual integration in the Mauthner cell enhances escape probability and reduces response latency. Scientific Reports, 12(1), 1097.

Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., & Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nature Neuroscience, 21(9), 1281–1289.

McIntyre, C., & Preuss, T. (2019). Influence of stimulus intensity on multimodal integration in the startle escape system of goldfish. Frontiers in Neural Circuits, 13, 7.

Medan, V., & Preuss, T. (2014). The Mauthner-cell circuit of fish as a model system for startle plasticity. Journal of Physiology–Paris, 108(2–3), 129–140.

Meredith, M. A., & Stein, B. E. (1986). Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. Journal of Neurophysiology, 56(3), 640–662.

Nevin, L. M., Robles, E., Baier, H., & Scott, E. K. (2010). Focusing on optic Tectum circuitry through the lens of genetics. BMC Biology, 8, 1–10.

Ohyama, T., Schneider-Mizell, C. M., Fetter, R. D., Aleman, J. V., Franconville, R., Rivera-Alba, M., ... & Zlatic, M. (2015). A multilevel multimodal circuit enhances action selection in Drosophila. Nature, 520(7549), 633–639.

Olcese, U., Iurilli, G., & Medini, P. (2013). Cellular and synaptic architecture of multisensory integration in the mouse neocortex. Neuron, 79(3), 579–593.

Poulsen, R. E., Scholz, L. A., Constantin, L., Favre-Bulle, I., Vanwalleghem, G. C., & Scott, E. K. (2021). Broad frequency sensitivity and complex neural coding in the larval zebrafish auditory system. Current Biology, 31(9), 1977–1987.

Preuss, T., Osei-Bonsu, P. E., Weiss, S. A., Wang, C., & Faber, D. S. (2006). Neural representation of object approach in a decision-making motor circuit. Journal of Neuroscience, 26(13), 3454–3464.

Privat, M., Romano, S. A., Pietri, T., Jouary, A., Boulanger-Weill, J., Elbaz, N., ... & Sumbre, G. (2019). Sensorimotor transformations in the zebrafish auditory system. Current Biology, 29(23), 4010–4023.

Randlett, O., Wee, C. L., Naumann, E. A., Nnaemeka, O., Schoppik, D., Fitzgerald, J. E., ... & Schier, A. F. (2015). Whole-brain activity mapping onto a zebrafish brain atlas. Nature Methods, 12(11), 1039–1046.

Robles, E., Smith, S. J., & Baier, H. (2011). Characterization of genetically targeted neuron types in the zebrafish optic Tectum. Frontiers in Neural Circuits, 5, 1.

Rowland, B. A., Quessy, S., Stanford, T. R., & Stein, B. E. (2007). Multisensory integration shortens physiological response latencies. Journal of Neuroscience, 27(22), 5879–5884.

Sajovic, P., & Levinthal, C. (1982). Visual cells of zebrafish optic Tectum: mapping with small spots. Neuroscience, 7(10), 2407–2426.

Scott, E. K., & Baier, H. (2009). The cellular architecture of the larval zebrafish Tectum, as revealed by gal4 enhancer trap lines. Frontiers in Neural Circuits, 3, 712.

Semmelhack, J. L., Donovan, J. C., Thiele, T. R., Kuehn, E., Laurell, E., & Baier, H. (2014). A dedicated visual pathway for prey detection in larval zebrafish. eLife, 3, e04878.

Stanford, T. R., Quessy, S., & Stein, B. E. (2005). Evaluating the operations underlying multisensory integration in the cat superior colliculus. Journal of Neuroscience, 25(28), 6499–6508.

Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. Nature Reviews Neuroscience, 9(4), 255–266.

Stein, B. E., Stanford, T. R., Ramachandran, R., Perrault, T. J., Jr., & Rowland, B. A. (2009). Challenges in quantifying multisensory integration: alternative criteria, models, and inverse effectiveness. Experimental Brain Research, 198(2–3), 113–126.

Stevenson, R. A., Ghose, D., Fister, J. K., Sarko, D. K., Altieri, N. A., Nidiffer, A. R., ... & Wallace, M. T. (2014). Identifying and quantifying multisensory integration: a tutorial review. Brain Topography, 27, 707–730.

Sumby, W. H., & Pollack, I. (1954). Visual contribution to speech intelligibility in noise. The Journal of the Acoustical Society of America, 26(2), 212–215.

Temizer, I., Donovan, J. C., Baier, H., & Semmelhack, J. L. (2015). A visual pathway for looming-evoked escape in larval zebrafish. Current Biology, 25(14), 1823–1834.

Thompson, A. W., Vanwalleghem, G. C., Heap, L. A., & Scott, E. K. (2016). Functional profiles of visual-, auditory-, and water flow-responsive neurons in the zebrafish Tectum. Current Biology, 26(6), 743–754.

Van Atteveldt, N., Murray, M. M., Thut, G., & Schroeder, C. E. (2014). Multisensory integration: flexible use of general operations. Neuron, 81(6), 1240–1253.

Vanwalleghem, G., Heap, L. A., & Scott, E. K. (2017). A profile of auditory-responsive neurons in the larval zebrafish brain. Journal of Comparative Neurology, 525(14), 3031–3043.

Wallace, M. T., Meredith, M. A., & Stein, B. E. (1993). Converging influences from visual, auditory, and somatosensory cortices onto output neurons of the superior colliculus. Journal of Neurophysiology, 69(6), 1797–1809.

Whitchurch, E. A., & Takahashi, T. T. (2006). Combined auditory and visual stimuli facilitate head saccades in the barn owl (Tyto alba). Journal of Neurophysiology, 96(2), 730–745.

Zottoli, S. J. (1977). Correlation of the startle reflex and Mauthner cell auditory responses in unrestrained goldfish. Journal of Experimental Biology, 66(1), 243–254.

Zottoli, S. J., Hordes, A. R., & Faber, D. S. (1987). Localization of optic tectal input to the ventral dendrite of the goldfish Mauthner cell. Brain Research, 401(1), 113–121.

Zottoli, S. J., Newman, B. C., Rieff, H. I., & Winters, D. C. (1999). Decrease in occurrence of fast startle responses after selective Mauthner cell ablation in goldfish (Carassius auratus). Journal of Comparative Physiology A, 184, 207–218.

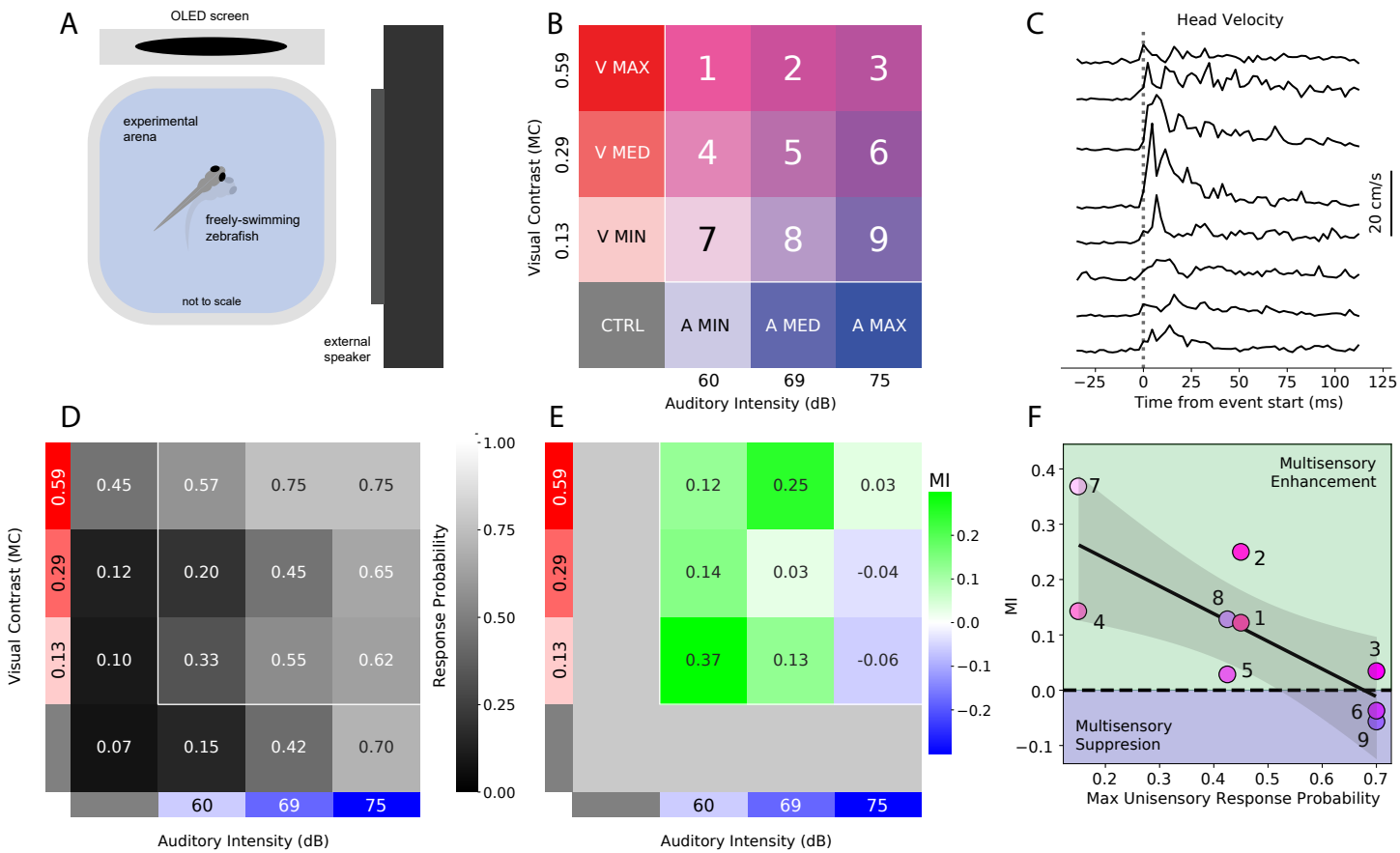


Figure 1. Multisensory enhancement of escape probability is inversely related to stimulus intensity. (A) Experimental arena used to present visual and auditory stimuli to freely-swimming zebrafish (N = 80). (B) Stimulus grid containing three visual contrasts (V min, V med, V max), three auditory intensities (A min, A med, A max), their nine combinations (1-9) and a blank stimulus (Control, zero visual contrast and zero auditory amplitude) condition. Multisensory stimuli are numbered to be used as reference in Fig. 1F. MC: Michelson Contrast. (C) Eight representative examples of motor events. Fast events are defined as those which surpass a threshold of 6 cm/s. Traces are aligned to the event's start time (vertical dotted line). (D) Escape response probability for each unisensory stimulus and for each combination of visual and auditory intensities (n = 40). (E) Multisensory Integration coefficient (MI) for each combination of visual and auditory intensities. Color code indicates MI strength. (F) MI for each multisensory condition, as a function of maximum unisensory response probability. Each point represents a single multisensory condition corresponding to a combination in Fig. 1B, averaged across all fish. Black line represents a linear fit and its confidence interval.

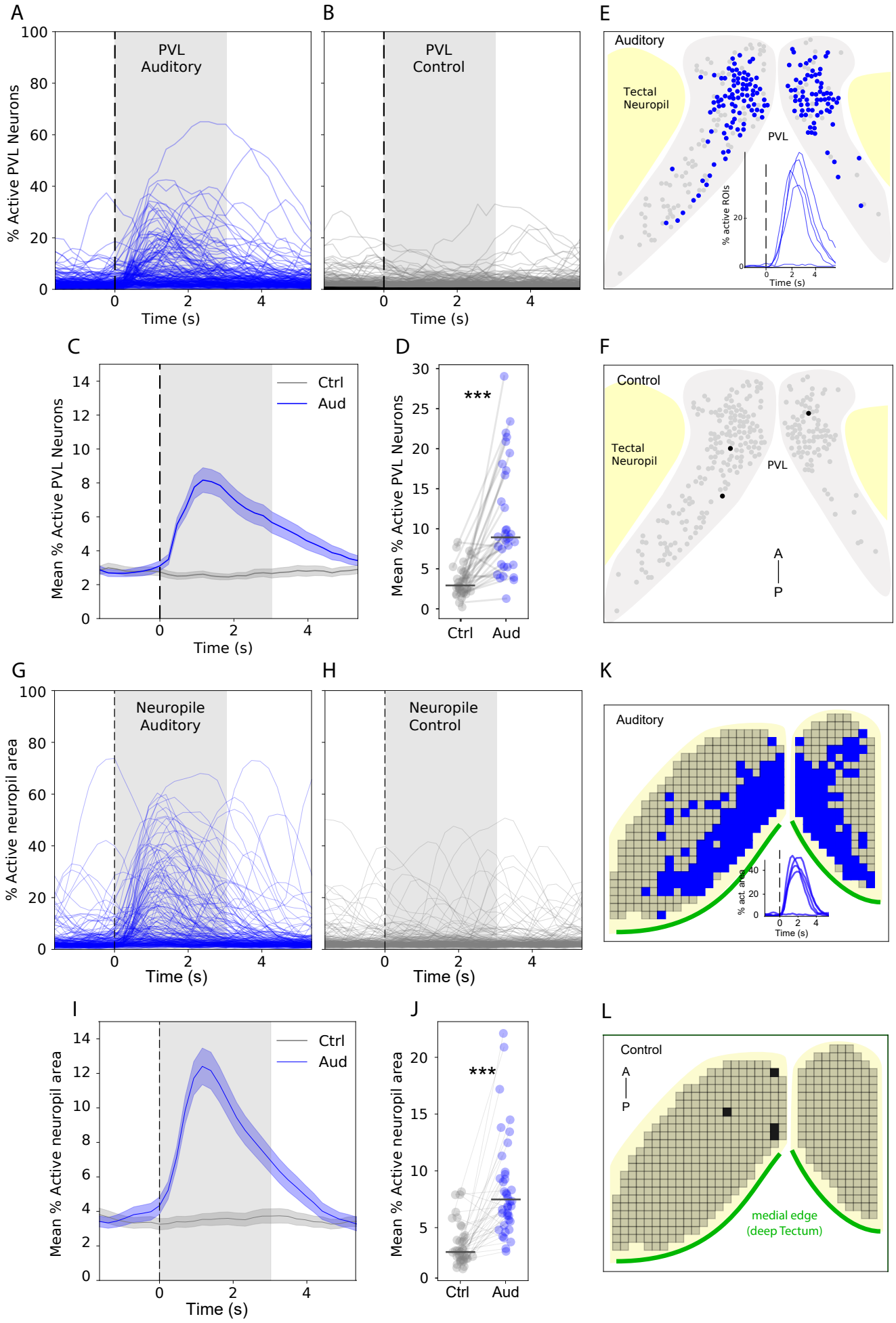

A
PVL
Auditory
% Active PVL Neurons
Time (s)
B
PVL
Control
Time (s)
E
Auditory
Tectal
Neuropil
PVL
% active ROIs
Time (s)
C
Ctrl
Aud
Mean % Active PVL Neurons
Time (s)
D
***
Mean % Active PVL Neurons
Ctrl
Aud
F
Control
Tectal
Neuropil
PVL
A
P
G
Neuropile
Auditory
% Active neuropil area
Time (s)
H
Neuropile
Control
Time (s)
K
Auditory
% act. area
Time (s)
I
Ctrl
Aud
Mean % Active neuropil area
Time (s)
J
***
Mean % Active neuropil area
Ctrl
Aud
L
Control
A
P
medial edge
(deep Tectum)

Figure 2. Auditory stimulation elicits neural responses in the Optic Tectum of zebrafish. (A) Percentage of active PVL Neurons during auditory trials. Each line represents a trial. Auditory responses are analyzed over a 3-s period (shaded area) after sound presentation (vertical dashed line, 5 trials per imaged plane, nplanes = 33, ntrials = 165). (B) Same as A for control trials (blank stimulus). (C) Mean (+/- SE) proportion of active ROIs in the PVL over time as a response to auditory (blue) and control (gray) trials, across experiments. (D) Mean proportion of active ROIs in the PVL after stimulation (shaded period in A-B) in auditory and control trials for each experiment. Connected data points correspond to single experiments, averaged across 5 trials for each condition. ***$p<0.001$. (E) Distribution of active PVL ROIs (blue dots) during the analysis window (3.5 s after end of stimulus presentation) for a representative example. % of active ROIs for each trial of the depicted example is shown in the inset. PVL is outlined in grey while the yellow area represents the neuropile. (F) Distribution of active PVL ROIs (black dots) during control trials of the same experiment as in E. (G) Auditory stimuli increase the proportion of active tectal neuropil area, compared to spontaneous activation (H) (5 trials per animal, n = 33). Inset shows neuropil regions which were activated in at least one trial for the illustrated example. (I) Mean proportion of active neuropil area over time as a response to auditory (blue) and control (gray) trials, across experiments. (J) Proportion of active neuropil area during the stimulus window in control and auditory trials, for each experiment. Connected data points correspond to single experiments, averaged across 5 trials for each condition. ***$p<0.001$. (K) Distribution of active neuropile area (blue squares) during the analysis window (3.5 s after end of stimulus presentation) for a representative example. % of active area for each trial of the depicted example is shown in the inset. (L) Distribution of active neuropile area during control trials of the same experiment as in K. Green curves show the medial edge of the tectal neuropils, corresponding to deeper layers of the Tectum.

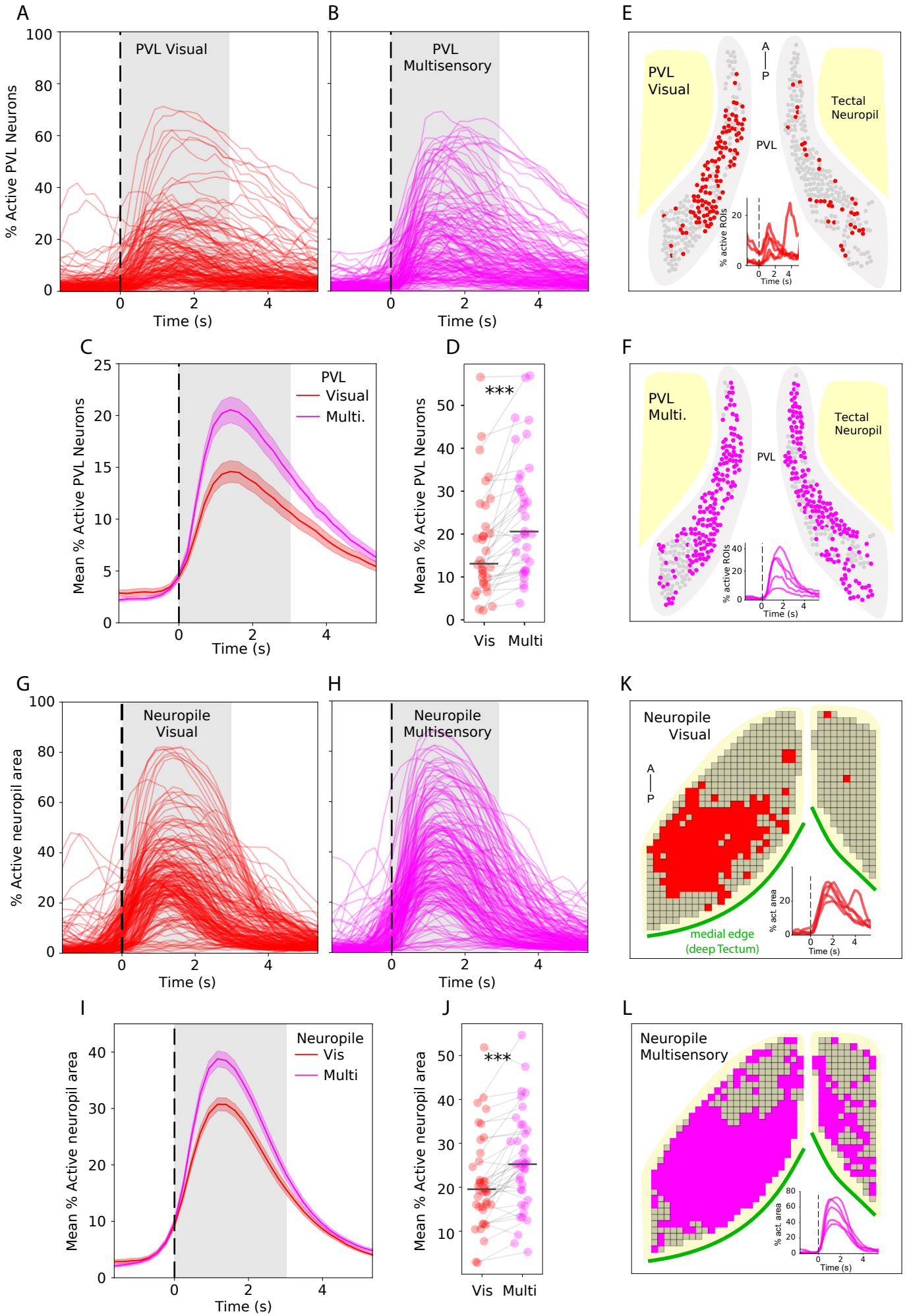

A
PVL Visual
% Active PVL Neurons
Time (s)
B
PVL Multisensory
Time (s)
E
PVL Visual
A
P
Tectal Neuropil
PVL
% active ROIs
Time (s)
C
PVL
Visual
Multi.
Mean % Active PVL Neurons
Time (s)
D
***
Mean % Active PVL Neurons
Vis
Multi
F
PVL Multi.
Tectal Neuropil
PVL
% active ROIs
Time (s)
G
Neuropile Visual
% Active neuropil area
Time (s)
H
Neuropile Multisensory
Time (s)
K
Neuropile Visual
A
P
medial edge (deep Tectum)
% act area
Time (s)
I
Neuropile
Vis
Multi
Mean % Active neuropil area
Time (s)
J
***
Mean % Active neuropil area
Vis
Multi
L
Neuropile Multisensory
% act area
Time (s)

Figure 3. Multisensory stimulation enhances neural responses over a unisensory visual stimulus in the Tectum of zebrafish. (A) Percentage of active PVL Neurons during visual trials. Each line represents a trial, 5 trials per imaged plane, nplanes = 33, ntrials = 165). The vertical dashed line represents the end of the visual loom expansion (B) Multisensory trials (5 trials per imaged plane, nplanes = 33, ntrials = 165) show significant activation after the stimulus presentation (vertical dashed line: sound presentation concurrent with end of looming expansion). (C) Mean proportion of active ROIs in the PVL over time as a response to multisensory (magenta) and visual (red) trials, across experiments. (D) Mean (+/- SE) proportion of active ROIs in the PVL after stimulation time (shaded window in A-B) in visual and multisensory trials for each experiment. Pairs of points correspond to single experiments, averaged across 5 trials for each condition. ***$p<0.001$. (E) Distribution of active PVL ROIs (red dots) during the analysis window (3.5 s after end of visual presentation) for a representative example. Percentage of active ROIs for each trial of the depicted example is shown in the inset. (F) Distribution of active PVL ROIs (magenta dots) during multisensory trials of the same experiment as in E showed a more distributed and bilateral activation. Percentage of active ROIs for each trial of the depicted example is shown in the inset. (G) Visually-evoked activity in the neuropile. (H) Multisensory trials evoke a significant activation of the neuropile. (I) Mean (+/- SE) proportion of active neuropil area over time as a response to visual (red) and multisensory (H, magenta) trials, across experiments. (J) Mean proportion of active neuropil area during the stimulus window in visual and multisensory trials for each experiment. Pairs of points correspond to single experiments, averaged across 5 trials for each condition. ***$p<0.001$. (K) Distribution of active neuropile area (red squares) during the analysis window (3.5 s after end of visual stimulus) for a representative example. Percentage of active area for each trial of the depicted example is shown in the inset. (L) Distribution of active neuropile area during multisensory trials of the same experiment as in K. Green curves show the medial edge of the tectal neuropile.

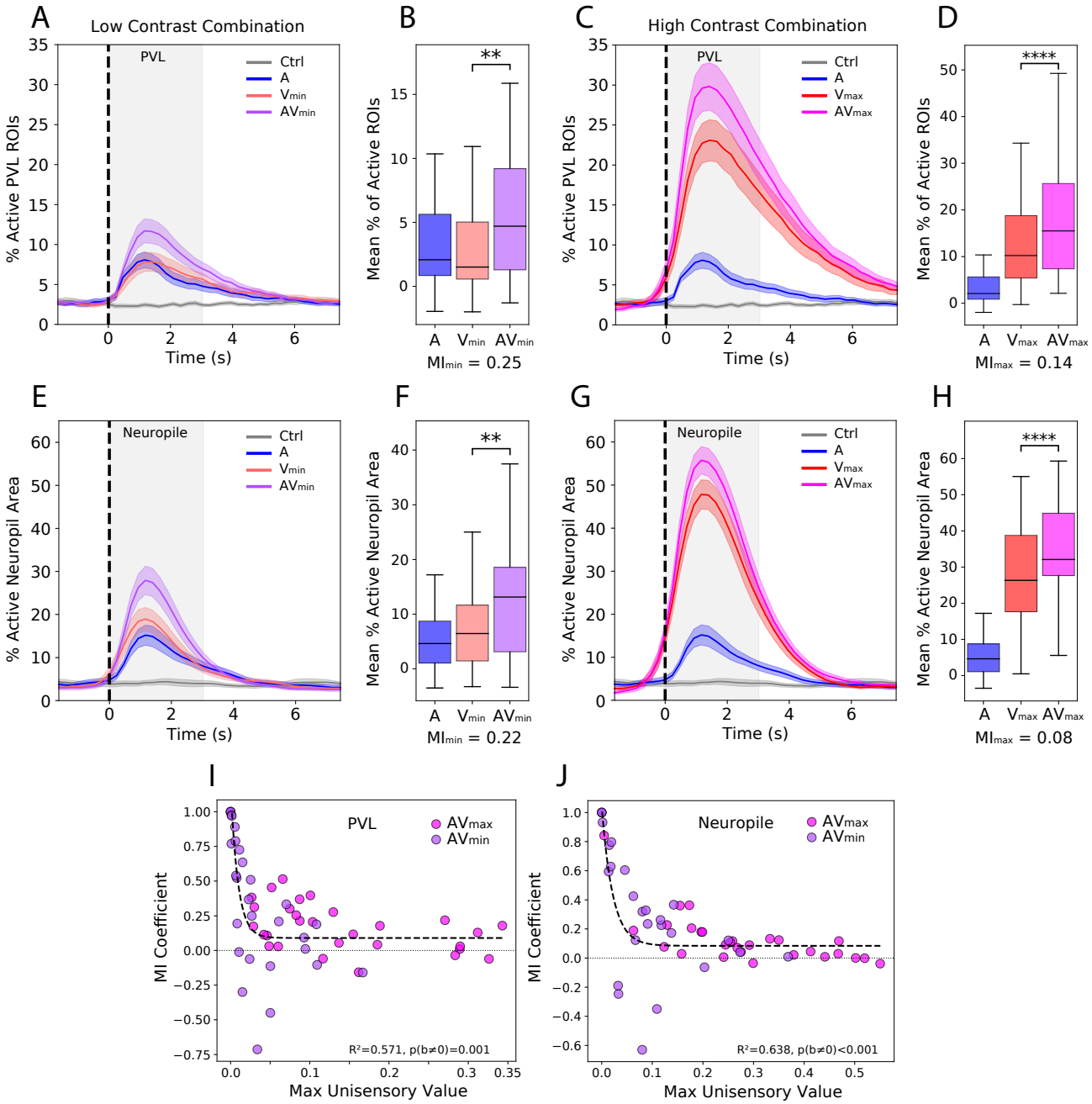


Figure 4. Inverse effectiveness in the zebrafish Tectum. (A) Mean proportion of active ROIs over time in response to control (grey), auditory (blue), weak visual (Vlow, light red) and multisensory (AVlow, purple) stimuli. (B) Mean proportion of active ROIs during the stimulation window subtracted from the control. **p=0.004. (C) Mean proportion of active ROIs over time in the PVL as a response to control (grey), auditory (blue), Vhigh (red) and multisensory (AVhigh, magenta) stimuli. (D) Mean proportion of active ROIs during the stimulation window subtracted from the control. ***p<0.001. (E-F) Same as A-B for the neuropile. (G-H) Same as C-D for the neuropile. (I) MI in the PVL plotted against the maximum unisensory activation probability, for each experiment's low contrast (AVlow, purple) and high contrast multisensory stimulus (AVhigh, magenta). Linear regression was performed on each subset (dashed purple and magenta lines). (J) Same as I for the neuropile.

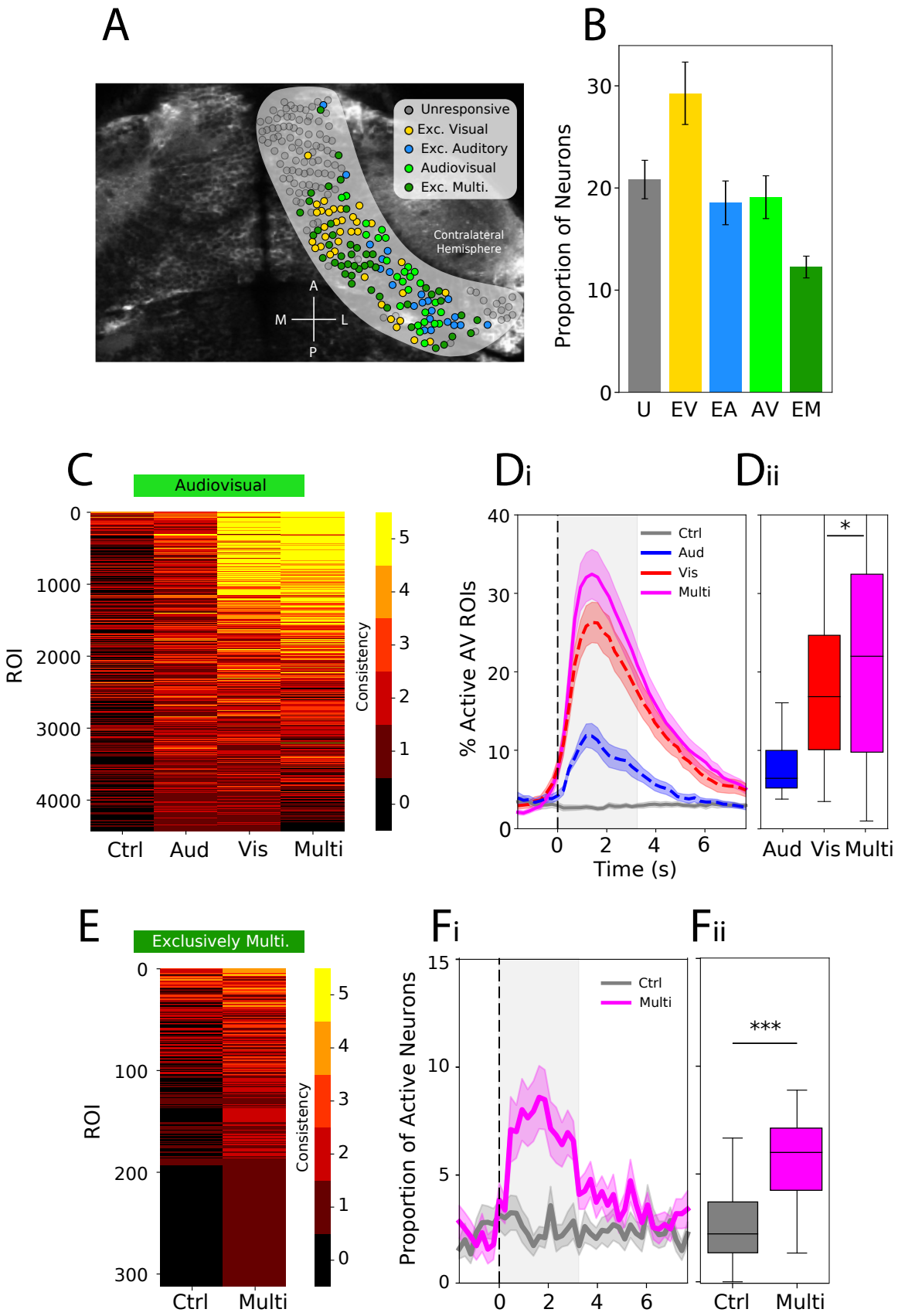
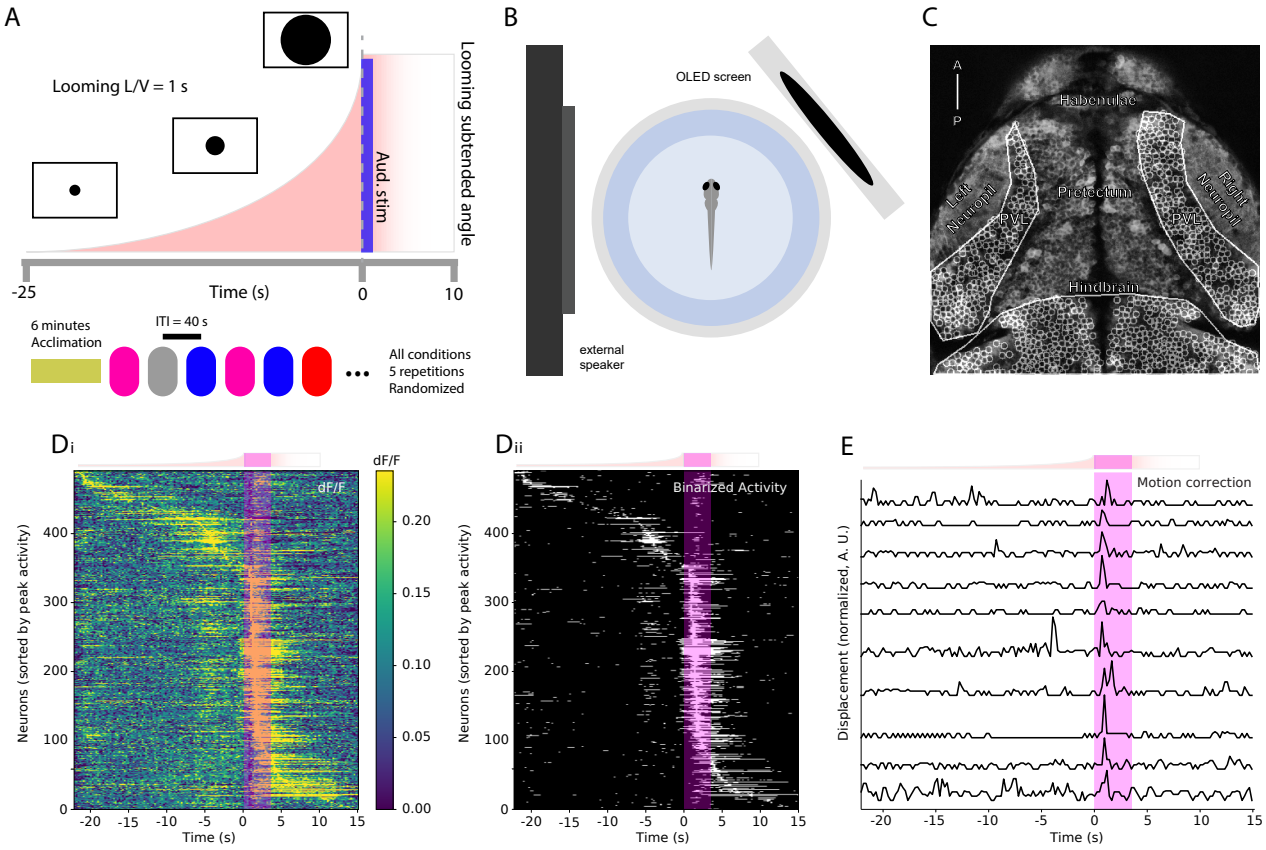


Figure 5. Multisensory stimulation recruits unresponsive PVL neurons and boosts response probability of audiovisual neurons. (A) Representative example of a PVL's neurons colored according to their response profile for the hemitectum contralateral to the visual stimulus. ROIs are overlaid on top of the template used for anatomical segmentation. (B) Proportion of neurons in each of the five neuron classes across experiments (mean +/- SD: U: 21 +/- 2, EA: 19 +/- 2, EV: 29 +/- 3, AV: 19 +/- 2, EM: 12 +/- 1). (C) Response consistency for all AV ROIs, for the control, auditory, visual and multisensory conditions. ROIs are sorted by their mean consistency across stimuli. (Di) Mean (+/- SE) proportion of active AV ROIs over time for each stimulation condition. (Dii) Mean peak proportion of active AV ROIs for auditory, visual and multisensory trials. *p = 0.012. (E) Response consistency for all EM ROIs, for the control and multisensory conditions. ROIs are sorted by their mean consistency across stimuli. (Fi) Mean (+/- SE) proportion of active EM ROIs over time for each stimulation condition. (Fii) Mean peak proportion of active EM ROIs for control and multisensory trials. ***p<0.001.



Supplementary Figure 1. Visual and auditory stimuli evoke tectal activity and motor events in agarose-embedded fish. (A) Experimental procedure and stimulus coordination (see Methods). Ovals refer to control (grey), auditory (blue), visual (red), multisensory (magenta) randomized trials. (B) Setup diagram for imaging experiments. (C) Representative example of ROI segmentation. Visually-defined polygons for the Hindbrain and PVL are shown, with semi-automatic segmentation of circular ROIs inside each polygon. Tectal neuropil was not segmented using this method but subdivided in a grid pattern. ROIs located outside selected polygons were ignored, as were ROIs which were never observed to be active. Hindbrain, PreTectum and Habenulae are labeled for reference. (D) Representative example of neural activity. (Di) Individual traces of ΔF/F for each PVL ROI (n = 491) during a multisensory trial, sorted from top to bottom by time of peak activity. The region shaded in magenta represents the time window used to define neural response to stimulation. (Dii) Binarized ΔF/F, as described in Methods, is shown for each PVL ROI during the same multisensory trial. (E) Ten representative examples of image displacement traces, captured by a motion-correction algorithm as described in Methods, in trials where a motion event was detected during the stimulation window (magenta shading).

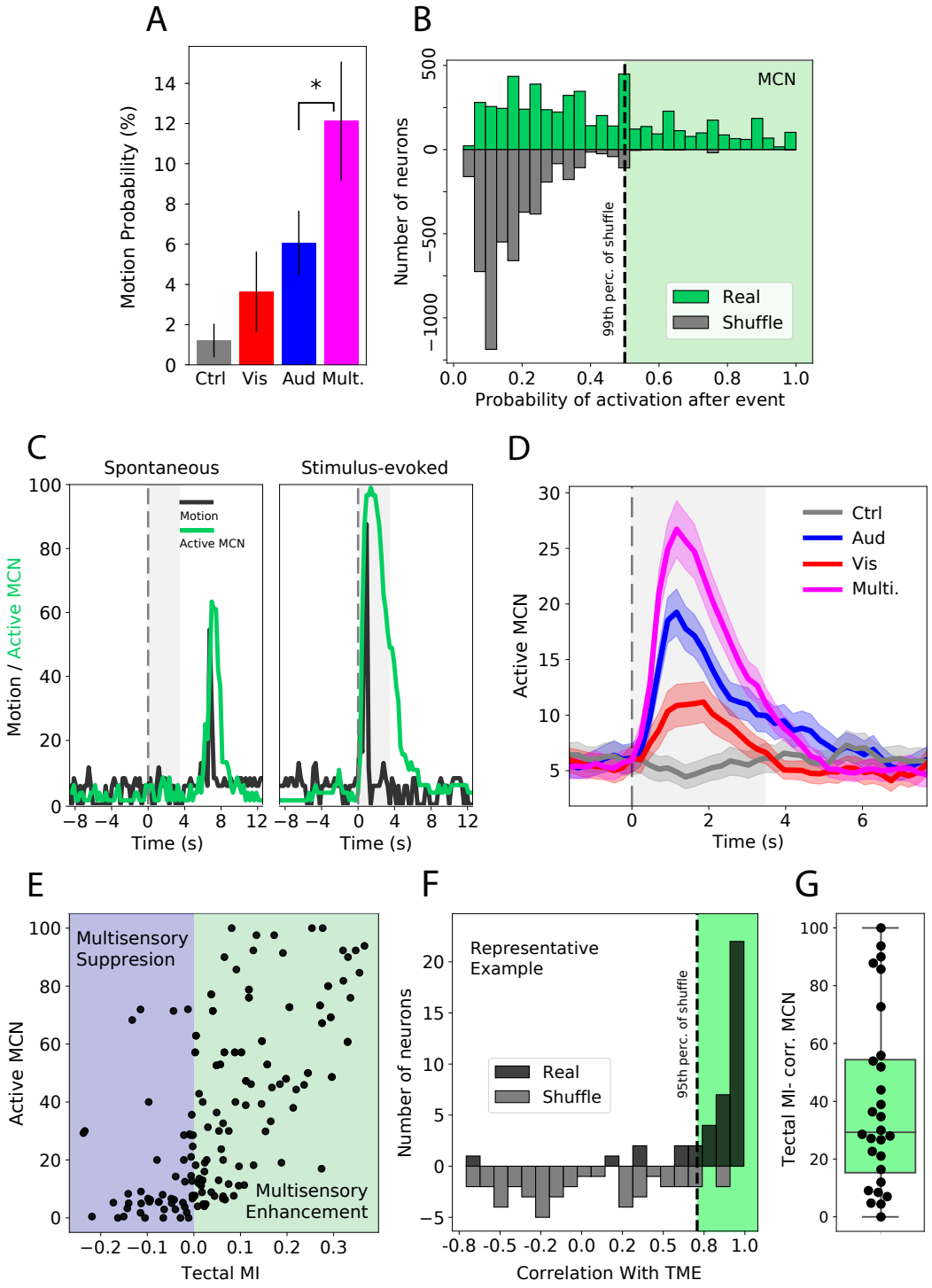
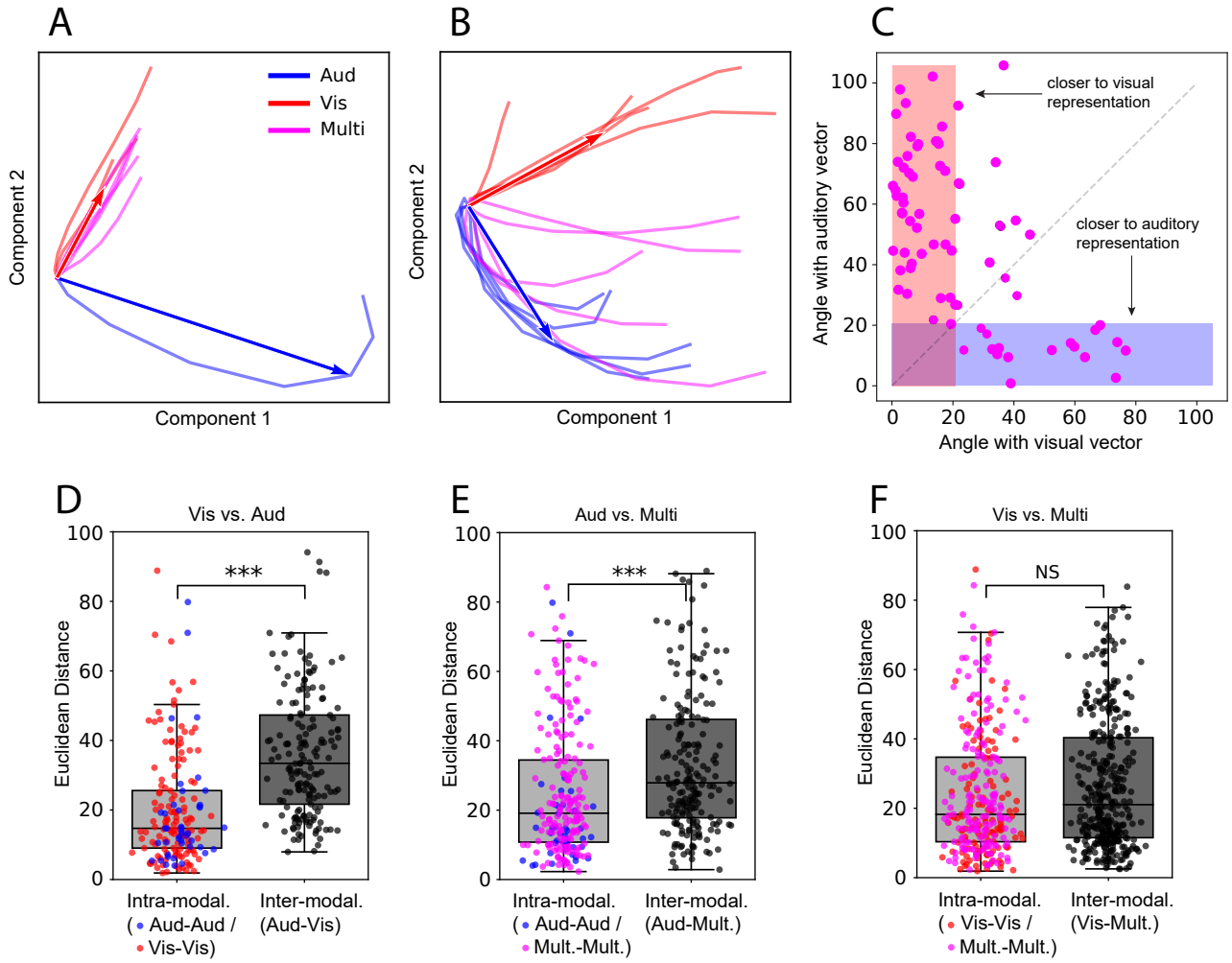


Figure 6. Motion-correlated neurons in the Hindbrain are mostly responsive to multisensory stimulation when tectal activity shows multisensory enhancement. (A) Motion event probability for each stimulation condition. Bars represent the mean motion probability across all experiments. *p = 0.029. (B) Distribution of the activation probability after a motion event, for all recorded Hindbrain neurons across experiments, computed for real events (green) or when shuffling times (gray). Only events outside the stimulation window were considered, to avoid biasing this analysis with the response to stimulation. Vertical dashed line represents 99th percentile of shuffle distribution. (C) The synchronous activation of motion-correlated neurons in the Hindbrain is highly indicative of spontaneous and stimulus-evoked motion events. Left: representative example of a spontaneous event (occurring outside the stimulation window). The black trace shows the displacement of the image acquired by the confocal microscope (arbitrary units), while the green trace shows the proportion of active motion neurons in the Hindbrain. Right: Representative example of a stimulus-evoked event (occurring shortly after stimulus presentation). (D) Mean proportion of active MCNs over time, across all experiments, as a response to control, auditory, visual or multisensory stimuli. (E) Activation of MCNs in Hindbrain as a function of Tectal MI. Each black dot represents a trial. Progressively stronger multisensory enhancement in the Tectum correlates with a higher proportion of active MCN. (F) A proportion of MCNs in Hindbrain are associated with Tectal MI. Correlation between ΔF/F and Tectal MI for each MCN, for each of the 5 multisensory trials in a representative experiment. Black bars represent the distribution of correlation coefficients, while gray bars show the results of shuffling trial order to remove all significant associations. Vertical dashed line represents 95th percentile of shuffle distribution. Green-shaded region shows neurons with correlation coefficients in the top 5% of what would be expected by chance. (G) Proportion of MCNs whose ΔF/F was significantly correlated with Tectal MI across multisensory trials. Each datapoint is a single experiment.



Supplementary Figure 2. Multisensory representation in the tectal PVL usually aligns closely to a single unisensory representation. (A-B) Representative examples of tPCA results for two different experiments. Trajectories of visual (red), auditory (blue), and multisensory (magenta) trials in the space defined by the first two principal components, in the 1.86 seconds after stimulation time, for each representative experiment. Arrows represent the median direction of visual and auditory trajectories. (C) Angle between each multisensory trajectory and the visual (x axis) and auditory (y axis) median directions for its corresponding experiment. Each point represents a single multisensory trial. The gray dashed line represents equal distance to both representations. (D) Euclidean distance between response trajectories of the same unisensory stimulus (Intra-modal, visual vs. visual, red, or auditory vs. auditory, blue) or between unisensory stimuli (Inter-modal., auditory vs. visual, black). Dots represent distances between individual trajectories compared to all other trajectories of the target class within an experiment. Boxplots represent the means of Intra-modality and Intermodality distances. ***p<0.001. (E) Same as D between response trajectories of the multisensory stimulus and auditory stimulus. ***p<0.001. (F) Same as D between response trajectories of the multisensory stimulus and visual stimulus. NS: statistically not significant.

Variable	Mean	Max	Min	Range	Feature Count
Head Velocity	x	x			2
End-Tail Velocity	x	x			2
Angular Velocity	x	x			2
Sum of Tail Angles	x	x		x	3
Mid-Tail Velocity	x	x			2
Head-Tail Distance	x		x		2
Head Acceleration	x	x		x	3
End-Tail Acceleration	x	x		x	3
Angular Acceleration	x	x		x	3
Head Jerk	x	x		x	3
End-Tail Jerk	x	x		x	3
Sum of Curvature	x	x		x	3
Angle Velocity Heading	x	x			2
Curvature Rate	x	x		x	3
Total					**36**

Supplementary Table 1: Kinematic descriptors of escape behaviors extracted from pose estimation data. Head Velocity: instantaneous speed of the head, calculated as the displacement of the midpoint between both eyes over time. End-Tail Velocity: instantaneous speed of the tail, derived from the movement of the mean point between the last two tail segments. Angular Velocity: rate of rotational change in the direction of the head, calculated from the angle between vectors connecting the head to a reference body point. Sum of Tail Angles: cumulative angular deviation between consecutive segments along the tail. Mid-Tail Velocity: instantaneous speed of the sixth tail segment, calculated from its displacement over time. Head-Tail Distance: linear distance between the head midpoint and the mean point of the tail's last segments. Head Acceleration: rate of change of the head's velocity. Tail Acceleration: rate of change of the tail's velocity. Angular Acceleration: rate of change in angular velocity. Head Jerk: rate of change of head acceleration. End-Tail Jerk: rate of change of tail acceleration. Sum of Curvature: cumulative discrete triangular curvature of the tail segments. Angle Velocity Heading: angle between the head's velocity vector and a reference vector from the head toward the body, indicating directional alignment of motion. Curvature Rate: rate of change of tail curvature over time. Note that from each of these parameters the Mean, Maximum, Minimum and/or Range was computed as individual variables used in the dimensionality reduction analysis.