



OPEN

Audiovisual integration in the Mauthner cell enhances escape probability and reduces response latency

Nicolás Martorell¹ & Violeta Medan^{1,2}✉

Fast and accurate threat detection is critical for animal survival. Reducing perceptual ambiguity by integrating multiple sources of sensory information can enhance perception and reduce response latency. However, studies addressing the link between behavioral correlates of multisensory integration and its underlying neural basis are rare. Fish that detect an urgent threat escape with an explosive behavior known as C-start. The C-start is driven by an identified neural circuit centered on the Mauthner cell, an identified neuron capable of triggering escapes in response to visual and auditory stimuli. Here we demonstrate that goldfish can integrate visual looms and brief auditory stimuli to increase C-start probability. This multisensory enhancement is inversely correlated to the salience of the stimuli, with weaker auditory cues producing a proportionally stronger multisensory effect. We also show that multisensory stimuli reduced C-start response latency, with most escapes locked to the presentation of the auditory cue. We make a direct link between behavioral data and its underlying neural mechanism by reproducing the behavioral data with an integrate-and-fire computational model of the Mauthner cell. This model of the Mauthner cell circuit suggests that excitatory inputs integrated at the soma are key elements in multisensory decision making during fast C-start escapes. This provides a simple but powerful mechanism to enhance threat detection and survival.

When confronted with potential threats in their environment, animals have to decide if and how to perform evasive movements to avoid harm. This is a powerful evolutionary force that led to the development of specialized sensory organs that extract qualitatively different information from a given event. The process of binding the different sensory signals associated to a single coherent event is called multisensory integration^{1,2}. Behaviorally, response improvement due to multisensory integration is often quantified by evaluating differences in the accuracy and speed of detection, localization and identification of stimuli^{1,3}. At the cellular level, integration performed by multisensory neurons is determined by synaptic convergence of sensory afferents belonging to different modalities, the neural operations that produce an “integrated” output and the interactions with other elements of the circuit or other brain areas². Multisensory integration becomes critically important for threat detection, when small reductions in sensory ambiguity can have a huge impact in the survival of the animal⁴. After detecting a potential threat, animals can perform a variety of protective behaviors. Depending on the perceived level of danger, animals might seek refuge, display freezing or alarm responses and when danger is extreme, execute fast escape behaviors^{5–7}. However, studies where behavioral correlates of multisensory integration can be directly tied to activity in an identified neuronal circuit are very rare and limited by the complexity and distribution of the neuronal networks involved.

Teleost fish can perform different evasive behaviors, the most explosive of which is known as a C-start fast escape. The C-start is initiated by a bilateral pair of reticulospinal neurons called Mauthner cells which determine the occurrence, latency and direction of the escape⁸. Mauthner cells have two main dendrites: disynaptic auditory input arriving from the inner ear contacts the lateral dendrite of the Mauthner cell^{9–11} while polysynaptic visual input coming from the optic tectum contacts the ventral dendrite^{12,13}. Both auditory and visual stimuli can activate the Mauthner cell with a probability and response latency which are a function of the stimulus

¹ Instituto de Fisiología, Biología Molecular y Neurociencias (IFIBYNE), CONICET-Universidad de Buenos Aires, Buenos Aires, Argentina. ²Departamento de Fisiología, Facultad de Ciencias Exactas y Naturales, Biología Molecular y Neurociencias, Ciudad Universitaria, Universidad de Buenos Aires, 1428 Buenos Aires, Argentina. ✉email: violetamedan@fbmc.fcen.uba.ar

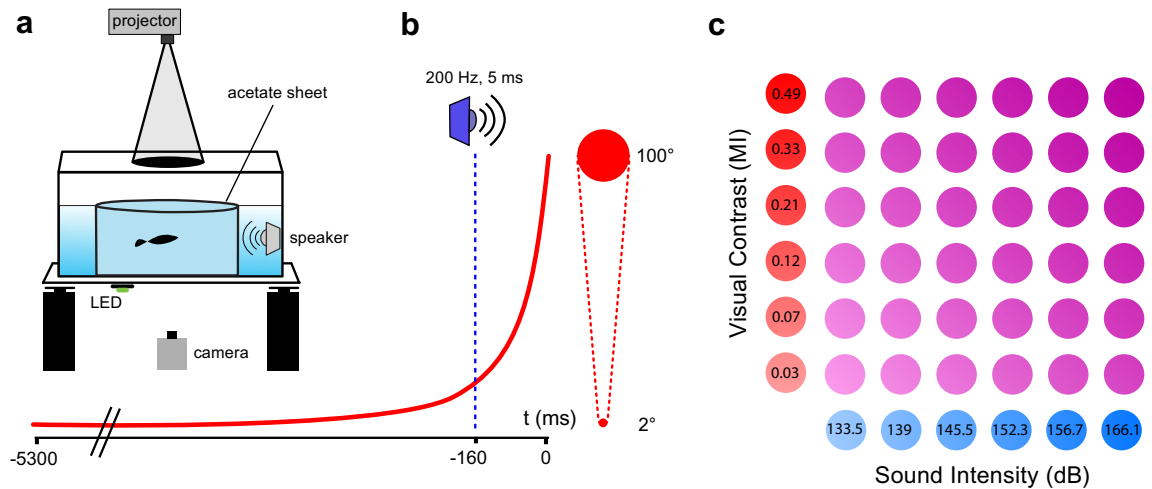


Figure 1. Experimental setup and stimulus design. **(a)** Behavioral setup. Computer-generated visual and auditory stimuli were delivered through a video projector and a water-proof speaker, respectively. A clear acetate sheet delimited the experimental arena. An LED was turned on synchronously with auditory stimulus onset while a high speed camera recorded fish behavior, visual stimuli and the state of the LED for the duration of each trial. **(b)** The visual loom expanded from a subtended angle of 2° to 100° in 5.3 s ($L/V=0.192$). When visual and auditory stimuli were combined, the auditory stimulus onset preceded the end of visual expansion by 160 ms. **(c)** Six intensities of visual stimuli (Michelson Index, MI) and 6 intensities of auditory stimuli (amplitude in dB, re 1 μ Pa) were combined into 36 different multisensory stimuli, thus creating a 2D space of multisensory intensities.

salience^{11,14,15}. Auditory salience depends on the amplitude of the sound wave while visual salience has been tied to stimulus dynamics and contrast^{14,16–19}. The temporal structure of auditory or visual stimuli that consistently trigger C-starts differs^{17,20}. Auditory C-starts are triggered by intense and brief auditory pips while visual C-start responses are typically evoked by fast expanding (looming) dark disks^{14,16,21}. The one-to-one relationship between a C-start and the firing of the Mauthner cell offers a unique opportunity to study how multisensory integration in a single neuron impacts on a fast escape behavior.

Here we demonstrate that, in goldfish, auditory and visual stimuli can be integrated to enhance C-start probability and reduce response latency. We also show that an integrate-and-fire computational model of the Mauthner cell reproduces fish responses to multisensory stimuli, making a direct link between behavioral data and its underlying neural mechanism.

Results

Risk assessment is affected by prior motor state. Fish reacted to visual looms and auditory pips (Fig. 1) with a variety of motor behaviors that were not limited to fast C-starts but included a range of slower escapes and alarm responses such as zigzagging, backward swimming, darting and freezing^{5,6,20,22}. As risk assessment and the resulting behavioral decision can be influenced by the animal's state *prior* to sensory stimulation, we analyzed motor behavior immediately before and after stimulation⁷.

Combined responses of 180 animals to auditory, visual or multisensory stimuli showed that behavior expressed after stimulation was dependent on the previous motor state (χ^2 Test of homogeneity, $p < 0.0001$). In about two thirds of all trials (63.3%) fish were either actively swimming or remained still but beating fins (non-freezing animals), while in the remaining trials fish were freezing before the stimulation began (36.7%, Fig. 2). Freezing had a strong impact on the behavioral outcome: in only 24% of trials in which animals were freezing a C-start was observed, while the rest remained freezing (in 1% of trials individuals performed an alarm response). In contrast, in trials from non-freezing animals, 54% percent performed a C-start while 17% responded with an alarm behavior. Conversely, almost 80% of all C-starts were performed by non-freezing animals. Therefore, although freezing does not abolish completely the possibility of performing a C-start, it sharply reduces its probability.

To further explore the influence of previous motor activity on C-start occurrence, we calculated swimming velocity during a 420 ms window, prior to auditory stimulation or 1.5 s before the end of visual expansion, since by this time no fish had responded (Supplementary Fig. S1a inset, grey area). Most animals swam at low velocities (< 20 mm/s) and swimming velocity distribution did not differ between animals that subsequently responded with a C-start or not (Supplementary Fig. S1a, binomial GLM, $p = 0.686$). This suggests that, although freezing reduces C-start probability, swimming does not enhance per se response probability. In fact, non-freezing animals with swimming velocities lower than 1 mm/s had a C-start response probability three times larger than those which were freezing (Supplementary Fig. S1b, 56% for non-freezing animals, 18% for freezing animals, binomial GLM, $p < 0.0001$). To test if visual contrast or auditory intensity affected the kinematics of the C-start, we calculated the maximum bend angle reached during the initial turn, as well as its duration, in a subset of trials. Neither contrast nor sound amplitude affected the maximum bend angle during the initial bend of the



Figure 2. Risk assessment is affected by prior motor state. Alluvial diagram of fish motor behavior in each trial before and after stimulation. The diagram shows the motor state before (Not Freezing, $n = 1003$, $N = 156$ or Freezing, $n = 580$, $N = 121$) and after sensory stimulation (C-start, $n = 675$, $N = 164$; Alarm, $n = 176$, $N = 98$; No change in motor behavior, $n = 732$, $N = 166$) combining all stimulus conditions. Each motor state is further divided by its “future” or “past” action (numbers within colored boxes). For example, 54% of non-freezing animals executed a C-start, while 60% of animals that did not changed their motor behavior were freezing before stimulation.

C-start (Supplementary Fig. S1c, d, Gaussian GLM, $p = 0.07$ [visual], $p = 0.3$ [auditory]) which showed a bimodal distribution peaking at $\pm 60^\circ$ (reflecting animals turning left or right). Duration of the initial bend lasted about 25 ms and was not affected by changes in stimulus intensity (Gaussian GLM, $p = 0.57$ [visual], $p = 0.63$ [auditory]).

Unisensory stimulus intensity modulates C-start response in non-freezing animals. To establish how varying unisensory stimulus salience affects risk assessment we evaluated the fish responses to increasing loom contrast or sound intensity (Fig. 3). In non-freezing animals, C-start response probability increased with stimulus intensity, covering a similar range in visual and auditory presentations (Fig. 3a–b, visual: from 10 to 71%, binomial GLM, $p < 0.0001$; auditory: 0% to 76%, binomial GLM, $p < 0.0001$). Unisensory stimuli produced alarm responses (Fig. 3c, d) although C-start probability was consistently higher than alarm probability, both for visual stimuli (red, Fig. 3a vs. c, Fisher’s Exact Test, $p < 0.0001$) and auditory stimuli (blue, Fig. 3b vs. d, Fisher’s Exact Test, $p = 0.0008$). Consistent with previous work from our lab¹⁹, alarm response probability was not dependent on stimulus intensity, neither for visual stimuli (Fig. 3c, binomial GLM for non-freezing trials, $p = 0.868$) nor for auditory stimuli (Fig. 3d, binomial GLM for non-freezing trials, $p = 0.141$; binomial GLM for freezing trials, $p = 0.164$).

Freezing animals had a lower C-start response probability across all visual or sound intensities (Fig. 3a red vs. light red, binomial GLM, $p < 0.0001$; Fig. 3b, blue vs. light blue, binomial GLM, $p < 0.0001$), did not perform alarm responses during visual stimulation (Fig. 3c) and only once to sound stimuli (Fig. 3d $n = 1/71$). Finally, while C-start probability is independent of visual contrast in freezing animals (Fig. 3a, binomial GLM, $p = 0.189$), there is a significant correlation with auditory intensity (Fig. 3b, binomial GLM, $p = 0.003$), although this correlation is only observed at larger amplitudes.

Stimulus intensity modulates response time. We next analyzed if C-start response time was affected by stimulus intensity. Our rationale was that as stimulus salience increases, the decision threshold should be reached faster. Responses to visual looms spanned within a wide time window, from 960 before to 100 ms after the end of the visual stimulation (Fig. 3e). However, as contrast grows there is a left-shift in the response time density distribution, i.e. responses occur earlier. Statistical analysis confirmed that higher contrast looms evoked C-starts with a shorter response time (gaussian GLM, $p = 0.026$). Response time was also influenced by freezing and by the position of the animal in the experimental arena at the time of stimulation (Supplementary Fig. S2b). Freezing delayed C-start onset (gaussian GLM, $p = 0.005$), and animals which experienced larger initial subtended angles (lighter grey points, Fig. 3e) exhibited shorter response times (gaussian GLM, $p = 0.002$).

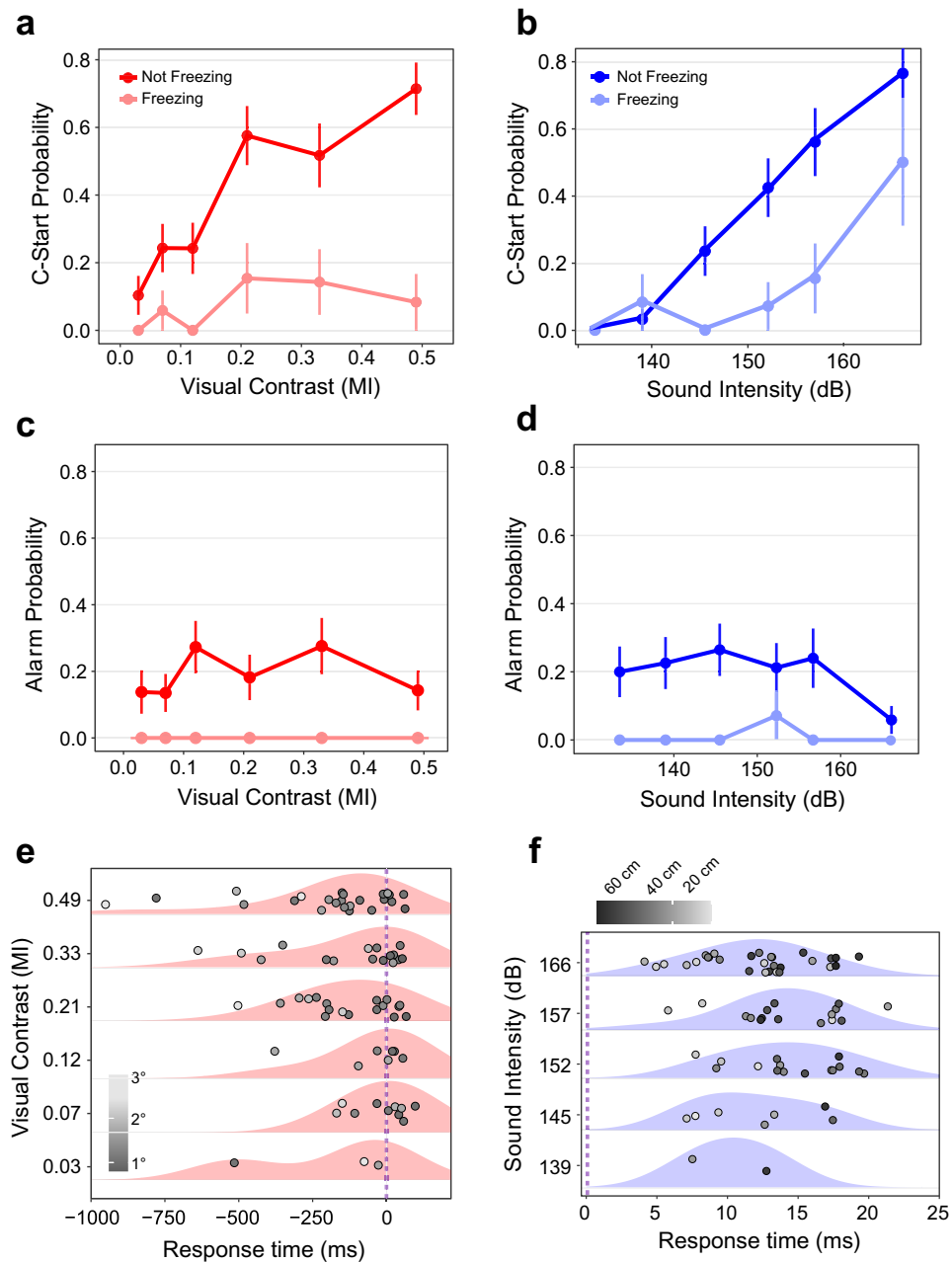


Figure 3. Unisensory stimulus intensity modulates C-start response in non-freezing animals. Unisensory C-start probability as a function of visual contrast (**a**, $N = 118$) or auditory intensity (**b**, $N = 108$) for non-freezing and freezing fish. Number of trials for each intensity vary between $n = 29$ – 37 for non-freezing and $n = 12$ – 19 for freezing trials. Alarm probability as a function of intensity for visual (**c**, $N = 118$) or auditory (**d**, $N = 108$) unisensory stimuli for non-freezing and freezing fish. Number of trials for each intensity varies between $n = 25$ – 34 for non-freezing and between $n = 8$ – 14 for freezing trials. In (**a**–**d**) bars represent standard error for the proportion. (**e**) Individual response times and density distributions for C-starts grouped by visual contrast ($n = 83$, $N = 53$). Dots are shaded according to the looming's subtended angle ($^{\circ}$) from the perspective of the animal, measured 1250 ms before the end of expansion. Temporal scale is zeroed at the end of the loom expansion (dashed line). (**f**) Individual response times and density distributions for C-starts grouped by sound intensity ($n = 70$, $N = 47$). The lowest intensity yielded no responses. Dots are shaded according to the animal's distance to the speaker (cm), 1 s before stimulus onset. Temporal scale is zeroed at the onset of auditory stimulation (dashed line). Shaded red or blue areas show density distributions of response times computed using a gaussian kernel. Vertical scatter was added for clarity in (**e**) and (**f**).

In contrast to the wide temporal distribution of visual responses, auditory C-starts occurred between 4 and 21 ms after sound presentation (Fig. 3f). The density distributions of auditory responses peaked between 10 and 15 ms, a value well within that reported in the literature²³ and were not shifted by sound intensity (gaussian GLM, $p=0.737$) nor were they affected by freezing (gaussian GLM, $p=0.122$). However, fish closer to the sound source responded earlier (gaussian GLM, $p<0.0001$, lighter grey points, Fig. 3f)²⁴. Animals located between 0 and 30 cm from the speaker had a mean response time of 10.64 ± 0.82 ms while those between 31 and 60 cm had a response time of 14.75 ± 0.67 ms (Two Sample *T*-Test, $p=0.0003$).

Multisensory integration enhances C-start probability. We also wanted to analyze unisensory vs. multisensory processing to assess how animals integrate multisensory information during threat detection. To this aim we designed sensory stimuli that mimic some aspects of a predator approaching from above (e.g. a bird), in direct collision course (looming stimulus). As the aerial predator finally breaks the surface of the water, a loud splash (auditory pulse) briefly precedes the end of the looming expansion (Fig. 1b). Besides being an ecologically relevant situation, this stimulus structure allowed us to specifically analyze the effect of a sudden addition of information (sound) after a period of gradual evidence accumulation (loom) in a potentially dangerous scenario. Combination of 6 levels of visual contrast and 6 levels of sound intensity resulted in 36 distinct multisensory stimuli, each with its unique combination of visual contrast and sound amplitude, where sound onset was kept at 160 ms before the end of visual expansion, unless otherwise specified (Fig. 1c). Figure 4 shows C-start probability for each combination for non-freezing (a) and freezing (b) animals. In line with the unisensory effects, it reveals that C-start probability increases with visual contrast and sound amplitude while freezing reduces response probability (binomial GLM, C-start probability \sim Amplitude + Contrast + Freezing, for each effect $p<0.0001$). For non-freezing animals, C-start probability covers a range between 0.07 and 1 (mean of 0.69, Fig. 4a), while in freezing animals, response probability ranges between 0 and 0.875 (mean of 0.26, Fig. 4b). The effects of contrast and amplitude over C-start probability were still significant for split data (binomial GLM non-freezing animals; amplitude: $p<0.0001$; contrast: $p<0.0001$; binomial GLM freezing animals; amplitude: $p<0.0001$; contrast: $p=0.003$). Alarm response probability for multisensory trials ranges between 0 and 0.39, decreases with sound amplitude but it is not affected by changes in contrast (Supplementary Fig. S3a).

Multisensory integration reduces response time. Timing is critical during a predatory evasion where a delayed escape can mean death. In this context, any additional process that speeds up threat assessment would become functionally advantageous⁴. We thus asked if C-start response time was reduced in multisensory vs. visual unisensory conditions, i.e. when a brief auditory stimulus is combined with ongoing visual stimulation. Figure 4c shows response times corresponding to the same 36 multisensory combinations grouped by contrast and colored by sound amplitude. We excluded responses which occurred before auditory stimulus onset, as these are unisensory visual responses (15% of the total).

In sharp contrast with unisensory experiments, where density distributions of response time showed a single peak (Fig. 3e, f), multisensory experiments show a more complex structure of response times with three distinct temporal windows. There is a high concentration of responses (between 82 and 100%, depending on the loom contrast) during the first 40 ms after auditory stimulus onset, which we consider to be true multisensory escapes and thus named this time period the Multisensory Integration interval (MSI). This is followed by a low response probability window (LP) of at least 45 ms where response probability drops to 0. Finally, a small proportion of the responses (10% of the total) occur on a ± 80 ms period centered on the end of the visual expansion (unisensory visual window, UV). Most of these late UV responses were produced in trials where the loom was combined with a low intensity sound. In contrast, multisensory stimuli that evoked responses during the MSI period had a mid to high intensity for the sound component (Supplementary Fig. S3b). These results clearly show that the addition of an extremely brief sound pip is enough to markedly shift the response distribution, effectively decreasing response time compared to visual only conditions^{3,25}. In addition, multisensory response latencies are much less variable than their unimodal visual counterparts: instead of spreading during several hundreds of milliseconds, most responses are locked to the occurrence of the auditory pip²⁶.

To explicitly test if the onset of the LP period was dependent on the moment of auditory stimulus presentation, we performed an additional experiment where the delay between the auditory pip and the end of the visual expansion was systematically varied (-40 , -60 , -160 , -260 , -360 and -460 ms relative to the end of visual expansion) while keeping loom contrast and sound intensity constant at an intermediate value (Michelson index of 0.16 and sound amplitude of 149 dB). Figure 4d shows the response times for each of these delays. Although in some trials fish escaped before the auditory stimulus onset (i.e. before the proper multisensory event starts, grey dots, Fig. 4d), most responses were triggered following the auditory stimulus (MSI period, cyan dots) while a small proportion of late responses were observed towards the end of the expansion (green dots). Interestingly, the duration of the LP period (grey area) was not fixed but dependent on the delay of the auditory pip, with longer delays producing longer LP periods.

Varying the delay between the auditory pulse and the end of the expansion does not change response probability (binomial GLM, $p=0.28$) nor the proportion of responses occurring during MSI (binomial GLM, $p=0.51$, cyan area of stacked bars, Fig. 4d). Overall, the results suggest that a relatively weak sound stimulus occurring up to half a second before the end of visual expansion enhances response probability and shifts responses to the moment of its presentation, effectively reducing C-start response delay compared to the visual-only condition.

Multisensory integration enhances response probability. Although C-starts were concentrated within the MSI, it could be that animals were either responding just to the loom or just to the sound, thus not necessarily implying integration. To infer actual multisensory integration, we compared response probabilities

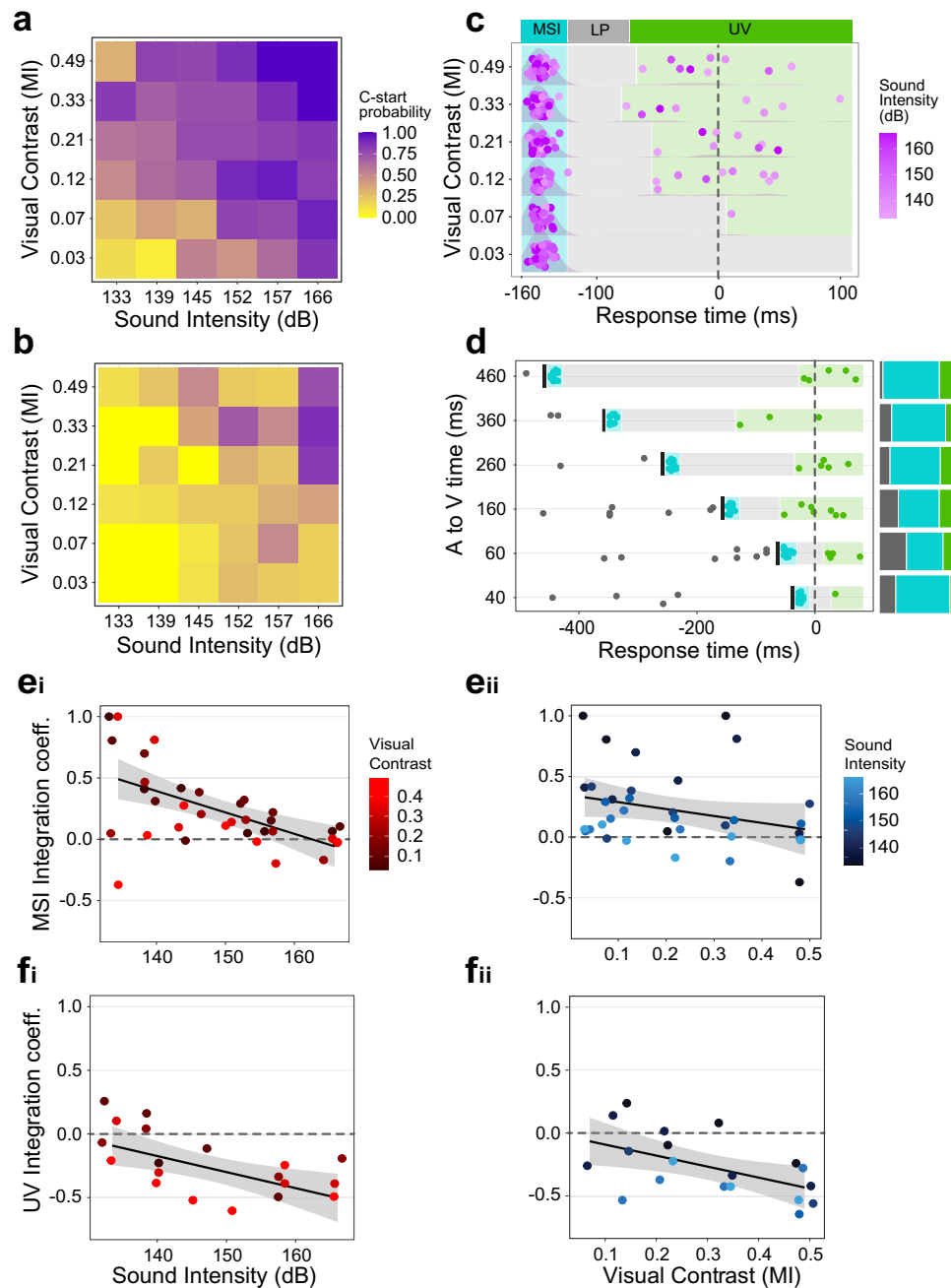


Figure 4. Multisensory integration enhances C-start probability and reduces response time. (a) C-start probability for non-freezing trials ($n=416$, $N=80$) and (b) freezing trials ($n=260$, $N=65$). (c) Response times and overlaid density distribution of C-starts during multisensory stimulation, grouped by visual contrast. Vertical scatter added for clarity. The purple shade of each dot represents the amplitude of the auditory stimulus presented 160 ms before the end of visual expansion (dashed line). The 250 ms period after the auditory stimulus shows three distinct phases: a high concentration of responses right after auditory presentation (MSI, cyan), an interval of low response probability (LP, light grey) and unisensory visual (UV, green) responses occurring around the end of visual expansion ($n=340$, $N=85$). (d) Multisensory response time was modulated by the delay between the auditory and visual stimuli. Sound intensity (149 dB) and visual contrast (MI 0.16) were invariant while the delay between auditory cue onset and end of loom was varied. Black bars signal the auditory stimulus. Response times follow the same pattern as in c, but the LP interval increases with the auditory-visual delay ($n=136$, $N=39$). Stacked bars show response probability before the auditory stimulus (dark grey), in the MSI (cyan) and UV (green) periods for each delay. (e) ICs during the MSI period as a function of sound intensity colored by visual contrast (e_i) or as function of visual contrast (e_{ii}) colored by auditory intensity (horizontal jitter is applied to ease visualization). ICs during the UV period as function of sound intensity and colored by visual contrast (f_i) or as function of visual contrast colored by auditory intensity (f_{ii}) (horizontal jitter applied to ease visualization). Coefficients were calculated only for combinations which yielded responses within this period (19 out of 36). In (e–f) the horizontal dashed line represents no integration and the solid line and shaded area represent a linear fit to the data and 95% confidence intervals.

in multisensory trials with those observed during unisensory trials. For each combination of stimuli, we calculated an Integration Coefficient (IC) for the MSI period based on the ORPs and the ERPs, assuming independent processing of visual and auditory inputs (see Methods). Figure 4e shows ICs plotted as a function of sound intensity (Fig. 4ei) or as function of visual contrast (Fig. 4eii) during the MSI window for each of the 36 multisensory combinations tested (or as a 2D matrix, Supplementary Fig. S3c). Most ICs (28/36) during the MSI period are positive, and their mean significantly differs from 0 (One sample *T*-test, $p = 0.0002$), implying that there is an overall multisensory enhancement of the C-start response, i.e. response probability exceeds what is predicted for independent processing of unisensory cues. ICs show a negative relationship both with sound intensity and visual contrast (Fig. 4e, sound: $p = 0.0001$, visual: $p = 0.03$) although the latter effect is weaker. This negative relationship implies that audiovisual multisensory integration shows inverse effectiveness with stimulus salience. In other words, multisensory enhancement is maximized when a weak auditory cue is combined with a low contrast loom (Supplementary Fig. S3c), but the enhancement vanishes as unisensory cues grow in salience^{25–27}.

We calculated the same coefficients for the UV period (Fig. 4fi, ii). ORPs during the UV period were in most cases lower than their respective ERPs, yielding negative coefficients (15/19 coefficients lower than 0), and their mean significantly differed from 0 (One sample *T*-test, $p = 0.0002$). Therefore, C-start probability during the UV period is less than what would be expected if no integration occurs, and even lower than expected for unimodal visual stimulation. We hypothesize that this reduction in the ICs during UV window is explained by the auditory stimulus shifting responses to the MSI period. If this were true, stronger sound stimuli should be more likely to shift responses to the MSI period, decreasing the number of responses observed in the LP and UV windows (Supplementary Fig. S3b). As goldfish do not perform multiple C-starts during a trial, if an animal C-starts during the MSI it won't do it again later during the LP and UV periods. Indeed, a gaussian GLM tested the dependence of ICs with respect to stimulus intensity ($IC \sim \text{Amplitude} + \text{Contrast}$) in the UV period. ICs show a negative relationship with sound amplitude (Fig. 4fi, $p = 0.008$) and visual contrast (Fig. 4fii, $p = 0.015$).

Overall, these results confirm that a brief auditory cue combined with a low contrast visual loom enhances escape responsiveness compared to unisensory stimulation and reduces response time compared to a visual-only condition. However, this enhancement is tightly restricted to a 40 ms window following the presentation of the auditory pip. When responses reappear after the LP period, ICs show negative integration. These results prompted the question of the mechanisms underlying the multisensory enhancement of the escape response and the auditory-evoked effect on visual processing to effectively decrease response time. Since rapid detection of a potential threat is critical to survival, our rationale was that all relevant information should be rapidly conveyed to the decision node for escape. We thus hypothesized that such network should integrate both visual and auditory stimuli with a very short latency and translate salience of integrated stimuli not into a graded response but into the probability of performing an all-or-none response such as a C-start. We therefore tested the hypothesis that a computational model representing a single element, the Mauthner cell, could accomplish this task.

An Integrate-And-Fire neuron model explains multisensory enhancement of C-start behavior.

In teleost fish, a variety of motor networks have been implicated in initiation and execution of evasive maneuvers^{17,28–30}. However, initiation of fast-start escape responses in response to auditory pips or fast looms is in most cases triggered by a single action potential of the Mauthner cell^{17,20,23,30,31}. The Mauthner cell receives excitatory visual input from the tectum and auditory input from the 8th nerve in addition to shunting feedforward inhibition^{32,33}. Both excitatory and inhibitory inputs are integrated at the soma and, if inputs are strong enough, an action potential initiates the C-start response. In search of a mechanism that might underlie the multisensory integration observed experimentally, we implemented a Leaky Integrate and Fire unit to produce a simplified model of the Mauthner cell^{34,35}. It's important to note that we did not attempt to produce a detailed model of the Mauthner cell circuit nor to explain the full complexity of the escape behavior, but to determine the minimal neural features that could reproduce our empirical data.

Brief square pulses or ramps of increasing current fed into the model neuron represented an approximation to the excitatory component of the auditory and visual inputs that would reach the Mauthner cell soma, respectively (Fig. 5a). The intensity of six “visual” and six “auditory” current inputs were adjusted to obtain a unisensory C-start probability for each stimulus that matched the response probabilities that were observed empirically (Fig. 5bi, ii; compare with Fig. 3a, b). As in the behavioral experiments, modelled C-start probability increases both with auditory intensity (binomial GLM, $p < 0.0001$) and visual intensity (binomial GLM, $p < 0.0001$). Next, we ran the simulation for each of the 36 combinations of multimodal stimuli introducing a fixed delay of 160 ms between the auditory input and the end of the visual input to obtain multisensory response probabilities (Fig. 5c). Increasing auditory or visual intensity produced a significant increase of response probability (binomial GLM, C-start prob. $\sim \text{Amplitude} + \text{Contrast}$, $p < 0.0001$ for both effects). Comparing each combination's probability with the corresponding probability from empirical data, we find that although the model produces slightly higher C-start probabilities for the least intense combinations (C-start probability ranges between 0.35 and 0.95), the overall difference is not significant, matching experimental results (Paired Samples *T*-test, $p = 0.47$, Fig. 5C).

ICs calculated within the MSI period are positive (One sample *T*-test, $p < 0.0001$) and decrease with auditory (Fig. 5di, binomial GLM, $p < 0.0001$) but not with visual intensity (Fig. 5dii, binomial GLM, $p < 0.357$), paralleling the inverse effectiveness observed for experimental data. Also, as in experimental data, ICs computed for the late window (i.e. those which occurred after the MSI) are negative (Fig. 5ei, ii One sample *T*-test, $p < 0.0001$). The agreement between modelled and experimental data reveals that inverse effectiveness can be explained by solely considering the summation of excitatory visual and auditory signals in the Mauthner Cell.

We further asked if the model reproduced the temporal distribution of responses observed in behavioral experiments. Specifically we asked if the introduction of the “auditory” square pulse during the ramped “visual” input shifted the moment the model cell crossed threshold (i.e. the cell fired). Figure 6a shows response times

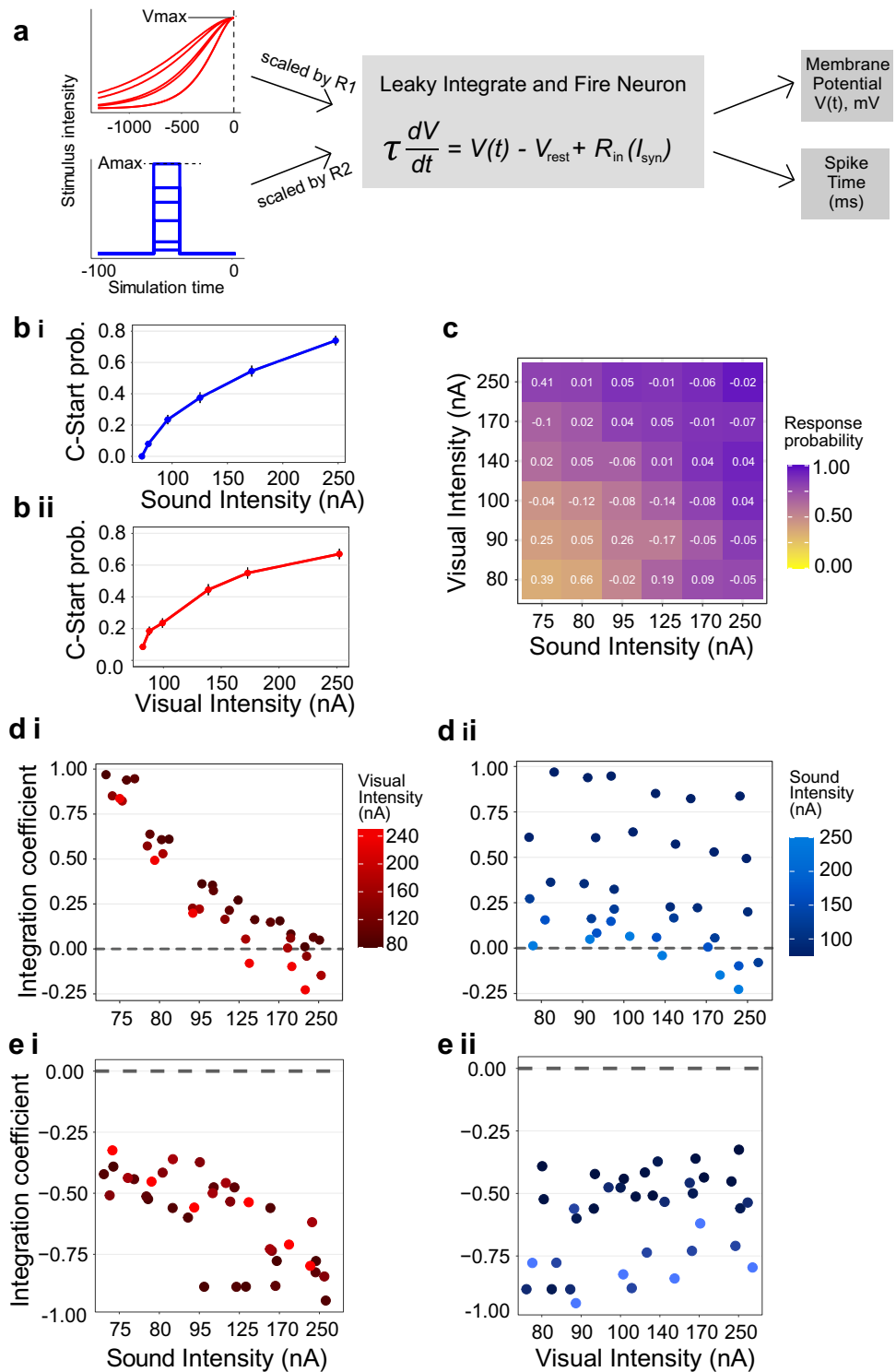


Figure 5. An integrate-and-fire neuron model explains multisensory enhancement of C-start behavior. (a) Model Diagram. Functions representing visual (upper left) or auditory (lower left) input converge onto the Mauthner cell model (central panel). Output variables measured are membrane potential and spike time when threshold is reached. (b) Response probability as a function of stimulus intensity for unisensory trials. Six auditory (b_i) or visual (b_{ii}) intensities were fitted so that response probability matched behavioral results for non-freezing animals (n = 1200). (c) C-start probability for multisensory trials. Numbers in white represent the relative difference with the experimental response probability, calculated as (p(mod) - p(exp))/(p(mod) + p(exp)) (cf., Fig. 4a). (d) ICs as a function of sound intensity and colored by visual intensity (d_i) or as function of visual intensity and colored by auditory intensity (d_{ii}). As in behavioral experiments, the model shows enhanced multisensory integration during the MSI period and inverse effectiveness for auditory intensity, although not for visual intensity. (e) Modelled ICs during the UV period as a function of sound intensity and colored by visual intensity (e_i) or as a function of visual intensity colored by auditory intensity (e_{ii}) (horizontal jitter applied to ease visualization). The dashed horizontal line represents no integration.

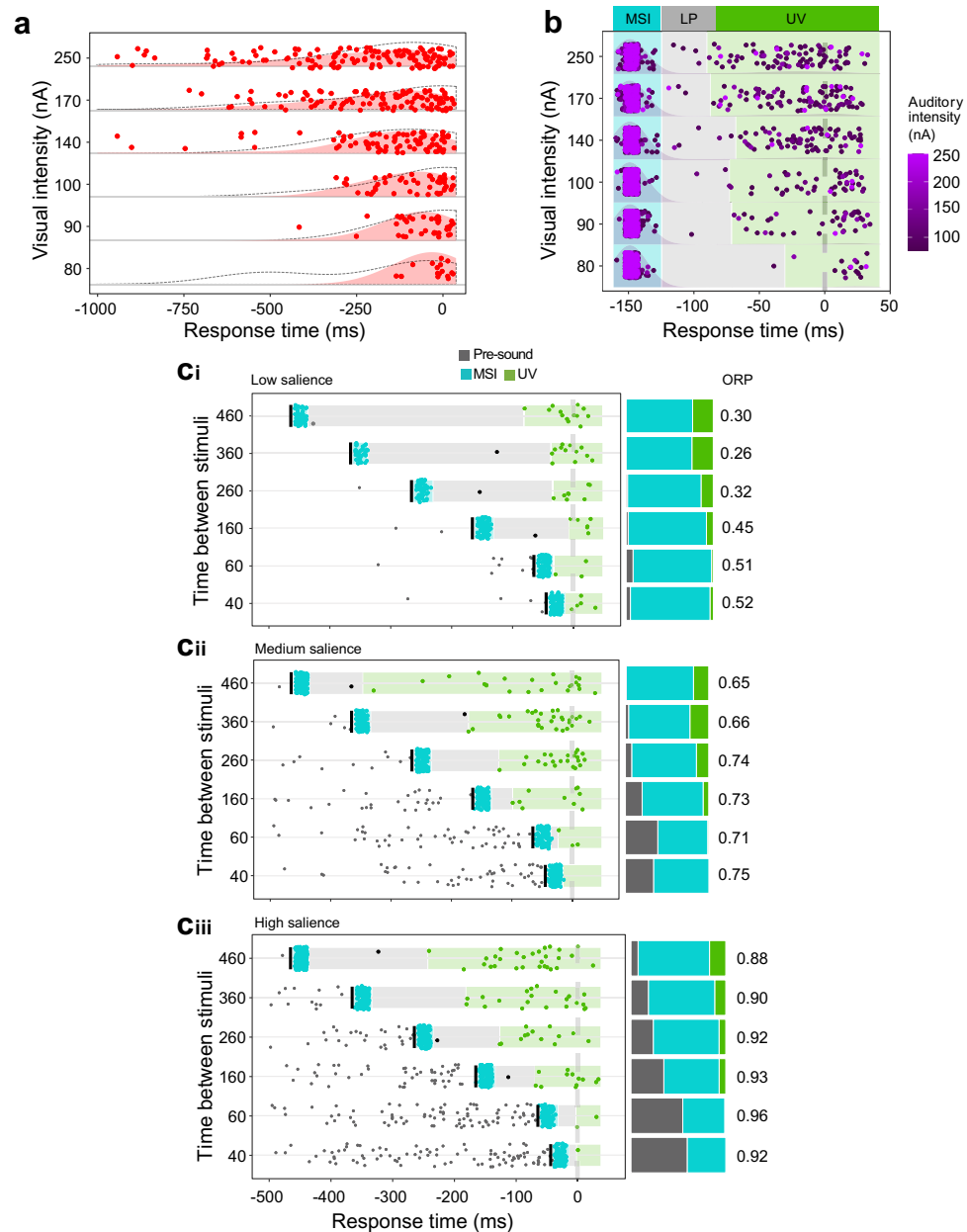


Figure 6. Model reproduces empirical response time distribution. (a) Response times and overlaid density distribution (light red) for visual unisensory trials, grouped by visual intensity. The grey dotted line is the experimental density distribution shown for comparison (cf. Figure 3e). (b) Multisensory response times and overlaid density distribution, grouped by visual intensity. The color of each dot represents auditory intensity. The vertical dotted line represents the end of visual stimulation. Responses are distributed in three distinct periods as in Fig. 4c–d ($n = 5089$). The boundary between LP and UV intervals is set by 5 and 95% of the responses. (c) Response times for multisensory trials with variable delay in ms between the auditory stimulus and the end of visual stimulus for low (c_i) medium (c_{ii}) or high (c_{iii}) saliency. Black vertical segments represent the time of auditory stimulus presentation ($n = 845$). Stacked bars to the right show response probability before the auditory stimulus (dark grey), MSI (cyan) and UV (green) for each delay. Numbers to the right indicate overall observed response probability for the simulation (ORP). Simulation was ran 100 times for each condition.

for the six intensities of unisensory visual stimuli used. The density distributions peak towards the end of the stimulus (at time 0) while higher intensities produce a higher proportion of early responses, i.e. the distribution tails to the left. Both effects, as well as the range of the model response times, matched those observed in the behavioral data (Fig. 3e).

As in our empirical data, simulated multisensory responses after auditory stimulus onset were divided into three distinct periods (Fig. 6b). A high concentration of responses (between 82 and 97% depending on the contrast) occurred during the MSI period. This was followed by a sharp decrease in response probability during

the LP period, which got progressively longer as visual intensity decreased. Finally, a relatively low number of events occurred during the UV period (compare with Fig. 4c). The model (run ten times more than behavioral trials) shows robustly that response probability in the LP and UV windows is not fixed but it increases for higher visual intensities, a trend that was present but subtle in the behavioral experiments. It also shows that most UV responses correspond to stimuli in which the sound component was weak.

We next explored the effect of systematically changing the delay between the auditory and visual inputs at three different salience levels. Figure 6ci–iii shows response times for combinations of auditory and visual stimuli of low, medium and high salience, where the elapsed time between auditory stimulus onset and the end of visual expansion was varied as in Fig. 4d. The simulation reproduced the empirical temporal structure of response times (compare Fig. 6cii with Fig. 4d). It also made evident that as visual salience increases, the proportion of early (visual unimodal) responses grows (grey stacked bars, Fig. 6ci–iii). While behaviorally we tested different delays of a single combination of intermediate sound intensity and visual contrast, modelling shows that comparable results should be expected when salience is lower (Fig. 6ci) or higher (Fig. 6ciii). Taken together, computational results strongly suggest that the multisensory enhancement of the C-start response and its inverse effectiveness features can be explained by only considering the interaction of excitatory visual and auditory signals in the Mauthner Cell. Moreover, the temporal structure of the C-start response distribution can also be explained by a minimal model that accumulates excitation until reaching threshold.

Discussion

The main questions posed in this paper are how goldfish integrate sensory information during risk assessment and whether this varies with the salience of the multisensory stimulus. We found that the addition of a brief sound pulse is capable of enhancing detection while speeding up the response to a visual threat. Multisensory enhancement disappears as unimodal salience increases to make single stimuli strong enough to bring the Mauthner cell to threshold. Providing putative mechanistic grounds for these observations, we found that behavioral results are reproduced by an Integrate and Fire model neuron. Noticeably, it was enough to combine excitatory input currents with dynamics matching the temporal structure of the empirical auditory and visual stimuli to reproduce the multisensory response enhancement and the shift in response time.

Inverse effectiveness of multisensory stimuli is strongly dependent on auditory intensity. Hallmark of multisensory integration, inverse effectiveness is traditionally explained in terms of the ambiguity of the situation^{1,2}. When any component of a multisensory percept is very salient, additional information contributed by secondary components will make only a modest contribution to the response. However, when unisensory stimulus strength is lowered, the value of combining information from independent sources grows. Although we matched the range of saliences of our visual and auditory stimuli (using response probability as a proxy for salience), increasing sound intensity produced a steeper decrease in multisensory enhancement than increasing contrast (i.e. a stronger inverse effectiveness, compare fit slope in Fig. 4ei vs. e_{ii}). From a functional perspective, one could argue that a blunt noise is a much less ambiguous signal of immediate danger than a gradual size increase of a dark spot. More mechanistically, considering the neural substrate of the C-start, a simplified model of the Mauthner cell circuit reproduces the asymmetry on the inverse effectiveness of visual and auditory salience (Fig. 5di vs. d_{ii}). This model consists of a single compartment that integrates excitatory currents (Fig. 5a), and thus all presynaptic circuit effects (such as inhibition), as well as dendritic filtering that may vary between auditory and visual processing, are omitted¹³. The only difference remaining between the “visual” and “auditory” components is their temporal structure. It is therefore not necessary to postulate complex differences between auditory and visual processing (even though they may exist) to explain their impact in response probability, multisensory integration and inverse effectiveness: in this case, it suffices to account for distinct temporal dynamics. This underscores that the temporal structure of a sensory stimulus adds critical information to its meaning.

Reduction in response time to multisensory stimuli is determined by sound onset. The auditory stimulus not only had a relatively stronger inverse effectiveness but also definitive influence on response time. The temporal distribution of unisensory visual responses peaks towards the end of the loom but has a tail of earlier responses which gets increasingly skewed as contrast increases (Figs. 3e and 6a). When a sound pip is added even as early as 460 ms before the end of the expansion, most responses concentrate in a narrow MSI time window following the pip. This effect is observed irrespective of loom contrast, sound intensity (Figs. 4c and 6b) or the interval between the sound and the end of the expansion (Figs. 4d and 6c). However, the MSI period is not just a response to an acoustic stimulus. The width of the response time distribution for multisensory trials doubles the duration of the auditory unisensory responses (40 vs. 20 ms, Figs. 3f vs. 4c). Functionally, adding a brief auditory pip to a low contrast looming expansion increases response probability and reduces reaction time compared to a visual only condition. In a real-life predatory encounter, the importance of binding these two pieces of ambiguous information becomes evident⁴.

Following the MSI period, we consistently observed a time window where no responses were recorded (LP) and where multisensory integration becomes negative. Auditory-evoked feedforward inhibition of the Mauthner cell could be implicated in the lack of responses during the LP period, but it cannot by itself fully explain the characteristics of the LP period. Feedforward inhibition triggered by acoustic stimuli decays after 50–60 ms^{13,15} and thus could only explain the shortest LP periods. In addition, although we cannot rule out that inhibition is delaying the UV responses, our computational model was able to reproduce the pattern of an MSI period followed by an LP window without having specified such a mechanism. Indeed, we explicitly tested the effect of adding feedforward inhibition to the model by introducing a negative current mirroring the excitatory inputs

delayed by 7 ms. The results were qualitatively identical with the only effect of increasing the intensity of the stimuli necessary to bring the cell to threshold. These observations could be empirically tested by decreasing (or increasing) inhibition in the circuit and assessing the shift in response thresholds, using pharmacological agents or recording the Mauthner cell with a chloride-filled electrode^{36–38}. We also modelled the effect of freezing as an inhibitory component in the computational model, although surely freezing operates at multiple levels in the escape circuit. We did so by adding a constant inhibitory input to the Mauthner cell and obtained that the temporal distribution of responses was unaffected while sensory thresholds increased. Therefore, the model suggests that both feedforward inhibition and freezing might increase sensory thresholds (Fig. 3) while leaving latency response distribution unchanged³⁶.

To illustrate the mechanism we propose to explain how multisensory computation is implemented in the Mauthner cell, Fig. 7 shows two scenarios where visual looms and auditory pips depolarize the cell. Sensory input will depolarize the Mauthner cell to trigger a C-start if threshold is crossed. The strength of the excitation resulting from a fixed external input (high, left or low, right) will differ across animals as a result of internal state, attention, position with respect to the stimulus, intrinsic properties of the Mauthner cell, etc. Therefore, a fixed stimulus will only trigger a response in a proportion of trials and response time will differ across animals (i.e. responses are not deterministic but probabilistic). Although individual response variability will affect both sensory modalities, the distribution of response times will be much broader for visual looms than for brief auditory pips, reflecting their substantial temporal differences (Fig. 7, upper vs. middle panels). Inverse effectiveness predicts that combination of high intensity visual and auditory inputs should not produce substantial MSI (Fig. 7, left panels). By adding the two components, threshold would be reached virtually 100% of the times (Fig. 7, lower left panel). This might appear to be a substantial improvement from a 70 or 80% chance of evoking a C-start for the unisensory stimuli. However, it is crucial to note that multisensory response probabilities should not be compared to unisensory probabilities alone, but to their ERP, which in this example is 94% (i.e. one would expect a response 94% of the times if stimuli were presented together but processed independently). In light of this analysis, 100% does not represent a great enhancement due to MSI, a fact captured by an IC of 0.03. Notwithstanding, all responses would occur when the auditory stimulus is presented, thus producing a left-shift in response time compared to the visual only condition.

In contrast, if two low stimuli produce individually a modest depolarization of the Mauthner cell, it would be expected that their combination produce a response enhancement (Fig. 7, right panels). Assuming weak unisensory components that only trigger the Mauthner cell in 10% of the population, their combined ERP would be 19% (Fig. 7, lower right panel). However, we have shown that MSI enhances response probability when stimuli have a low salience. In this example, an ORP of around 50% would more than double the ERP (19%, resulting in an IC of 0.45, see Fig. 4ei). Note that this is due to summed inputs being much more capable of reaching the threshold than the inputs themselves, which is not true for stronger unisensory stimuli, already reaching the threshold most of the time. Again, a substantial reduction in response time can be observed, compared to the visual only condition. Therefore, this simple conceptual framework that only assumes temporal integration of two different excitatory inputs with specific temporal profiles is enough to explain our empirical results. Although additional neural circuitry surely plays a role in multisensory integration and escape execution^{39–41}, our findings suggest that excitatory inputs integrated at the Mauthner cell soma are key elements in multisensory decision making during fast C-start escapes.

Conclusions

That multisensory integration in single cells translates to actual behavioral advantage is a presumption that has received little empirical evidence. Here we show that goldfish integrate weak multisensory cues to enhance threat detection and reduce escape latency. In a real life predatory encounter, the functional role of this multisensory process is evident. Moreover, we show that a very simple neural implementation of this process in the Mauthner cell can reproduce behavioral observations. This sheds light on the computational basis of multisensory decision mechanisms likely operating in many other organisms.

Methods

Animals. One hundred and ninety adult goldfish (*Carassius auratus*) of both sexes and 7–10 cm of standard body length were used in this study. Fish were purchased from local aquaria (Daniel Corralelo, Buenos Aires, Argentina) and allowed to acclimate for at least a week after transport. Fish were kept in rectangular glass holding tanks (30×60×30 cm; 95 l) in groups of 10. Tanks were supplied with filtered and dechlorinated water and maintained at 18 °C. Ambient light was set to a 12-h light/dark photoperiod. Animals were fed floating pellets (Sera, Germany) five times a week. All procedures and protocols were performed in accordance with the guidelines and regulations of the Institutional Animal Care and Use Committee of Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and approved by the same committee (protocol #70). Reporting in the manuscript conforms to the ARRIVE guidelines pertinent to this study⁴².

Experimental setup. Goldfish were tested in a rectangular experimental tank (80 cm length, 70 cm width, and 15 cm height) standing on an anti-vibration table with its external walls covered with black opaque cardboard to avoid external visual stimulation. A cylindrical enclosure (60 cm diameter, 13 cm height) made of clear acetate was placed inside the tank to confine fish to the experimental arena (Fig. 1a). The tank was filled with filtered dechlorinated water up to a height of 12 cm. Auditory stimuli were produced by an underwater loudspeaker (UW-30, University Sound, Buchanan, MI) placed outside the experimental arena and supported within a 6 cm thick layer of foam. Visual stimuli were produced by a video projector (Epson Powerlite S6+, 60 Hz) secured 130 cm above the tank. The tank was covered by a lid made of white wax paper sheet, which served as a

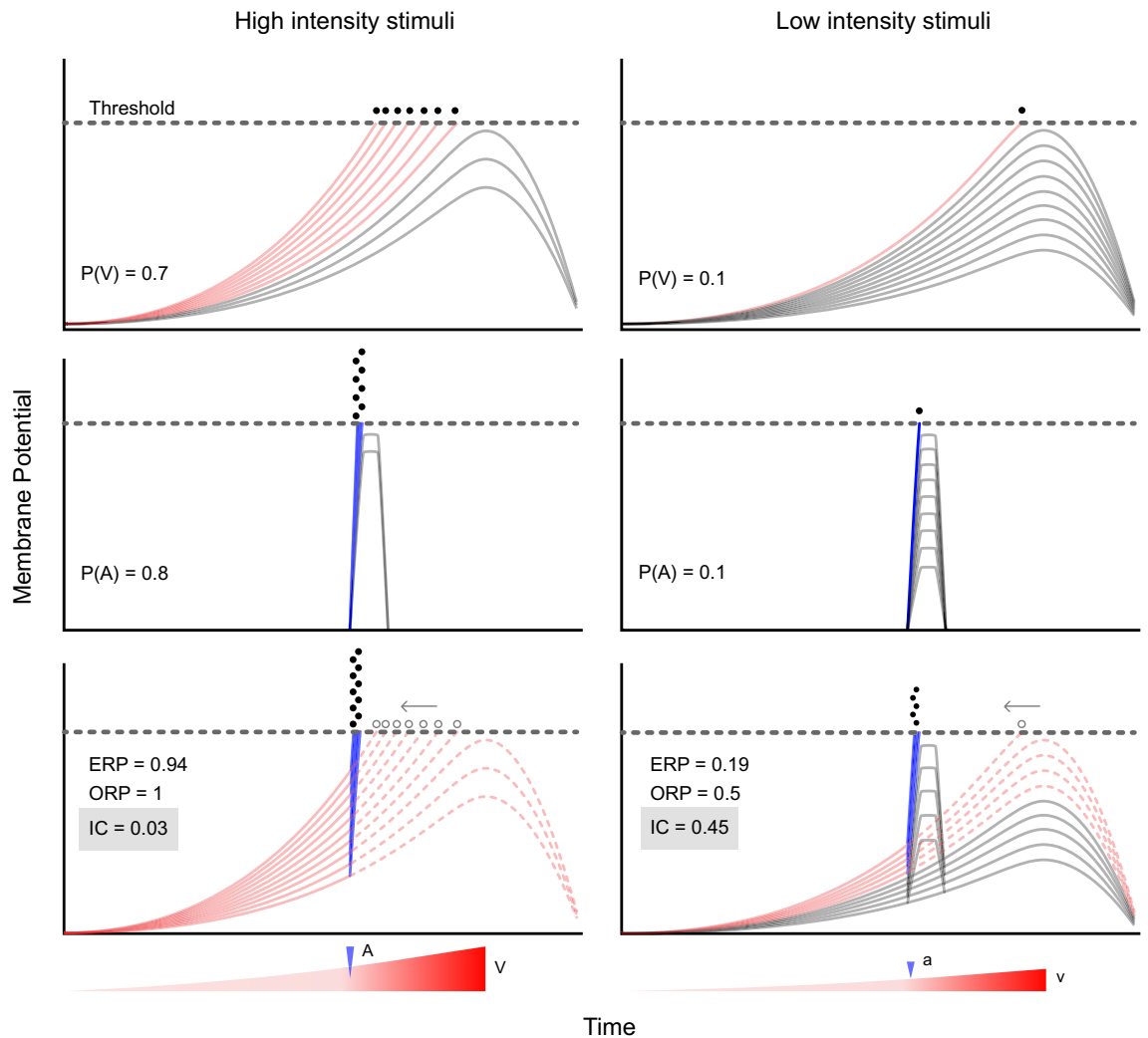


Figure 7. Conceptual model of multisensory integration in the Mauthner cell. Mauthner cells showed a ramped depolarization as a consequence of the increasing visual expansion (red triangles, bottom) or brief auditory pips (blue arrowhead, bottom). Differences in cell excitability or other factors will produce different rates of depolarization in different cells (different curves in the same panel). In the example, 7 out of 10 cells reached threshold with a strong loom (upper left) while only 1 out of 10 of those cells reached threshold with a low intensity loom (upper right). A similar scenario can be observed when comparing strong and weak auditory stimuli (left and right middle panels), yielding 0.8 and 0.1 response probability, respectively. The effect of combining two strong or two weak stimuli is exemplified in the bottom panels. A strong multisensory stimulus will depolarize all cells until threshold is reached (locked to the presentation of the auditory stimulus, black dots) but the relative enhancement of response probability (0.94–1) would be minimal (6%, IC = 0.03). In sharp contrast, two weak stimuli having a unisensory response probability of 0.1 might be combined to produce a summed depolarization that drives the Mauthner cell to threshold on 50% of the population (raising the IC = 0.45). The time point where each cell would have crossed threshold in the visual only condition (empty gray dots, lower panels) shifts to the moment when the auditory stimulus occurs (black dots, lower panels). Curves indicate Mauthner cell depolarization in response to a visual loom colored red if it reaches threshold or grey if it does not. Each plot indicates either the visual ($P(V)$) or auditory ($P(A)$) unimodal response probability or the Expected (ERP) and Observed (ORP) response probability and the resulting Integration Coefficient (IC).

projection screen for the visual stimuli. The bottom of the tank was transparent, allowing for simultaneous video recording of fish behavior and stimulus presentation at 240 fps (432 × 320 pixels, Casio EX ZR100, Tokyo, Japan). A green LED located below the experimental tank (hidden from animal's view) was used to allow the camera to record the time of auditory stimulus onset. Experiments were made in a silent room with ceiling lights off. Presentation of visual and auditory stimuli as well as camera acquisition and LED activation was controlled by MATLAB (R2016a, www.mathworks.com). Triggering of video acquisition occurred simultaneously with visual stimulus onset and stopped 8.25 s after the end of visual stimulation.

Stimuli. Computer-generated dark disks that expand over a lighter background (loom) efficiently elicit C-start escapes^{14,16,21}. In our study, looming stimuli expanded exponentially from a diameter of 0.27 cm until reaching a

diameter of 18.15 cm 5.3 s later. The subtended angle, as measured from the center of the tank, grew from 2.06° to 100.84° (Fig. 1b). Looming stimulus can also be defined by the size of the approaching object of equal width and height (L), the approach velocity (V), and the apparent distance covered. The combined term L/V is an indicator of the expansion of the subtended angle with time and is used to allow comparison to looming-evoked responses in other animals. The looming used here yield an $L/V = 0.192$. We chose a fast expanding loom as it has been shown that low L/V values preferentially recruit fast escapes triggered by the Mauthner cell whereas looms of higher L/V values might recruit networks driving responses of longer latency^{17,20,30}.

To obtain visual stimuli of different salience, we modified the grayscale value of the disk stimulus (8-bit values between 70 and 118) while keeping the background constant (8-bit value of 120), obtaining different contrasts^{19,43}. To characterize the luminance of each component of our stimuli, we used the irradiance sensor (J1812) of a Tektronix J17 photometer (Wilsonville, Oregon, MI, USA) positioned in the center of the tank while projecting images on the wax paper lid. During these measurements, all pixels in the screen were set to the grayscale 8-bit value that we were currently testing. Using these irradiance values, we determined the contrast for each stimulus, calculated as the Michelson index (MI), where contrast is defined as $(I_{\text{disk}} - I_{\text{background}})/(I_{\text{disk}} + I_{\text{background}})$, and I_{disk} and $I_{\text{background}}$ refer to the irradiance of the expanding disk and the background, respectively. We used six different visual stimuli, characterized by Michelson indices of 0.03, 0.07, 0.12, 0.21, 0.33 and 0.49. Auditory stimuli consisted of a single cycle of a 200 Hz sine wave (5 ms in duration), whose amplitude was changed to modify its salience. We used six different auditory stimuli of amplitudes of 133.5, 139, 145.5, 152.3, 156.7 and 166.1 dB re. 1 μPa when recorded with a hydrophone (Sensor SQ34) placed 10 cm away from the speaker. Intensity of auditory and visual stimuli was selected to obtain a range of unimodal C-start response probabilities spanning from about 0–80% (see below).

We analyzed 12 unisensory stimuli (6 visual, 6 auditory) as well as all 36 multisensory combinations of stimuli (Fig. 1c). When visual and auditory stimuli were combined (multisensory presentations), auditory component onset occurred 160 ms before the end of visual expansion, unless otherwise specified (Fig. 1b).

Stimulation protocol. Individual fish were placed in the experimental tank and allowed to acclimate for 15 min. The animals then received between 9 and 13 stimuli that included at least one unisensory visual (V), one unisensory auditory (A) and, when performing a multisensory protocol, 7–9 multisensory (V + A) stimuli. From the 36 possible multisensory stimuli each animal was tested with a subset that included low, medium and high intensity combinations. Stimulus presentation was randomized and inter-trial interval was set to 4 min to minimize the effect of habituation. However, a slight decrease in response probability was observed, both for C-start responses (21% of decrease by trial 13, Binomial GLM, $p = 0.004$) and alarm responses (4.6% of decrease by trial 13, Binomial GLM, $p = 0.002$, Supplementary Fig. S1a).

We also found a strong thigmotactic behavior, with 95% of the fish remaining within 15 cm from the enclosure walls (distances to the center of the arena were not consistent with a uniform distribution: Chi-Square Test of Homogeneity, $p < 0.0001$; mean distance to center \pm standard deviation were of 29.4 ± 4.8 cm, Supplementary Fig. S2b, c). As visual looms were projected in the center of the arena, fish position affected the stimulus' subtended angle perceived at the retina, which could in turn affect C-start response probability. Indeed, C-start probability was higher for animals closer to the center of the arena (and to the expanding loom) and decreased as the animal's position approached the enclosure walls (Supplementary Fig. S2d). However, this dependence could have been related to the position with respect to the center of the arena itself and not to the fish's relative position to the center of the loom expansion. To test that this was not the case, the visual expansion was centered on one extreme of the enclosure, and thus did not coincide with the center of the tank. When presented this way, C-start probability was uncorrelated with distance to the center of the arena (binomial GLM, $p = 0.911$). Throughout all trials, C-start probability was (negatively) correlated with the distance between the fish and the center of the looming expansion (binomial GLM, $p < 0.0001$, Supplementary Fig. S2d). As for auditory stimuli, distance to the speaker was not correlated with response probability (binomial GLM, $p = 0.719$, Supplementary Fig. S2e).

C-Start escape and alarm responses. Videos were analyzed offline using VirtualDub (1.10.4, www.virtualdub.org) and custom-made code built upon the OpenCV library in Python (3.4.10, www.opencv.org). For trials which included a visual stimulus, the last frame of expansion of the looming was set as time point 0 ms. Therefore, C-start responses occurring before the end of visual expansion have a negative response time, whereas those occurring after rendered a positive response time (Fig. 1b). For auditory only trials, response time was considered from the start of the auditory stimulus, which in the videos was recorded by the activation of an LED not visible to the fish. Animals which never performed a C-start response were excluded from analysis ($N = 10/190$). Videos were also inspected and scored by two independent observers to analyze the occurrence of C-start escapes or other behaviors suggesting increased arousal or alarm. Alarm responses consist on a variety of subtle but robust motor reactions including accelerating or decelerating swimming, darting (a single fast acceleration in one direction with the use of the caudal fin), erratic movements/zigzagging (representing fast acceleration bouts in rapid succession), rapid abduction of fins with no body displacement and freezing (which consists of a complete cessation of movement, except for gills and eyes)^{5,6,44}. An alarm response was recorded when scoring of occurrence and description of behavior matched for both observers.

Data analysis. In order to analyze the effect of multisensory integration for different treatments, we first asked what would be the expected response probability for a combination of visual and auditory stimuli if there was no integration (i.e. if visually-evoked responses and auditory-evoked responses were processed as independent phenomena, as opposed to the effect of a visual stimulus depending on the co-occurrence of an auditory stimulus and vice versa). Given a response probability for a visual stimulus $P(V)$, and a response probability for

an auditory stimulus $P(A)$, using the Addition Rule for the Probabilities of Independent Events, the Expected Response Probability (ERP) in the absence of integration upstream of escape decision-making can be calculated as $P(V \text{ or } A) = P(V) + P(A) - P(V) \times P(A)$. We compared the ERP with the Observed Response Probability (ORP) for each multisensory combination of stimuli V and A . The relative difference between ORP and ERP is a measure of the effect of integration over response probability, since it is the difference between what we observed and what is expected of a system which does not integrate the sensory signals. We define an Integration Coefficient (IC) as $\frac{ORP-ERP}{ORP+ERP}$. This value ranges from -1 to 1 , it is 0 if integration has no effect over response probability, it is positive if integration increases response probability and it is negative if integration decreases response probability. We calculated the IC for each of the 36 multisensory combinations of stimuli.

Statistical analysis. R (version 4.0.2, www.r-project.org) and RStudio (version 1.1.456, www.rstudio.com) were used for statistical analysis. A significance level of $\alpha = 0.05$ was used throughout the study. Effects of explanatory variables over response variables were assessed using Generalized Linear Models (binomial GLMs in the case of binary response variables, gaussian GLMs in the case of continuous response variables). Sample size is denoted by N when it refers to the number of animals or n when it refers to the number of trials. When working with probability data, error bars represent standard error of a proportion.

Computational model. We used NEST (2.20.0, www.nest-simulator.org)⁴⁵ running on Python (3.8.3) to create a computational model of the Mauthner Cell (Fig. 5a) using a Leaky Integrate-And-Fire neuron model with exponential post synaptic currents (*iaf_psc_exp* on the documentation). This model follows a differential equation that can be expressed as $\tau \frac{dV}{dt} = V(t) - V_{rest} + R_{in} I_{syn}$, where τ is the time constant, $V(t)$ is the membrane potential, V_{rest} is the resting potential, R_{in} is the input resistance and I_{syn} is the sum of the input currents. The parameters of this model were adjusted to match reported intrinsic properties of the Mauthner cell. Resting membrane potential was set to -80.0 mV and so was the reset value for the potential after a spike. Firing threshold was set to -65.0 mV, C to 2500 pF/cm² and the time constant to 0.5 ms^{13,46}.

The Mauthner cell receives excitatory mechanosensory inputs through its lateral dendrite and excitatory visual inputs through its ventral dendrite which are integrated at the soma. We modelled these two types of inputs while retaining stimulus characteristics from our experiments. Auditory stimuli were designed as 20 ms square pulses of current, as an approximation of low-pass filtered somatic recordings of Mauthner Cells during similar auditory stimulation^{11,13,24}. We designed 6 auditory stimuli with intensities that ranged between 75 and 250 nA. The intensities were adjusted to reproduce the observed unisensory auditory probability. Similarly, 6 visual stimuli were designed as ramped currents^{14,47} which followed the equation $v(t) = \frac{e v_{max}(s-t)}{s e^{(s-t)/s}}$, where e is Euler's constant, t is time, v_{max} is the maximum value of the "visual" input current and s is a parameter which controls the slope of the curve. v_{max} was varied to modify stimulus intensity, and ranged between 90 and 220 nA to match experimental C-start probabilities of unisensory visual trials. The value of s was picked randomly for each simulated trial from a Gamma distribution with a mean of 200 and a standard deviation of 150. This generates a family of curves which both resemble the temporal dynamic of our looming stimuli and of previous recordings of goldfish Mauthner Cells during looming stimulation¹⁴. To add variability to the model, v_{max} and a_{max} (maximum value of "auditory" input) were multiplied by a number (R_1 and R_2 , respectively) picked from a uniform random distribution in the interval $(0; 1]$ before the beginning of each trial. Thus, reported stimulus intensities for auditory or visual stimuli represent the maximum possible input current ($R_i = 1$). The neuron received auditory and visual inputs through two distinct excitatory synapses with weights equal to 1.

For each combination of v_{max} and a_{max} the simulation was run 200 times. Although experimental trials lasted 5.3 s, behavioral responses were only observed in the last 1000 ms of visual stimulation, therefore computational rounds were only simulated for 1300 ms. The end of visual exponential increase was set to 300 ms before the end of the trial, and the auditory square pulse was presented 160 ms before the end of the visual input, unless otherwise specified.

We calculated response probability for a given stimulus as the proportion of trials where threshold was reached. Response time was recorded as the time elapsed between the threshold was crossed and the end of visual stimulation (or to auditory stimulus onset in unisensory auditory trials).

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Received: 10 August 2021; Accepted: 3 January 2022

Published online: 20 January 2022

References

- Stein, B. E. & Stanford, T. R. Multisensory integration: Current issues from the perspective of the single neuron. *Nat. Rev. Neurosci.* **9**(4), 255–266 (2008).
- van Atteveldt, N., Murray, M. M., Thut, G. & Schroeder, C. E. Multisensory integration: Flexible use of general operations. *Neuron* **81**(6), 1240–1253 (2014).
- Rowland, B. A., Quessy, S., Stanford, T. R. & Stein, B. E. Multisensory integration shortens physiological response latencies. *J. Neurosci.* **27**(22), 5879–5884 (2007).
- Nair, A., Nguyen, C. & McHenry, M. J. A faster escape does not enhance survival in zebrafish larvae. *Proc. Biol. Sci.* **284**(1852), 20170359 (2017).
- Smith, R. J. F. Alarm signals in fishes. *Rev. Fish Biol. Fish.* **2**(1), 33–63 (1992).

6. Kalueff, A. V. *et al.* Towards a comprehensive catalog of zebrafish behavior 1.0 and beyond. *Zebrafish* **10**(1), 70–86 (2013).
7. Evans, D. A., Stempel, A. V., Vale, R. & Branco, T. Cognitive control of escape behaviour. *Trends Cogn. Sci.* **23**(4), 334–348 (2019).
8. Korn, H. & Faber, D. S. The Mauthner cell half a century later: A neurobiological model for decision-making?. *Neuron* **47**(1), 13–28 (2005).
9. Zottoli, S. J. & Faber, D. S. Properties and distribution of anterior VIIIth nerve excitatory inputs to the goldfish Mauthner cell. *Brain Res.* **174**(2), 319–323 (1979).
10. Pereda, A. E., Bell, T. D. & Faber, D. S. Retrograde synaptic communication via gap junctions coupling auditory afferents to the Mauthner cell. *J. Neurosci.* **15**(9), 5943–5955 (1995).
11. Szabo, T. M., Weiss, S. A., Faber, D. S. & Preuss, T. Representation of auditory signals in the M-cell: Role of electrical synapses. *J. Neurophysiol.* **95**(4), 2617–2629 (2006).
12. Zottoli, S. J., Hordes, A. R. & Faber, D. S. Localization of optic tectal input to the ventral dendrite of the goldfish Mauthner cell. *Brain Res.* **401**(1), 113–121 (1987).
13. Medan, V., Mäki-Marttunen, T., Sztarker, J. & Preuss, T. Differential processing in modality-specific Mauthner cell dendrites. *J. Physiol.* **596**, 667 (2017).
14. Preuss, T., Osei-Bonsu, P. E., Weiss, S. A., Wang, C. & Faber, D. S. Neural representation of object approach in a decision-making motor circuit. *J. Neurosci.* **26**(13), 3454–3464 (2006).
15. Neumeister, H., Szabo, T. M. & Preuss, T. Behavioral and physiological characterization of sensorimotor gating in the goldfish startle response. *J. Neurophysiol.* **99**(3), 1493–1502 (2008).
16. Dunn, T. W. *et al.* Neural circuits underlying visually evoked escapes in larval zebrafish. *Neuron* **89**(3), 613–628 (2016).
17. Bhattacharyya, K., McLean, D. L. & MacIver, M. A. Visual threat assessment and Reticulospinal encoding of calibrated responses in larval Zebrafish. *Curr. Biol.* **27**(18), 2751–2762.e6 (2017).
18. Heap, L. A. L., Vanwalleghem, G., Thompson, A. W., Favre-Bulle, I. A. & Scott, E. K. Luminance changes drive directional startle through a thalamic pathway. *Neuron* **99**(2), 293–301.e4 (2018).
19. Otero Coronel, S., Martorell, N., Beron de Astrada, M. & Medan, V. Stimulus contrast information modulates sensorimotor decision making in goldfish. *Front. Neural Circuits.* **14**, 23 (2020).
20. Burgess, H. A. & Granato, M. Sensorimotor gating in larval zebrafish. *J. Neurosci.* **27**(18), 4984–4994 (2007).
21. Temizer, I., Donovan, J. C., Baier, H. & Semmelhack, J. L. A visual pathway for looming-evoked escape in larval Zebrafish. *Curr. Biol.* **25**(14), 1823–1834 (2015).
22. Pfeiffer, W. The fright reaction of fish. *Biol. Rev. Camb. Philos. Soc.* **37**, 495–511 (1962).
23. Eaton, R. C., Lee, R. K. & Foreman, M. B. The Mauthner cell and other identified neurons of the brainstem escape network of fish. *Prog. Neurobiol.* **63**(4), 467–485 (2001).
24. Casagrand, J. L., Guzik, A. L. & Eaton, R. C. Mauthner and reticulospinal responses to the onset of acoustic pressure and acceleration stimuli. *J. Neurophysiol.* **82**(3), 1422–1437 (1999).
25. Stein, B. E., Stanford, T. R. & Rowland, B. A. Development of multisensory integration from the perspective of the individual neuron. *Nat. Rev. Neurosci.* **15**(8), 520–535 (2014).
26. McIntyre, C. & Preuss, T. Influence of stimulus intensity on multimodal integration in the startle escape system of goldfish. *Front. Neural Circuits.* **13**, 7 (2019).
27. Kadunce, D. C., Vaughan, J. W., Wallace, M. T. & Stein, B. E. The influence of visual and auditory receptive field organization on multisensory integration in the superior colliculus. *Exp. Brain Res.* **139**(3), 303–310 (2001).
28. Eaton, R. C., Lavender, W. A. & Wieland, C. M. Alternative neural pathways initiate fast-start responses following lesions of the mauthner neuron in goldfish. *J. Comp. Physiol.* **145**(4), 485–496 (1982).
29. Eaton, R. C., Nissanov, J. & Wieland, C. M. Differential activation of Mauthner and non-Mauthner startle circuits in the zebrafish: Implications for functional substitution. *J. Comp. Physiol.* **155**(6), 813–820 (1984).
30. Marquart, G. D., Tabor, K. M., Bergeron, S. A., Briggman, K. L. & Burgess, H. A. Preoptine non-giant neurons drive flexible escape behavior in zebrafish. *PLoS Biol.* **17**(10), e3000480 (2019).
31. Weiss, S. A., Zottoli, S. J., Do, S. C., Faber, D. S. & Preuss, T. Correlation of C-start behaviors with neural activity recorded from the hindbrain in free-swimming goldfish (*Carassius auratus*). *J. Exp. Biol.* **209**(Pt 23), 4788–4801 (2006).
32. Diamond, J., Roper, S. & Yasargil, G. M. The membrane effects, and sensitivity to strychnine, of neural inhibition of the Mauthner cell, and its inhibition by glycine and GABA. *J. Physiol.* **232**(1), 87–111 (1973).
33. Faber, D. S., Korn, H. & Lin, J. W. Role of medullary networks and postsynaptic membrane properties in regulating Mauthner cell responsiveness to sensory excitation. *Brain Behav. Evol.* **37**(5), 286–297 (1991).
34. Izhikevich, E. M. *Dynamical Systems in Neuroscience* 522 (MIT Press, 2007).
35. Bahl, A. & Engert, F. Neural circuits for evidence accumulation and decision making in larval zebrafish. *Nat. Neurosci.* **23**(1), 94–102 (2020).
36. Weiss, S. A., Preuss, T. & Faber, D. S. Phase encoding in the Mauthner system: Implications in left-right sound source discrimination. *J. Neurosci.* **29**(11), 3431–3441 (2009).
37. Mirjany, M. & Faber, D. S. Characteristics of the anterior lateral line nerve input to the Mauthner cell. *J. Exp. Biol.* **214**(Pt 20), 3368–3377 (2011).
38. Curtin, P. C. P. & Preuss, T. Glycine and GABAA receptors mediate tonic and phasic inhibitory processes that contribute to prepulse inhibition in the goldfish startle network. *Front. Neural Circuits.* **9**, 12 (2015).
39. Koyama, M., Kinkhabwala, A., Satou, C., Higashijima, S. & Fetcho, J. Mapping a sensory-motor network onto a structural and functional ground plan in the hindbrain. *Proc. Natl. Acad. Sci. USA* **108**(3), 1170–1175 (2011).
40. Fetcho, J. R. & McLean, D. L. Some principles of organization of spinal neurons underlying locomotion in zebrafish and their implications. *Ann. N Y Acad. Sci.* **1198**, 94–104 (2010).
41. Severi, K. E., Böhm, U. L. & Wyart, C. Investigation of hindbrain activity during active locomotion reveals inhibitory neurons involved in sensorimotor processing. *Sci. Rep.* **8**(1), 13615 (2018).
42. du Sert, N. P. *et al.* The ARRIVE guidelines 20: Updated guidelines for reporting animal research. *PLoS Biol.* **18**(7), e3000410 (2020).
43. Evans, D. A. *et al.* A synaptic threshold mechanism for computing escape decisions. *Nature* **558**(7711), 590–594 (2018).
44. Brown, G. E., Godin, J. G. & Pedersen, J. Fin-flicking behaviour: A visual antipredator alarm signal in a characin fish. *Hemigrammus Erythrozonus. Anim. Behav.* **58**(3), 469–475 (1999).
45. Fardet, Tanguy, Vennemo, Stine Brekke, *et al.* NEST 2.20.0. Zenodo. (2020).
46. Funch, P. G. & Faber, D. S. Cable properties of goldfish Mauthner axon. *J. Neurophysiol.* **47**(6), 1196–1213 (1982).
47. Medan, V. & Preuss, T. The Mauthner-cell circuit of fish as a model system for startle plasticity. *J. Physiol. Paris.* **108**(2–3), 129–140 (2014).

Acknowledgements

We thank Dr. Lidia Szczupak, Dr. Martín Carbó Tano, Dr. Martín Berón de Astrada and Santiago Otero Coronel for discussion and for critically reading earlier versions of this manuscript. We also thank Ángel Vidal for invaluable technical assistance.

Author contributions

V.M. and N.M.: Conception of the study, experimental design and data analysis. N.M. Performed empirical and computational experiments. V.M. and N.M. Contributed to drafting and revising the manuscript and both authors approved the final version of the manuscript.

Funding

N.M. was funded by a doctoral fellowship from Universidad de Buenos Aires. This work was supported by ANPCyT (PICT 2016-0007) and Universidad de Buenos Aires (UBACyT 20020170200213BA), and CONICET (PIP2014 GI 11220130100729CO01) (V.M.).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-04998-2>.

Correspondence and requests for materials should be addressed to V.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022