

# Whole-brain neural substrates of behavioral variability in the larval zebrafish

Jason Manley, Alipasha Vaziri 

Laboratory of Neurotechnology and Biophysics, The Rockefeller University, New York, United States • The Kavli Neural Systems Institute, The Rockefeller University, New York, United States

 [https://en.wikipedia.org/wiki/Open\\_access](https://en.wikipedia.org/wiki/Open_access)

 Copyright information

Reviewed Preprint

v2 • May 12, 2025

Revised by authors

Reviewed Preprint

v1 • May 30, 2024

## eLife Assessment

Manley and Vaziri introduce an **important** new method for brain-wide imaging of cellular activity in zebrafish and provide evidence for the applicability of this technique. They use this method to explore the question of how neural variability gives rise to variability in behavior. The analyses used are mostly **convincing**, although questions regarding spatial and temporal imaging resolution and their effects on the study's interpretations and conclusions suggest only partial support for some of the central results.

<https://doi.org/10.7554/eLife.97014.2.sa3>

## Abstract

Animals engaged in naturalistic behavior can exhibit a large degree of behavioral variability even under sensory invariant conditions. Such behavioral variability can include not only variations of the same behavior, but also variability across qualitatively different behaviors driven by divergent cognitive states, such as fight-or-flight decisions. However, the neural circuit mechanisms that generate such divergent behaviors across trials are not well understood. To investigate this question, here we studied the visual-evoked responses of larval zebrafish to moving objects of various sizes, which we found exhibited highly variable and divergent responses across repetitions of the same stimulus. Given that the neuronal circuits underlying such behaviors span sensory, motor, and other brain areas, we built a novel Fourier light field microscope which enables high-resolution, whole-brain imaging of larval zebrafish during behavior. This enabled us to screen for neural loci which exhibited activity patterns correlated with behavioral variability. We found that despite the highly variable activity of single neurons, visual stimuli were robustly encoded at the population level, and the visualencoding dimensions of neural activity did not explain behavioral variability. This robustness despite apparent single neuron variability was due to the multi-dimensional geometry of the neuronal population dynamics: almost all neural dimensions that were variable across individual trials, i.e. the “noise” modes, were nearly orthogonal to those encoding for sensory information. Investigating this neuronal variability further, we identified two sparsely-distributed, brain-wide neuronal populations whose pre-motor activity predicted whether the larva would respond to a stimulus and, if so, which direction it

would turn on a single-trial level. These populations predicted single-trial behavior seconds before stimulus onset, indicating they encoded time-varying internal modulating behavior, perhaps organizing behavior over longer timescales or enabling flexible behavior routines dependent on the animal's internal state. Our results provide the first whole-brain confirmation that sensory, motor, and internal variables are encoded in a highly mixed fashion throughout the brain and demonstrate that de-mixing each of these components at the neuronal population level is critical to understanding the mechanisms underlying the brain's remarkable flexibility and robustness.

## Introduction

Animals are not deterministic input-output machines, but instead display highly flexible behavioral responses even under sensory invariant conditions. A holistic understanding of the neuronal mechanisms underlying decision making requires explaining such variability in behavior at the single trial level. However, there are many putative sources of trial-to-trial neuronal variability. In particular, it is unclear whether such variability results from deterministic sources, such as internal states and long-timescale dynamics, or stochasticity. Individual neurons exhibit non-negligible noise from a variety of sources, including electrical noise due to ion channel fluctuations, synaptic noise due to biochemical processes, jitter in the timing of action potentials, and “digitization” noise due to the all-or-nothing nature of the action potential (Faisal et al., 2008 [↗](#)).

However, trial-to-trial variability in neuronal firing is generally not independent across neurons, but instead highly correlated over repeated stimulus presentations within certain groups of neurons (Kohn et al., 2015 [↗](#); Shadlen and Newsome, 1998 [↗](#); Zohary et al., 1994 [↗](#)), suggesting that variability in neural circuits may not be dominated by the stochasticity of individual neurons. Such structured covariation is referred to as “noise correlations” (Cohen and Kohn, 2011 [↗](#)) because they are unrelated to the often experimenter-controlled external variable that is simultaneously encoded within the neuronal population. While certain structures of noise correlations are known to limit the information encoding capacity of the neuronal population (Moreno-Bote et al., 2014 [↗](#); Bartolo et al., 2020 [↗](#); Romyantsev et al., 2020 [↗](#)), there is little consensus about both their origin (Kanitscheider et al., 2015 [↗](#); Zylberberg et al., 2016 [↗](#)) and behavioral impact (Zohary et al., 1994 [↗](#); Huang and Lisberger, 2009 [↗](#)). In particular, there are opposing reports in different species and under different task conditions as to whether such inter-neuronal correlations can improve (Cafaro and Rieke, 2010 [↗](#); Zylberberg et al., 2016 [↗](#); Valente et al., 2021 [↗](#)) or interfere (Cohen and Maunsell, 2009 [↗](#); Ruda et al., 2020 [↗](#); Kafashan et al., 2021 [↗](#)) with the neural computations underlying decision making. As such, a general framework capable of reconciling such widely varying reports and explaining the functional impact of noise correlations on neural computation and behavior is currently missing.

Further, few experimental studies have linked noise correlations in sensory regions to pre-motor dynamics and ultimately behavior (Valente et al., 2021 [↗](#)). Instead, most reports have mainly focused on the scale and structure of such covariations across sensory neurons in anesthetized animals or without regard to the animal's behavioral state. However, recent studies have shown that many species encode the animal's own movements and decisions in a highly fashion, including across early sensory regions (Kauvar et al., 2020 [↗](#); Musall et al., 2019 [↗](#); Stringer et al., 2019 [↗](#); Steinmetz et al., 2019 [↗](#)). These brain-wide, motor-related fluctuations must then drive at least some of the observed neuronal covariation across stimulus presentations. Thus, noise correlations are not merely stochastic “noise”, but instead may encode important variables that are stimulus-independent and perhaps related to the animal's proprioception and context, such as long-timescale behavioral and internal state dynamics.

Such highly overlapping representations of sensory information and additional contextual variables may underly the brain's remarkable ability to flexibly control behavior and select the most appropriate action at any given moment based on the animal's goals and internal state. As such, naturalistic decision making is a function of not just incoming sensory stimuli, but also internal variables such as the animal's motivations and recent experience or learned behaviors. For example, foraging has been shown across species to be modulated by hunger (Filosa et al., 2016 [DOI](#)), long-timescale exploration versus exploitation states (Flavell et al., 2013 [DOI](#); Marques et al., 2020 [DOI](#)), and other hidden internal variables (Sternson, 2020 [DOI](#); Torigoe et al., 2021 [DOI](#); Flavell et al., 2022 [DOI](#)). This evidence suggests that answering longstanding questions regarding the neuronal mechanisms underlying naturalistic decision making will require understanding the intersection of numerous neural circuits distributed throughout the vertebrate brain.

Thus, probing the structure of the neuronal mechanisms underlying behavioral variability requires high spatiotemporal resolution recording of brain-wide neuronal dynamics and behavior across many trials. While it is often necessary to average across trials to deal with both the inherent variable firing of single neurons and experimental measurement noise, such trial averaging inherently precludes any investigation of the link between neural and behavioral variability on the trial-by-trial level. Instead, methods that utilize the entire neuronal population dynamics have shown that pooling information across neurons can allow for successful decoding of information that is not possible from individual neurons. This is because simultaneous multi-neuron recordings can leverage statistical power across neurons to capture the most salient aspects of population dynamics within single trials. Such population-level approaches have revolutionized the fields of motor planning and decision making, providing highly robust brain-machine interfaces (Kao et al., 2015 [DOI](#); Pandarinath et al., 2018 [DOI](#)) that combine information from at least a few recording sites and making it possible to predict decisions (Wei et al., 2019 [DOI](#); Lin et al., 2020 [DOI](#)), reaction time (Afshar et al., 2011 [DOI](#)), and additional behavioral states (Kaufman et al., 2015 [DOI](#); Marques et al., 2020 [DOI](#)) at the single trial level. Recently, optical imaging techniques (Weisenburger and Vaziri, 2018 [DOI](#); Kim and Schnitzer, 2022 [DOI](#); Machado et al., 2022 [DOI](#)) have dramatically increased the attainable number of simultaneously recorded neurons within increasingly larger brain volumes at high speed and resolution. Thus, large-scale neural recording provides a promising approach to both screening for potential sources of variability across the brain and identifying robust population-level dynamics at the level of single trials. However, due to the difficulty of obtaining whole-brain optical access in many vertebrates, previous studies have not been able to quantify brain-wide noise correlations at the single neuron level and link such noise modes to variable decisions at the single trial level.

In this study, we sought to identify the neural loci which drive or encode information related to behavioral variability across trials. To do so, we performed whole-brain recording at cellular resolution in larval zebrafish engaged in ethologically relevant visually-evoked behaviors to identify neuronal populations that were predictive of behavioral variability under sensory invariant conditions. While it is well-known that larval zebrafish on average exhibit a target-directed prey capture response to small visual stimuli and a target-avoidance escape response to large visual stimuli, we demonstrated that these ethological, oppositely valenced, and highly divergent responses exhibited significant variability across trials of identical stimulus presentations. Turning first to visually-evoked neural populations, we found that despite trial-to-trial variability in the responses of single neurons, visual information was reliably encoded at the population level across trials of each stimulus. As such, the visual-encoding neuronal activity patterns did not account for the larvae's variable behavioral responses across trials.

A key feature of our system was that it allowed investigating the geometry of trial-to-trial variability of neuronal dynamics on the whole-brain-level, by simultaneous observation of brain-wide neural variability during a visual decision-making behavior. Thus, we were able to identify whole-brain modes of neural variability which were related to the larva's motor actions. In particular, we identified highly distributed neuronal populations whose pre-motor activity

predicted the larvae's responsiveness and turn direction on a single trial level, indicating that behavioral variability is not represented in particular anatomical neural loci. We found that this predictability exhibited two dominant timescales: a longer-timescale and pre-stimulus turn bias that was not time-locked to the stimulus presentations or motor timing; and a rapid increase in predictability and ramping activity about one to two seconds before movement initiation. Consistent with result from previous studies, we speculate that the former, longer-timescale bias is related to a circuitry that has evolved to optimize an exploration versus exploitation behavior in foraging while the latter, short-timescale ramping activity likely drives the downstream motor circuitry to execute the actual selected action in each trial. Our data suggest that behavioral variability in response to repetitions of the same sensory stimulus may not be driven by a single brain region. Rather, it is more likely generated by a combination of factors encoded by neurons throughout the brain, including a time-varying and internal turn direction bias, in addition to the well-studied visuomotor transformations.

## Results

### Fourier Light Field Microscopy enables high-speed imaging of neuronal population dynamics

In order to investigate the whole-brain neural correlates of behavioral variability, we designed and built a Fourier light field microscope (fLFM) that was optimized for high-resolution, whole-brain recoding of neuronal activity in the larval zebrafish. Conventional light field microscopy (cLFM) provides an elegant solution to volumetric snapshot imaging by placing a microlens array (MLA) in the image plane to encode both the position and angle of incidence of the light rays onto a camera sensor (Levoy et al., 2006). Combined with 3D deconvolution algorithms for computational reconstruction of the recorded volume (Agard, 1984; Broxton et al., 2013; Prevedel et al., 2014), LFM enables imaging of an entire volume without requiring any scanning. Due to its high speed and volumetric capabilities, cLFM has been widely applied to image the dynamics of *in vivo* biological systems, particularly neuronal dynamics (Prevedel et al., 2014; Aimon et al., 2019; Lin et al., 2020; Quicke et al., 2020; Yoon et al., 2020). Further, additional computational strategies have been devised to image deep into scattering tissue, such as the mouse brain (Nöbauer et al., 2017; Skocek et al., 2018; Nöbauer et al., 2023). However, a primary limitation of cLFM is a dramatic drop in resolution near the native image plane (NIP) because its spatio-angular sampling is uneven across depth and highly redundant near the NIP.

This low resolution at the center of the sample can only be partially overcome using techniques such as wavefront coding (Cohen et al., 2014). In contrast, fLFM has been shown to provide high-resolution imaging across a two-to three-fold extended depth by processing the light field information through the Fourier domain (Llavador et al., 2016; Scrofanì et al., 2017; Cong et al., 2017; Guo et al., 2019; Liu et al., 2020); i.e., placing the MLA conjugate to the back focal plane of the objective. In this configuration, the MLA segments the wavefront by transmitting the various spatial frequencies (which correspond to angular information) into images on different regions of the camera, providing a well-aliased sampling of the spatio-angular information without significant redundancy near the NIP.

We simultaneously optimized our design for simplicity and cost-effectiveness by using exclusively off-the-shelf elements, as opposed to previous implementations of fLFM which relied on custom optical components. Our design (Figure 1A, see Methods for details) enables imaging of an approximately  $750 \times 750 \times 200 \mu\text{m}^3$  volume, corresponding to roughly the whole brain of a 7 days post fertilization (dpf) larval zebrafish. 3D information is captured in the raw images (Figure 1B) because each lenslet in the MLA segments a region of the Fourier plane and thus forms an image of the sample from a unique perspective. Experimentally, the point spread function (PSF) full width at half maximum (FWHM) of our system measured  $3.3 \pm 0.21 \mu\text{m}$  (95% confidence

interval) laterally and  $5.4 \pm 0.4 \mu\text{m}$  axially (**Figure 1C**), consistent with theoretical estimates of  $3.1 \mu\text{m}$  and  $4.7 \mu\text{m}$ , respectively. This represents a substantial improvement in resolution compared to our previous cLFM, which had a PSF FWHM of  $3.4 \mu\text{m}$  laterally and  $11.3 \mu\text{m}$  axially (Prevedel et al., 2014). Additionally, the fLFM has a much more consistent resolution as a function of depth as compared to the cLFM (**Figure S1A**), which exhibits a dramatic drop in resolution near the NIP. As such, our design enabled cellular-resolution imaging across the whole brain of the larval zebrafish.

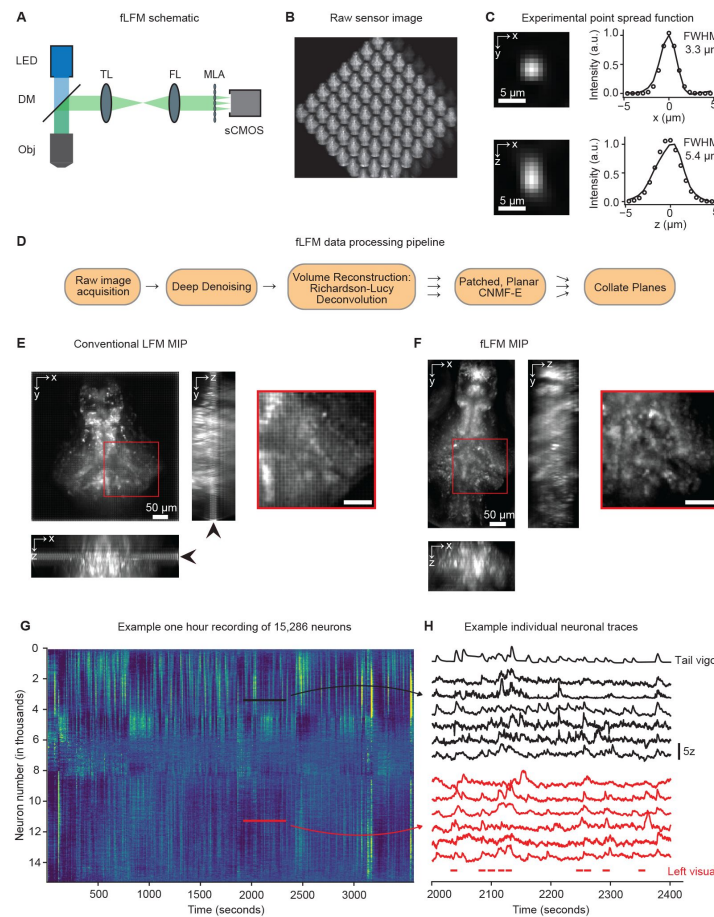
Next, we developed a custom pipeline (**Figure 1D**) for 3D reconstruction and neuronal segmentation to extract the dynamics of neurons across the field of view. This consisted of three main steps: denoising, 3D reconstruction, and neuronal segmentation. As one of the key limitations of LFM is signal-to-noise ratio (SNR) due to the distribution of emitted photons onto a sensor array, we first denoised the raw sensor images. To do so we trained a DeepInterpolation (Lecoq et al., 2021) deep neural network model, which has been shown to increase the SNR of the resulting neuronal timeseries by removing shot noise that is independent across consecutive frames (**Figure S1B**, see Methods for details). Next, the full 3D volume was reconstructed from denoised sensor images using Richardson-Lucy deconvolution (Agard, 1984; Broxton et al., 2013; Prevedel et al., 2014) and an experimentally measured PSF. In addition to enabling high-resolution imaging across an extended depth of field, a key advantage of fLFM is that the reconstruction of the volume images can be performed approximately 100 times faster than in cLFM, due to fLFM's shift-invariant point spread function. Finally, to identify neuronal regions of interest (ROIs) and their dynamics within the reconstructed volume, we utilized CNMF-E (Zhou et al., 2018), a constrained matrix factorization approach to extract *in vivo* calcium signals from one-photon imaging data. The CNMF-E algorithm was applied to each plane in parallel, after which the neuronal ROIs from each plane were collated and duplicate neurons across planes were merged (**Figure S1C**).

Finally, the effect of photobleaching was corrected according to a biexponential fit to the mean activity across neurons (**Figure S1D**).

To validate our setup, we imaged whole-brain dynamics from head-immobilized larval zebrafish expressing nuclear-localized GCaMP6s (NL-GCaMP6s) pan-neuronally (Vladimirov et al., 2014) while monitoring the larva's tail movements using a high-speed camera and presenting visual stimuli, which as we will discuss below consisted of single dots of 3-5 different sizes moving from the center toward the left or right visual field. Across our 1-2 hour recordings, we identified  $16,524 \pm 3,942$  neurons per animal (mean  $\pm$  95% confidence interval; range: 6,425 to 35,928) which showed a 3D distribution throughout the brain that was consistent with a Z-Brain atlas reference image (Randlett et al., 2015) labeled with nuclear-localized red fluorescent protein (**Figure S1E**). Thus, the combination of fLFM and the improvement in our reconstruction pipeline enabled the detection of significantly higher neuron numbers than the  $\sim 5,000$  neurons as reported in our previous cLFM realizations (Prevedel et al., 2014; Lin et al., 2020). The performance improvement over cLFM could also be visualized by inspecting the maximum intensity projection (MIP) of a zebrafish recording in each modality. While cLFM (**Figure 1E**) exhibits lower resolution and grid-like artifacts near the native image plane (black arrows) due to redundant sampling near the NIP, fLFM's well-aliased sampling of spatio-angular information provides cellular-resolution imaging throughout the entire volume (**Figure 1F**).

Within these observed large-scale dynamics, we found a diversity of neuronal activity patterns (**Figure 1G**). For example, we identified neurons tuned to the vigor the larva's tail movements (see Methods for details) and those tuned to the presentation of visual stimuli as described below (**Figure 1H**), providing a proof of principle that fLFM enables whole-brain imaging of the relevant neuronal population activity encoding sensory inputs and behavior.





**Figure 1**

## Fourier Light Field Microscopy (fLFM) provides a simple and cost-effective method for whole-brain imaging of larval zebrafish during behavior

**A.** Schematic of the fLFM system. The sample is illuminated with a 470 nm LED through a 20×/1.0-NA imaging objective (Obj). The fluorescence is sent to the imaging path via a dichroic mirror (DM). An Olympus-style f=180mm tube lens (TL) and f=180mm Fourier lens (FL) are used to conjugate the back focal plane of the objective onto a microlens array (MLA). An sCMOS sensor is positioned at the focal plane of the MLA to capture the raw fLFM images.

**B.** An example raw sensor image. Each lenslet in the 8×8 array forms an image of the sample from a slightly different perspective, allowing for reconstruction of the 3D volume.

**C.** Experimental measurement of the point spread function (PSF). Left: the x-y (top) and x-z (bottom) profiles of a reconstructed image of a 1 μm fluorescent bead. Right: corresponding cross sections of the PSF (points). A Gaussian profile (lines) was fit to these data to measure the full width at half maximum (FWHM), which was 3.3 μm laterally and 5.4 μm axially.

**D.** Schematic of our fLFM data processing pipeline (see Methods for detailed description).

**E.** A maximum intensity projection (MIP) of a conventional LFM (cLFM) larval zebrafish recording resulting in strong grid-like artifacts and low resolution near the native image plane (black arrows). A reconstructed volume of  $750 \times 750 \times 200 \mu\text{m}^3$  is shown.

**F.** An MIP of a fLFM recording. fLFM exhibits higher axial resolution and does not contain artifacts at the native image plane. A reconstructed volume of  $750 \times 375 \times 250 \mu\text{m}^3$  is shown.

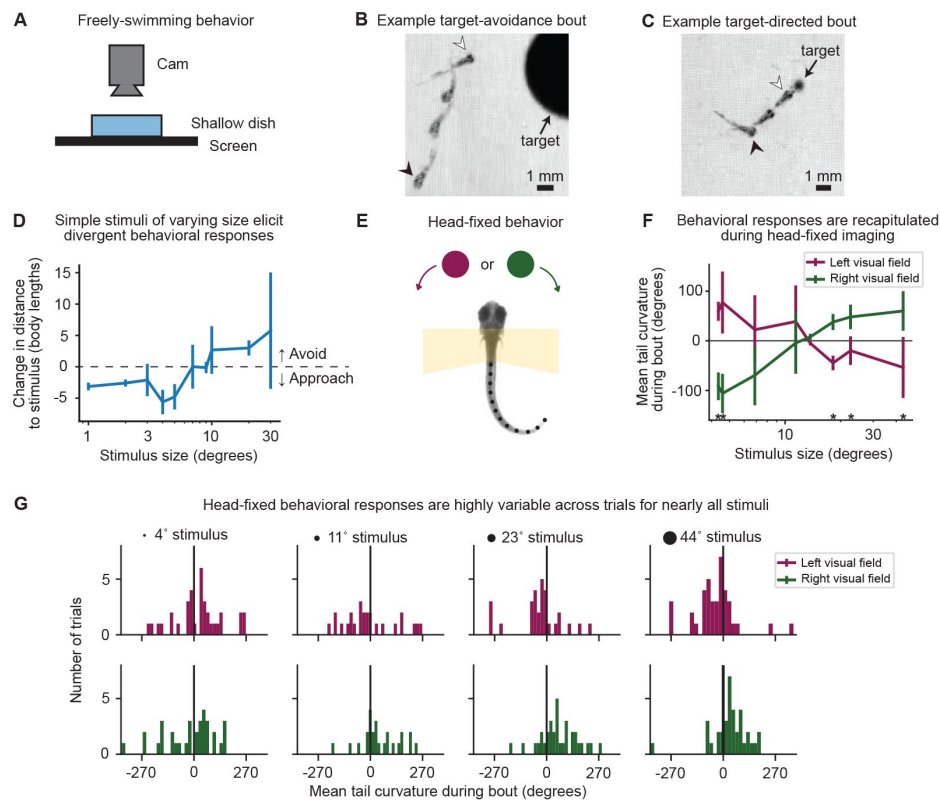
**G.** A heatmap of extracted neuronal activity from an example one hour recording of 15,286 neurons. Neurons are sorted using the rastermap algorithm (Stringer et al., 2019), such that nearby neurons exhibit similar temporal activity patterns. Multiple distinct activity patterns are seen, included strong activity during two sets of drifting grating-induced optomotor response trials (black arrows).

**H.** Example neurons tuned to tail movements and visual stimuli. In black are six example neuron traces from the designated region in panel G, which exhibited correlations with the GCaMP-kernel-convolved tail vigor (top trace, see Methods for details). In red are six example neuron traces from the designated region in panel G, which exhibited correlations with the visual stimuli presented in the left visual field (denoted by the red lines at the bottom of the plot).

## Larval zebrafish exhibit highly variable motor responses to visual stimuli

Whole-brain, cellular-resolution imaging at high speed enabled us to screen for regions across the brain which exhibited covariations related to variability in behavior and naturalistic decision making. We thus set out to investigate the neuronal basis of trial-to-trial variability in action selection during two different ethologically-relevant behaviors which are known to be generated by brain-wide sensorimotor circuitries (Chen et al., 2018 [DOI](#)). In the larval zebrafish, several visually-evoked behaviors such as prey capture (Borla et al., 2002 [DOI](#); Patterson et al., 2013 [DOI](#)) and escape response (Temizer et al., 2015 [DOI](#)) have been widely studied. In particular, varying a single stimulus parameter, the size of a visual object, can elicit dramatically distinct average behavioral responses (Barker and Baier, 2015 [DOI](#)), ranging from target-directed responses (e.g., prey capture) to small stimuli, to target-avoidance behaviors (e.g., escape response) to large stimuli. These distinct behaviors can be evoked in the laboratory using simple visual stimuli and in a head-fixed imaging preparation (Bianco et al., 2011 [DOI](#); Semmelhack et al., 2014 [DOI](#); Temizer et al., 2015 [DOI](#); Barker and Baier, 2015 [DOI](#); Bianco and Engert, 2015 [DOI](#); Dunn et al., 2016a [DOI](#); Filosa et al., 2016 [DOI](#); Förster et al., 2020 [DOI](#); Oldfield et al., 2020 [DOI](#)). As such, the neural circuitry underlying each of the individual behaviors is well understood (Portugues and Engert, 2009 [DOI](#); Bollmann, 2019 [DOI](#)) on the trial-averaged level. However, importantly these behaviors are not reflexive. They involve the processing of information across multiple brain regions and are subjected to neuromodulatory effects and longer time scale changes of internal states (Filosa et al., 2016 [DOI](#); Marques et al., 2020 [DOI](#)). As such, the larvae do not deterministically perform the same actions over multiple sensory invariant trials. Even under highly optimized and invariant conditions, prey capture is only observed in up to 30% of trials (Bianco et al., 2011 [DOI](#)), whereas escape response is more consistently observed in 80% of trials (Temizer et al., 2015 [DOI](#)). However, the neuronal populations and the underlying circuitry driving such variability in responsiveness and action selection in either of these paradigms have not been identified, yet these ethologically highly divergent behaviors can be elicited by tuning a single parameter. We thus hypothesized that modulating the size of the visual object should reveal a regime of behavioral variability between these highly divergent, naturalistic behaviors, providing a useful means and context to study questions such as how behavioral strategies may switch across trials or how neural and behavioral variability changes as a function of the ambiguity of the presented stimulus.

Thus, we first sought to characterize the trial-to-trial variability of the larvae's behavioral responses to repetitions of simple visual stimuli of various sizes. In a freely swimming context (**Figure 2A** [DOI](#)), we displayed moving dots of various sizes ranging from 1 to 40 visual degrees (see Methods for details) on a screen positioned below the larvae. Analyzing these behavioral recordings for different dot sizes, we indeed found characteristic target-avoidance bouts similar to escape response (**Figure 2B** [DOI](#)) and target-directed bouts similar to prey capture response (**Figure 2C** [DOI](#); see also Video 1). For each bout of behavior in which a larva was close to a visual object, we quantified whether it approached or avoided the target using the larva's change in distance from the object as a metric (see Methods for details). Consistent with previous work (Barker and Baier, 2015 [DOI](#)), on average the larvae exhibited target-directed behavior to small stimuli (1-5°), followed by a crossover point at which the larvae were not preferentially target-directed or target-avoiding for stimuli of 7° in size, and then a target-avoidance behavioral regime to stimuli at above 10° or larger which elicited characteristic escape responses on average (**Figure 2D** [DOI](#)). Moreover, across individual trials the larvae's responses exhibited significant variability, yielding many trials in which the larvae either did not respond or responded in the opposite direction of the average response, as indicated by the large standard deviation across bouts for the same stimulus size (**Figure 2D** [DOI](#)).



**Figure 2**

### Zebrafish exhibit highly variable motor responses to visual stimuli

**A.** Schematic of the freely-behaving experimental setup. A single larva is placed in a 90 mm petri dish with a shallow (approximately 3 mm) amount of water. The dish is placed on a screen which displays visual stimuli consisting of dots of various sizes drifting in a constant direction and speed. The behavior is monitored with a camera from above.

**B.** An example target-avoidance bout. A composite image taken of four frames from start (white arrow) to near the finish (black arrow) of a bout in which the larvae avoided a large visual stimulus. See also Video 1.

**C.** An example target-directed bout. A composite image taken of three frames from start (white arrow) to finish (black arrow) of a bout in which the larvae made a directed movement toward a small visual stimulus. See also Video 1.

**D.** The behavioral response of larvae to various size visual stimuli. For each size stimulus, all bouts of movement in which the larva was within 10 body lengths of a visual stimulus were considered. For each bout, the change in distance between the larvae and the stimulus during the movement was monitored. Thus, target-directed bouts exhibited negative values corresponding to a decrease in the distance to the stimulus, whereas target-avoidance bouts exhibited positive values. Stimuli less than 7° in size evoked target-directed responses on average, whereas stimulus greater than 10° evoked target-avoidance responses. Shown is the mean  $\pm$  standard deviation of  $n=9$  larvae.

**E.** Schematic of the head-fixed experimental setup. A larva is embedding in agarose in order to keep its head fixed during whole-brain imaging. The agarose around the tail and eyes are removed, such that the tail can be tracked (indicated by black dots along the tail) and visual stimuli can be presented. Visual stimuli of various sizes are presented moving from the center of visual field to either the left or right.

**F.** The behavioral response of larvae to various size visual stimuli during head fixation. Visual stimuli are presented in either the left (pink) or right (green) visual field. The directedness of tail movements is monitored by computing the mean tail curvature during a bout of movement, with positive values indicating leftward motions. Similar to freely-behaving experiments, visual stimuli of 7° or less evoked target-directed responses, whereas stimuli larger than 10° evoked target-avoidance. Shown is the mean  $\pm$  standard deviation of  $n=10$  larvae (inter- and intra-larva variability is described in [Figure S2](#)). Asterisks indicate stimuli with significant difference between presentations on the left and right visual field ( $p<0.05$ , paired t-test).

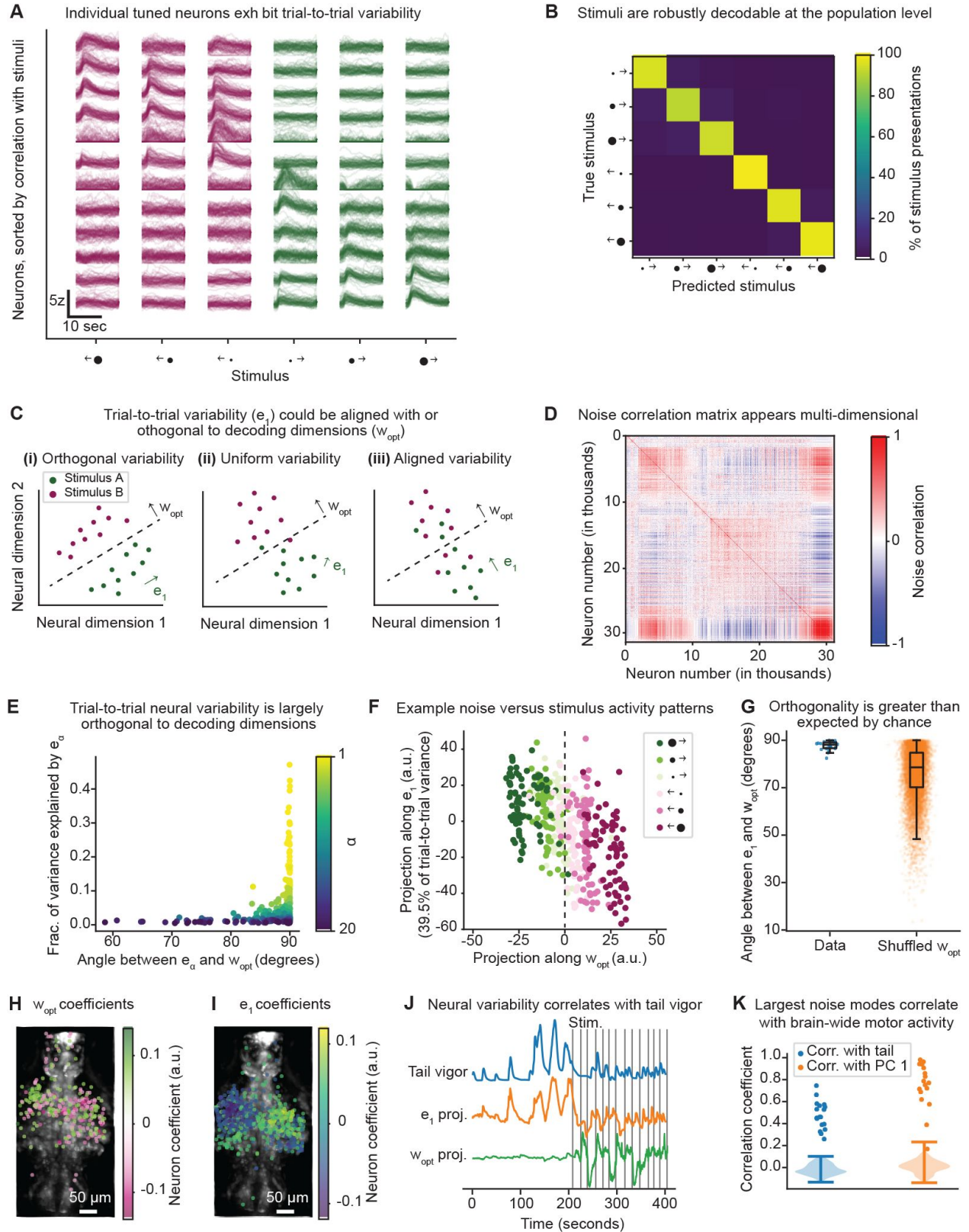
**G.** Example behavioral responses to various stimuli. For each of four stimulus sizes on the left and right visual fields, a histogram of the mean tail curvature during bouts is shown for an example larva. While stimulus-evoked behavioral responses are variable in all cases, they appear the least variable in the case of largely target-avoidance bouts to large 44° stimuli (rightmost column).



Next, we confirmed that these behaviors were preserved in a head-fixed imaging preparation, in which the larvae were embedded in agarose with their eyes and tail cut free (**Figure 2E** [↗](#)), such that their variable behavioral responses to visual stimuli could be tracked while performing whole-brain neuronal recordings using our fLFM system. To do so, we presented visual stimuli as dots of a certain size moving in either the left or right visual field of the animals (see Methods for details). Visual stimuli were created using a projector that projected onto the edge of the circular 35 mm petri dish in which the larvae were embedded. Utilizing curvature of the tail to the left or the right as a metric for target-directed versus -avoidance behavior, we again identified the stimulus-size dependence of the larvae's average behavioral response (**Figures 2F** [↗](#), **S2A** [↗](#)) that recapitulated what was observed in freely behaving larvae. We also confirmed that the corresponding opposite tail responses were observed when the stimuli were presented on the other side of the visual field, demonstrating that even during head-fixation the larvae exhibit on average directed behaviors that depend on the stimulus size. However, at the single trial level, we observed behavioral variability to nearly all stimulus sizes (**Figures 2G** [↗](#), **S2B** [↗](#)), even in the stimulus size regime known to optimally drive prey capture and escape response behaviors. The most widely variable responses were at 4° and 11°, corresponding to stimuli in the target-directed regime and near the crossover between target-directed and avoidance, respectively. This high degree of variability among responses to small and intermediate stimuli is consistent with previous literature estimates that even in optimized setups, the prey capture response is observed at rates less than 30% during head-fixation (Bianco et al., 2011 [↗](#)) and any motor response at all to small visual stimuli is observed at rates up to 60% of trials in free behavior (Filosa et al., 2016 [↗](#)). However, the neural mechanisms and circuitry underlying such trial-to-trial behavioral variability are not understood. In particular, it is unclear whether such variability within these ethologically relevant, yet divergent behaviors is driven by noise and inherent stochasticity in neuronal circuits, or represents modulation by time-varying internal states, such as motivation and hunger.

## Trial-to-trial variability in visually-evoked neurons is largely orthogonal to visual decoding dimensions

Previous studies have reported that individual neurons tuned to specific stimuli often exhibit significant variability in their responses over multiple presentations of the same stimulus (Montijn et al., 2016 [↗](#); Rumyantsev et al., 2020 [↗](#); Shadlen and Newsome, 1998 [↗](#); Zohary et al., 1994 [↗](#); Zylberberg et al., 2016 [↗](#)). Thus, one potential mechanism underlying our observed behavioral variability could be trial-to-trial deviations in the neural encoding of visual stimuli. Given that we had observed behavioral variability across the entire studied range of stimulus sizes, we proceeded to record whole-brain dynamics with the fLFM while presenting many repetitions of 3 to 6 different stimulus sizes spanning 1° to 40° (see Methods for details). Investigating the responses of individual neurons across the whole brain, we indeed found that even neurons highly tuned to a particular stimulus exhibited variability in their responses across multiple presentations of the same stimuli (**Figure 3A** [↗](#)). Given that downstream decision-making neuronal ensembles likely pool information from many visually tuned neurons, we asked whether an optimal population decoder could reliably extract information about the visual stimulus across trials. We proceeded to build various logistic regression classifiers to decode which visual stimulus was presented from the whole-brain neuronal activity pattern during the stimulus presentation period of each trial. We found that despite the observed variability on the single neuron level, the stimuli could be reliably decoded from the visually-tuned neurons identified within the whole-brain population activity (**Figure 3B** [↗](#)) at the single trial level with an accuracy of  $94 \pm 2\%$  (mean  $\pm$  95% confidence interval). This robust decodability suggests that the observed single neuron variability does not severely limit the information encoding capacity at the population level.



**Figure 3**

### **Trial-to-trial variability in visually-evoked neurons is largely orthogonal to visual decoding dimensions**

**A.** Individual visually tuned neurons exhibit trial-to-trial variability in their responses to stimuli. For each of the six visual stimuli (three object sizes on both the left and right visual fields), the two neurons exhibiting the highest correlation with the stimulus kernel (see Methods for details) are shown for an example larva. Each column represents a neuron, and each line represents its response to a given stimulus during a single trial. To visualize the neuronal activity during a given trial while accounting for the delay and kinematics of the nuclear-localized GCaMP (NL-GCaMP) sensor, a duration of approximately 15 seconds is extracted beginning at the onset of the 3-second visual stimulus period.

**B.** Visual stimuli are reliably decodable from whole-brain dynamics on the single trial level. The visual stimuli were decoded using a one-versus-rest, multiclass logistic regression classifier with lasso regularization (see Methods for details). For each larva, a confusion matrix is computed for the test trials during 6-fold cross-validation. Shown is the average confusion matrix across  $n=8$  larvae which were shown six visual stimuli.

**C.** Schematic of potential geometric relationships between sensory decoding and neural variability dimensions. In each plot, each dot represents the neural response during a single presentation of stimulus A or B. The decision boundary for an optimal classifier is denoted with a dashed line, and the optimal stimulus decoding direction is denoted by the vector  $\vec{w}_{\text{opt}}$ . The direction representing the maximal trial-to-trial variance is denoted by  $\vec{e}_1$ , and can be calculated by finding the first eigenvector of the noise covariance matrix (see Methods for details). These vectors can be: (i) orthogonal, such that neuronal variability does not limit the stimulus decoding; (ii) show little relationship, for example in the case of uniform variability; or (iii) aligned, such that variability likely limits the information encoding capacity along  $\vec{w}_{\text{opt}}$ .

**D.** The trial-to-trial noise correlation matrix appears multi-dimensional. Shown is the average noise correlation matrix across all stimulus types presented. The neurons are sorted using rastermap, which produces a one-dimensional embedding of the neurons, such that neurons which show similar correlation profiles are placed near to one another. A number of neuronal populations exhibiting correlations across trials are apparent from the clusters of high correlations near the diagonal.

**E.** Trial-to-trial variability in the visually-evoked neurons is largely orthogonal to visual decoding dimensions. The fraction of trial-to-trial variance explained by each noise mode  $\vec{e}_\alpha$  is plotted against the angle between  $\vec{e}_\alpha$  and the optimal stimulus decoding direction  $\vec{w}_{\text{opt}}$ . Shown are the  $\vec{e}_\alpha$  for  $n=8$  larvae, colored by their rank order  $\alpha$  based on the fraction of variance explained. The largest noise modes were approximately orthogonal ( $\sim 90^\circ$ ) to the stimulus decoding direction, whereas only a few of the smallest noise modes exhibited angles less than  $90^\circ$ .

**F.** Example projections of single trial neural activity along the stimulus decoding and noise dimensions. Each dot represents the average neural activity within a single trial projected along  $\vec{w}_{\text{opt}}$  and  $\vec{e}_1$  for an example larva. Each of the six visual stimuli, three object sizes presented on either the right (green) or left (pink) visual field, are robustly encoded along  $\vec{w}_{\text{opt}}$  across trials; however, in all stimuli there is strong orthogonal variability along  $\vec{e}_1$ , the largest noise mode representing 39.5% of the trial-to-trial variance.

**G.** The approximately orthogonal angles between  $\vec{w}_{\text{opt}}$  and  $\vec{e}_1$  across larvae are greater than expected by chance. Shuffled versions of  $\vec{w}_{\text{opt}}$ , representative of random vectors in the neural state space, were computed by permuting the stimulus labels before performing stimulus decoding. The distribution of angles between  $\vec{e}_1$  and  $\vec{w}_{\text{opt}}$  is significantly greater than  $\vec{w}_{\text{opt,shuff}}$  ( $p < 0.05$ , Wilcoxon rank-sum test).

**H.** Example neuron coefficients for  $\vec{w}_{\text{opt}}$ . Shown are the 366 neurons with the largest weights over a maximum intensity projection of the recorded volume. The visually-evoked neurons which encode stimulus information are concentrated within the optic tectum.

**I.** Example neuron coefficients for  $\vec{e}_1$ . Shown are the 828 neurons with the largest weights over a maximum intensity projection of the recorded volume. The visually-evoked neurons contributing to the largest noise mode are highly overlapping with the neurons contributing to  $\vec{w}_{\text{opt}}$  in panel G.

**J.** The neural noise modes are highly correlated with tail movements. Shown are both the GCaMP kernel-convolved tail vigor and the neuronal projection onto the first noise mode  $\vec{e}_1$  for a representative larva. Over the full two-hour recording, the tail vigor and noise mode projection exhibit a significant correlation of  $r=0.58$ ,  $p<10^{-6}$ . For comparison,  $\vec{w}_{\text{opt}}$  and the start time of the visual stimuli (black vertical lines) are shown.

**K.** The largest neural noise modes reflect brain-wide motor encoding. Shown are the correlations between the first noise mode  $\vec{e}_1$  and the tail vigor (blue) or the first principal component (PC 1, orange) of whole-brain data (see **Figure S3C**). Dots show data from individual larvae, whereas the violin plots below show the null distribution for temporally shuffled data, in which the tail vigor or PC 1 are circularly permuted.

Given the wealth of literature debating whether trial-to-trial noise correlations could interfere with stimulus encoding (Moreno-Bote et al., 2014; Kanitscheider et al., 2015; Zylberberg et al., 2017; Bartolo et al., 2020; Rummyantsev et al., 2020; Valente et al., 2021; Kafashan et al., 2021) and potentially drive behavioral variability, we next asked how the observed neuronal variability was structured to prevent degradation of the visual encoding. We began by considering the dimensions that maximally explain trial-to-trial neuronal variability within the high-dimensional space of neuronal population dynamics, which are sometimes referred to as “noise modes” (Rummyantsev et al., 2020). Historically, the term “noise” is used to describe such trial-to-trial variability that is not related to the external stimulus; however, such noise could represent significant internal neuronal dynamics and not just measurement shot noise, for example. If the primary noise modes (represented by neural activity along the vectors  $\vec{e}_\alpha$ ) are orthogonal to the optimal stimulus decoding direction ( $\vec{w}_{\text{opt}}$ ), then trial-to-trial variability will not hinder the visual stimulus decoding (**Figure 3C(i)**). In this case, the neuronal population may carry a multi-dimensional set of at least two variables and the visual stimulus encoding would be preserved. Trial-to-trial variability could also not be limited to a particular set of dimensions, for example if it represented noise that is uniformly distributed in the neural state space (**Figure 3C(ii)**), in which case pooling or averaging across sufficient neurons may still recover adequate stimulus information. Finally, trial-to-trial noise modes may be aligned with the stimulus decoding dimension (**Figure 3C(iii)**), which would maximally degrade the visual stimulus decoding.

To assess which of these scenarios applied to the geometry of trial-to-trial neural variability in our data, we first utilized partial least squares (PLS) regression to identify a low-dimensional subspace of the whole-brain neural activity that optimally preserved information about the visual stimulus (**Figure S3A**, see Methods for details), which we refer to as the visually-evoked neuronal activity patterns. Importantly, this approach ensures that we are identifying trial-to-trial neural variation that is contained within the visually-evoked neurons, as opposed to variability that is completely unrelated to visual encoding. Within these visually-evoked neurons, PLS regression also identifies the optimal stimulus decoding direction  $\vec{w}_{\text{opt}}$  (**Figure S3B**). Additionally, trial-to-trial variability of the visually-evoked neurons is summarized by their noise covariance matrix, which describes how the activity of neurons covaries across trials (Moreno-Bote et al., 2014; Kohn et al., 2015). The average noise correlation matrix across all stimuli (see Methods for details) appeared highly structured and multi-dimensional (**Figure 3D**), indicating the presence of multiple neuronal ensembles whose activity is strongly correlated across trials. We analyzed the structure of this trial-to-trial noise by finding the eigenvectors  $\vec{e}_\alpha$  of the average neural noise covariance matrix across all stimuli, following (Rummyantsev et al., 2020). As such, these noise modes  $\vec{e}_\alpha$  represented the dimensions of neuronal activity within the visually-evoked neurons that maximally covaried across trials. We found that the noise modes were strongly structured since the trial-to-trial variance was concentrated in relatively few dimensions. The single largest noise mode captured up to 50% of the variance across larvae (**Figure 3E**, y-axis), indicating that such

activity is likely not merely “noise” but potentially physiologically relevant. Thus, the observed neuronal variability was not independent across neurons, but strongly correlated. To assess the impact these noise modes could have on the stimulus decoding, we calculated the angles between the noise modes  $\vec{e}_\alpha$  and  $\vec{w}_{\text{opt}}$ . We discovered that the noise modes containing the majority of the variance were orthogonal to  $\vec{w}_{\text{opt}}$ , while only the smallest noise modes contained any overlap the stimulus decoding direction  $\vec{w}_{\text{opt}}$  (**Figure 3E** [↗](#)).

This finding demonstrates that the structure underlying trial-to-trial neuronal variability in the zebrafish visual system lies in a regime closely resembling the orthogonal variability as illustrated in **Figure 3C(i)** [↗](#), as also previously observed in a subregion of mouse primary visual cortex (Rumyantsev et al., 2020 [↗](#)). Thus, while even a small degree of observed overlap between sensory and noise modes can in principle limit the information encoding capacity of the visually-evoked neurons (Moreno-Bote et al., 2014 [↗](#)), this orthogonal structure minimizes this limitation and is consistent with our previous finding that the stimuli presented in this study are highly decodable from the neuronal population dynamics (**Figure 3B** [↗](#)). In addition, while the angle between the sensor and noise modes arises from the pattern of coefficients across neurons, we can visualize activity along these modes by projecting average neural activity within each trial onto the stimulus decoding direction  $\vec{w}_{\text{opt}}$  and the noise modes (**Figure 3F** [↗](#)). This demonstrates that that majority of the neural variability across trials (given by the projection onto  $\vec{e}_1$ ) is largely independent from the sensory activity (along  $\vec{w}_{\text{opt}}$ ). Finally, since these modes are defined in a multi-dimensional space spanned by the  $d$  PLS components, we must consider the chance that any two random vectors would appear orthogonal, which becomes increasingly common in higher-dimensional spaces. We found that the distribution of angles between  $\vec{e}_1$  and  $\vec{w}_{\text{opt}}$  across larvae was significantly greater than the angles between  $\vec{e}_1$  and randomly shuffled versions of  $\vec{w}_{\text{opt}}$  (**Figure 3G** [↗](#), see Methods for details), confirming that this nearly orthogonal structure is not simply a result of the multi-dimensional nature of the PLS space.

In order to confirm that the observed neural variability in the visually-evoked populations was not predominantly due to eye movements, such as saccades or convergence, we tracked the angle of each eye. We utilized DeepLabCut, a deep learning tool for animal pose estimation (Mathis et al., 2018 [↗](#)), to track keypoints on the eye which are visible in the raw fLFM images, including the retina and pigmentation (**Figure S3D(i)** [↗](#)). This approach enabled identification of various eye movements, such as convergence and the optokinetic reflex (**Figure S3D(ii-iii)** [↗](#)). Next, we extracted a number of various eye states, including those based on position (more leftward vs. rightward angles) and speed (high angular velocity vs. low or no motion). **Figure S3E(i)** [↗](#) provides example stimulus response profiles across trials of the same visual stimulus in each of these eye states, similar to a single column of traces in **Figure 3A** [↗](#) broken out into more detail. These data demonstrate that the magnitude and temporal dynamics of the stimulus-evoked responses show apparently similar levels of variability across eye states. If neural variability was driven by eye movement during the stimulus presentation, for example, one would expect to see much more variability during the high angular velocity trials than low, which is not apparent. Next, we asked whether the dominant neural noise modes vary across eye states, which would suggest that the geometry of neuronal variability is influenced by eye movements or states. To do so, the dominant noise modes were estimated in each of the individual eye conditions, as well as bootstrapped trials from across all eye conditions. The similarity of these noise modes estimated from different eye conditions (**Figure S3E(ii)** [↗](#), right) was not significantly different from the similarity of noise modes estimated from bootstrapped random samples across all eye conditions (**Figure S3E(ii)** [↗](#), left). Therefore, while movements of the eye likely contribute to aspects of the observed neural variability, they do not dominate the observed neural variability here, particularly given our observation that the largest noise mode represents a considerable fraction of the observed neural variance (**Figure 3E** [↗](#)).



However, in high-dimensional spaces, it becomes increasingly common that two random vectors could appear orthogonal. While this is particularly a concern when analyzing a neural state space spanned by tens of thousands of neurons, our application of PLS regression to identify a low-dimensional subspace of relevant neuronal activity partially mitigates this concern. In order to control for this confound, we compared the angles between  $\bar{w}_{\text{opt}}$  and  $\bar{e}_1$  across larvae to that computed with shuffled versions of  $\bar{w}_{\text{opt,shuff}}$  estimated by randomly shuffling the stimulus labels before identifying the optimal decoding direction. While it is possible to observe shuffled vectors which are nearly orthogonal to  $\bar{e}_1$ , the shuffled distribution spans a significantly greater range of angles than the observed data ( $p < 0.05$ , Wilcoxon rank-sum test), demonstrating that this orthogonality is not simply a consequence of analyzing multi-dimensional activity patterns.

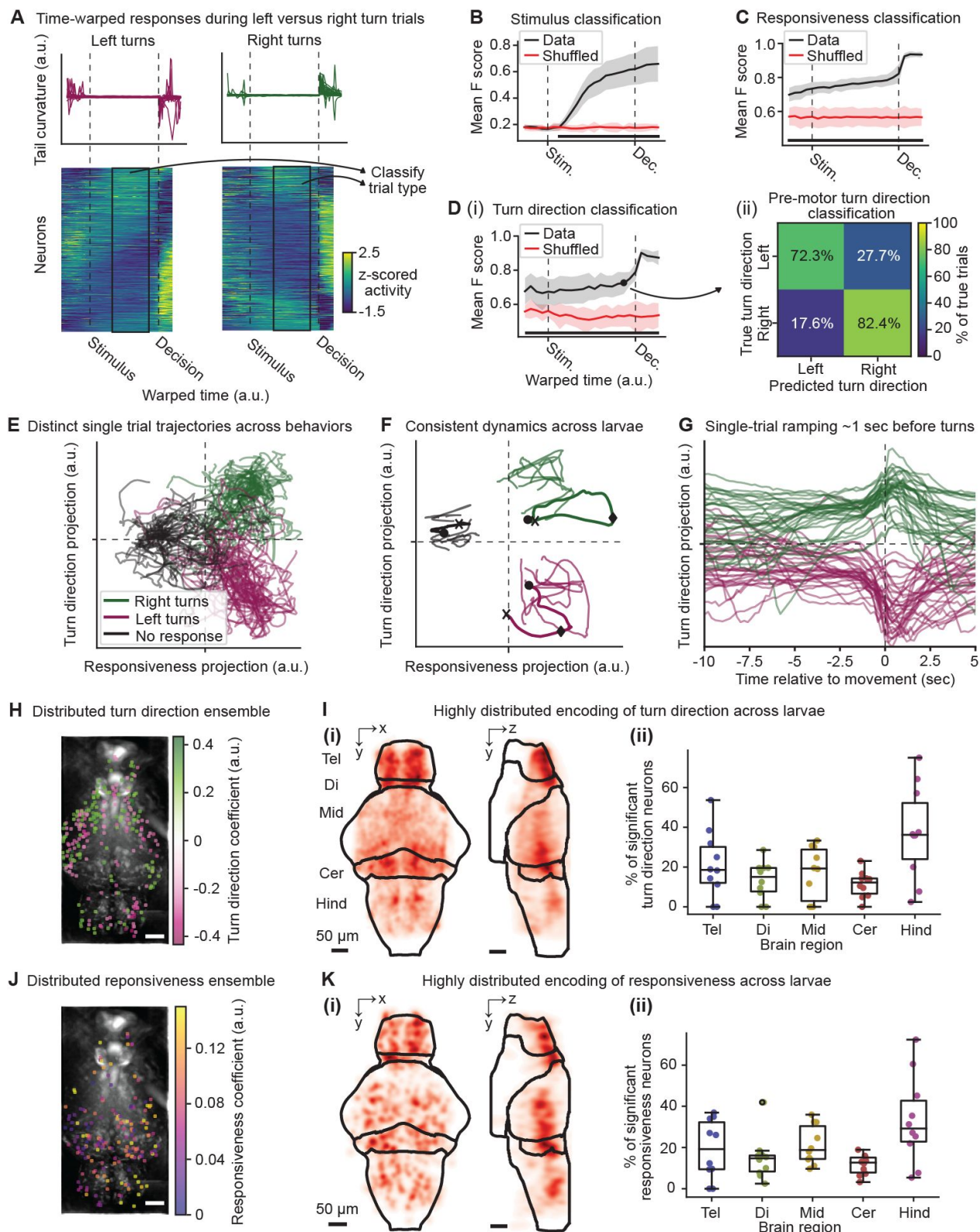
Next, we mapped these neural coefficient vectors defining the stimulus decoding direction and the noise modes onto their anatomical locations of the larvae's brain. As expected, the visually-evoked neurons were highly concentrated within the midbrain and specifically the optic tectum, such that both the stimulus decoding direction (**Figure 3G**) and noise modes (**Figure 3H**) highly overlapped. This confirms that the noise modes indeed represent fluctuations within the visual circuitry, as opposed to irrelevant activity from other brain areas. The significance of each neuron's contribution to  $\bar{w}_{\text{opt}}$  and  $\bar{e}_\alpha$  was assessed by comparing its coefficient to a distribution of coefficients derived from trial-shuffled data (see Methods for details). Overall,  $30 \pm 6\%$  of visually-evoked neurons that significantly contributed to the largest noise mode  $\bar{e}_1$  also exhibited significant contributions to the stimulus decoding direction  $\bar{w}_{\text{opt}}$  (mean  $\pm$  95% CI across  $n=10$  larvae). Therefore, while many visually-evoked neurons have response components along both  $\bar{e}_\alpha$  and  $\bar{w}_{\text{opt}}$ , the neuronal population dynamics are highly structured such that activity patterns encoding external visual stimuli are reliable and largely independent from any trial-to-trial noise modes.

Given that these noise modes represented highly-structured neuronal variability across trials, we next asked whether they were at all related to the larvae's behavior. First, we asked whether any noise modes were correlated with the vigor of the larvae's tail movements, defined as the absolute tail curvature time series convolved with the GCaMP response kernel (see Methods for details). We found that the largest noise modes were often correlated with the larvae's instantaneous tail vigor (**Figure 3J**), as well as other kinematic variables (**Figure S4A**), for example tail curvature (which includes a left versus right directedness). While we identified the noise modes within the visually-evoked neuronal population, these noise modes were also highly correlated with the largest principal component (PC) of the whole-brain dynamics (**Figure 3K**). This indicates that both the observed trial-to-trial neuronal variability in visual regions and elsewhere across the brain are dominated by behavior-related activity (**Figure S4B**). In fact,  $35 \pm 15\%$  of neurons with significant coefficients in the largest visual noise mode  $\bar{e}_1$  also significantly contributed to the first brain-wide PC (mean  $\pm$  95% CI across  $n=10$  larvae, see Methods for details). While the lower variance noise modes did not exhibit clear motor-related activity, we suspected they may be involved in additional internal computations that may modulate decision making across trials. Taken together, these results show that visual stimuli are faithfully represented on the population level, and that the observed behavioral variability is not driven by changes in the fidelity of sensory encoding across trials. Instead, behavioral variability may be explained by additional, orthogonal dimensions of neuronal activity either within the visually-evoked population or elsewhere in the brain that are stimulus-independent but predictive of the larvae's behavior.

## Pre-motor neuronal populations predictive of single-trial behavior

We hypothesized that there exists a pre-motor neuronal population that would be predictive of the larvae's behavioral response in any given trial *before* movement initiation. As such, we aimed to identify any neural locus predictive of the larvae's turn direction at various timepoints through each visual stimulus presentation. Thus, we looked beyond only the visually-evoked population and returned to the whole-brain activity patterns. In order to build predictive models from the neuronal dynamics despite varying reaction time across trials, we applied time-warping to the

neuronal timeseries within each trial such that the stimulus onset and movement initiation timepoints were aligned across trials (**Figure 4A** [↗](#), see Methods for details). We selected trials in which the larvae had a reaction time of at least one second from the stimulus onset (providing sufficient pre-motor neuronal data) and made its first tail movement with a defined minimum tail curvature to either the left or right, so as to separate clear left and right turns from simple forward locomotion. Further, we utilized trials from all visual object sizes to collect enough trials to reliably train models and thus identify populations predictive of decisions regardless of incoming sensory information. We found that there appeared to be neurons with differing time-warped activity patterns during left and right turn trials (**Figure 4A** [↗](#), bottom). Additionally, we extracted trials in which the larva did not respond following the stimulus presentation, which were time-warped according to a randomly selected reaction time from the distribution in responsive trials. As such, each trial could be categorized according to the stimulus presented, the larva's responsiveness (response/no response), and its turn direction (left/right) within responsive trials.



## Figure 4

### Pre-motor neuronal populations predictive of single-trial behavior.

**A.** Trials are time-warped to align stimulus and decision onsets before classifying the turn direction. Top: Time-warped tail curvatures for trials in which the fish performed leftward or rightward turns, on the left and right, respectively. Bottom: Trial-averaged and time-warped neuronal timeseries for 15,286 neurons during left and right turns. The neurons are sorted using the rastermap algorithm. A time window is swept across the stimulus and decision timepoints to train binary classification models to predict the turn direction from the associated neuronal dynamics.

**B.** Stimulus classification accuracy peaks after the onset of visual stimulation. The mean F score across  $n=7$  larvae is used to assess the performance of 6-way multiclass classification of the presented visual stimulus as a function of warped time surrounding the stimulus onset (Stim.) and decision timepoint (Dec.). Shown is the mean  $\pm$  95% confidence interval of the F score for the best time window ending at the given timepoint (Data), compared to shuffled data in which the class labels are randomized. The black bar at the bottom indicates timepoints where the data have a significantly higher F score than the shuffled data ( $p<0.05$ , paired t-test).

**C.** Binary classification of responsiveness, whether or not the fish responds in a given trial, is significant throughout all time periods but accuracy peaks near movement initiation. As in panel B, except for binary classification of responsiveness. Nonresponsive trials are time-warped by randomly selecting a reaction time from the response trials and applying the same transformation.

**D.** (i) Turn direction classification accuracy is significantly higher than shuffled data across the entire time-warped interval, but peaks near movement initiation. As in panel B, except for binary classification of turn direction. (ii) Single trial classification of turn direction across larvae. The mean confusion matrix across  $n=7$  larvae, which show an accuracy of  $77 \pm 4\%$  (mean  $\pm$  95% confidence interval).

**E.** Single trial trajectories are separated based on responsiveness and turn direction. Shown are neural activity trajectories during single trials in an example larva projected onto the brain-wide neural dimensions that optimally separated turn direction and responsiveness.

**F.** Consistent trial-averaged trajectories across larvae. As in panel F, except for the trial-averaged responses for  $n=6$  example larvae. For the one bold animal, timepoints across the trial are indicated by a circle for trial start, diamond for the decision timepoint, and an X for the trial end.

**G.** Real-time single-trial dynamics in an example larva. Along the turn direction neural projection, left and right trials are separated for many seconds before the decision timepoint, which is longer than the three second length of visual presentations. Activity along this dimension shows consistent ramping across trials approximately one second before movement.

**H.** Example turn direction neuronal ensemble. Shown are the coefficients for all neurons which showed significantly higher (one-tailed t-test,  $p<0.05$ ) absolute coefficients in the real models compared to shuffled data in which the turn direction labels are randomly permuted. Scale bar:  $50 \mu\text{m}$ .

**I.** Highly distributed encoding of turn direction across larvae. The significant turn direction neurons located within five major brain regions are shown for  $n=10$  larvae. (i) The 3D distribution of neurons across larvae aligned to the Z-Brain atlas. (ii) Quantification of the percentage of turn direction neurons located within each brain region shown in (i). There is no significant difference between the percentage of neurons across brain regions ( $p>0.05$ , paired t-test).

**J.** Example responsiveness neuronal ensemble. As in panel I, except for responsiveness. Shown are the coefficients for all neurons which showed significantly higher (one-tailed t-test,  $p<0.05$ ) absolute coefficients in the real models compared to shuffled data in which the turn direction labels are randomly permuted. Scale bar:  $50 \mu\text{m}$ .

**K.** Highly distributed encoding of responsiveness across larvae. As in panel J, except for responsiveness. There is no significant difference between the percentage of neurons across brain regions ( $p > 0.05$ , paired t-test).

We first asked whether the larvae's responsiveness and turn direction could be accurately classified from the whole-brain neuronal dynamics within single trials, as well as how this predictability behaved over the pre-motor period. To identify the timepoints throughout the trial within which these could be accurately classified, we varied the time window of neuronal dynamics used to build a binary logistic regression classification model (**Figure S5A** [↗](#)). Neuronal activity across all neurons was integrated from the window start to the window end separately for each trial and then a model was trained on 80% of the trials. To assess the prediction accuracy at various timepoints throughout the trial, we utilized the mean F score across larvae on held-out trials, which measures the harmonic mean of the precision and recall. As expected, the multiclass stimulus prediction accuracy peaked shortly after the stimulus onset and remained high throughout the pre-motor and decision periods (**Figure 4B** [↗](#)). Performing stimulus classification again using only neurons from individual brain regions, we found that only populatinos in the midbrain and cerebellum exhibited strong stimulus decoding (**Figure S5B** [↗](#)). For both the responsiveness (**Figure 4C** [↗](#)) and the turn direction (**Figure 4D(i)** [↗](#)), the mean F score was consistent across the pre-stimulus and stimulus periods, before quickly ramping near the movement initiation timepoint. Interestingly, the classification performance was significantly higher than shuffled data at all timepoints (**Figure 4C-D** [↗](#)), including before the start of the stimulus presentation, with an average accuracy of  $77.4 \pm 4.4\%$  (mean  $\pm$  95% CI across  $n=10$  larvae) during the pre-motor period (**Figure 4D(i)** [↗](#)). These results suggest there are two dominant timescales which contribute to the larva's behavioral response: a longer-timescale and relatively weak (but significant) activity pattern that does not appear aligned to stimulus or behavioral events, which we term the pre-stimulus turn bias; and a fast-timescale dramatic increase in predictability very near the decision, which we term the movement initiation dynamics.

Given that we observed the largest degree of behavioral variability in the responses to small-to-intermediate size stimuli (**Figure 2G** [↗](#)), we hypothesized that the pre-stimulus turn bias would be most influential (and thus most predictive of behavior) during such stimulus presentations, and less so during presentations of larger stimuli. Indeed, while the mean F scores for turn direction prediction were highly variable across larvae, on average the largest stimuli exhibited the lowest predictability from the pre-stimulus turn bias signal (**Figure S5C** [↗](#), blue line). This likely reflects a more salient stimulus drive during large stimulus trials, which in the natural environment could reflect predators and a potential life-or-death decision for the larvae. However, significant predictability was observed across all stimulus sizes during the movement initiation dynamics (**Figure S5C** [↗](#), orange line). We thus argue that the pre-stimulus turn bias is subsequently combined with incoming stimulus information during action selection. Ultimately, the larva's decision is then signaled by the dramatic increase in predictability during the movement initiation dynamics.

These single-trial predictions could be visualized by projecting the brain-wide neuronal activity during each trial onto the neural dimensions that optimally classified turn direction and responsiveness. Consistent with our classification performance, we found that the trajectories during single trials (**Figure 4E** [↗](#)) were highly separated according to the ultimate behavioral response. On average and across larvae, these trajectories remained localized in distinct quadrants throughout the entire trial (**Figure 4F** [↗](#)), including before and during the stimulus presentation period. This suggests that the observed behavioral variability is due in part to the initial internal state of the neural dynamics at the time of stimulus presentation and decision. Indeed, by removing the time-warping and observing the dynamics during single trials in real time, left and right turn trials appear separated for many seconds prior to the movement (longer than the 3 second visual trials). Ultimately, the decision to turn during the movement initiation dynamics appeared driven by a strong ramping activity approximately one second before the turn



(**Figure 4G** [↗](#)). Thus, the identified neuronal dimensions exhibited reliable decoding within single trials driven by an initial bias and ultimately resulting in the observed movement initiation dynamics.

Next, we asked which specific neuronal populations across the brain contributed to the accurate pre-motor classification of behaviors, and thus could be involved in driving the decision of when and how to act. We considered a neuron to be significantly contributing to the behavioral prediction if its distribution of absolute coefficients across 6-fold cross-validation was significantly higher (one-tailed t-test,  $p < 0.05$ ) when fit to real behavioral data as opposed to a null distribution created by randomly permuting the trial type labels. We found that both the ensemble of neurons which significantly contributed to the turn direction prediction (**Figures 4H-I** [↗](#)) and the responsiveness ensemble (**Figures 4J-K** [↗](#)) were highly distributed across the brain, with potentially dense localization in the telencephalon, cerebellum, and dorsal diencephalon (habenula). Both of these distinct neuronal ensembles showed no significant difference in their distribution across major brain regions: the telencephalon (Tel), diencephalon (Di), midbrain (Mid), cerebellum (Cer), and the remaining hindbrain (Hind). Further, while this population included visually-evoked neurons within the optic tectum, we found that our previously identified noise modes  $\vec{e}_\alpha$  were similarly predictive of single trial turn direction (**Figure S5D** [↗](#)), whereas the optimal stimulus decoding direction  $\vec{w}_{\text{opt}}$  (**Figure 3G** [↗](#)) was not able to predict single-trial turn direction (**Figure S5E** [↗](#)).

Finally, given that our turn direction and responsiveness classification required using lasso regularization to induce sparsity and achieve strong prediction performance, we aimed to identify whether a larger selection of neurons may contribute to or reflect these activity patterns by quantifying the correlation between each neuron and the projection of the whole-brain neural activity along the optimal turn direction or responsive prediction directions. A neuron was considered significantly correlated with either projection if its correlation was significantly higher than that of shuffled data obtained by circularly permuting the neuron's timeseries (see Methods for details). These distributions (**Figures S5F-G** [↗](#)) strongly resembled those in **Figures 4H-K** [↗](#) and supported the brain-wide nature of these representations. Further, the predictability of turn direction or responsiveness was similar across these brain regions (**Figures S5H-I** [↗](#)). Thus, our data suggest that when stimuli are ambiguous, single-trial action selection is largely explained by a widely-distributed circuit containing subpopulations encoding internal time-varying biases related to both the larva's responsiveness and turn direction, yet distinct from the sensory encoding circuitry.

What could be the origin of these trial-to-trial biases in the larvae's behavior? Given that our results demonstrate that it is possible to predict a behavioral bias before the visual stimulus is shown, we hypothesized these results could reflect the influence of the zebrafish neuronal circuitry for spontaneous turns, which includes the hindbrain oscillator (or anterior rhombencephalic turning region, ARTR) known to bias sequential turns towards a given direction (Dunn et al. 2016b [↗](#)). While our pre-motor neuronal population is not exclusively localized to the ARTR, we did identify many neurons in the cerebellum and hindbrain (**Figures 4J-K** [↗](#)) whose locations were consistent with these known populations. As such, we asked whether spontaneous turns could be predicted from this same turn direction neuronal ensemble (**Figure 5A** [↗](#)). To address this question, we selected spontaneous turns that occurred either during inter-trial intervals or within a 2-minute period at the beginning of each recording without any stimulus presentations, utilizing the same criteria to define left and right turns as previously (see Methods for details). We then asked whether the spontaneous turn direction during periods without sensory stimulation could be predicted from the same turn direction neuronal ensemble and coefficients previously identified in **Figure 4H-I** [↗](#). We found a significant correlation between the pre-motor turn direction predictions and the observed spontaneous tail curvature across larvae (**Figure 5B** [↗](#),  $r = 0.41$ ,  $p < 0.05$ ), suggesting this same ensemble is activated prior to both spontaneous and visually-evoked turns. This ensemble had a pre-motor turn direction

classification accuracy of  $70.2 \pm 6.0\%$  (mean  $\pm$  95% CI across  $n=5$  larvae, [Figure 5C](#)). Thus, these results demonstrate that a portion of the observed behavioral variability is overlapping with the larva's motor circuitry responsible for generation of spontaneous behavior.

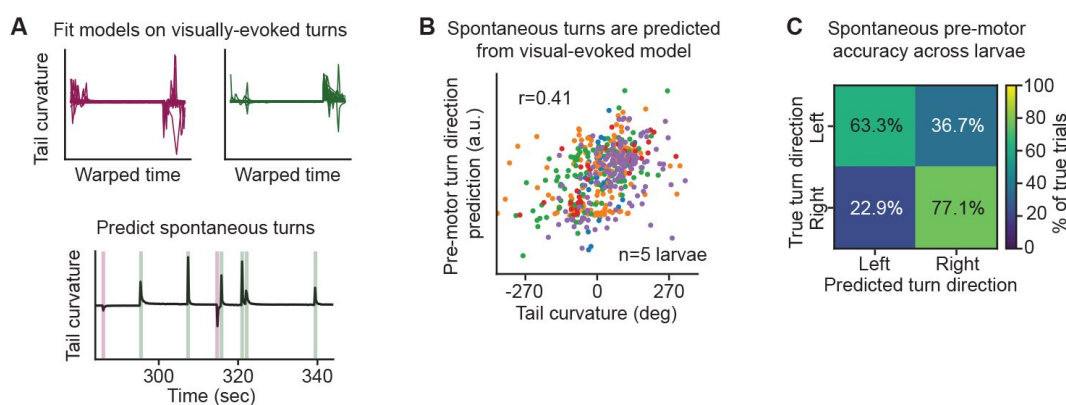
Our data highlight that the neural mechanisms involved in single-trial decision making are reflected in a highly distributed ensemble of neurons across the larval zebrafish brain. Further, behavioral variability, particularly in the larvae's responses to small-to-intermediate size stimuli, can be partially explained by an internal, task-independent turn bias that is unrelated to the visual stimulation, while their stimulus-driven responses to larger predator-like stimuli exhibited a comparatively lower level of behavioral variability. This suggests the importance of considering how internal biases and circuits governing spontaneous movements interact with the well-studied sensorimotor circuits to govern how the brain generates behavior in a naturalistic context.

## Discussion

In this study, we designed and built an optimized, high-resolution Fourier light field microscope (fLFM) to perform whole-brain imaging of the larval zebrafish in order to investigate the nature of trial-to-trial variability in neuronal dynamics and behavior on whole-brain yet cellular level. To do so, we studied visually-evoked turns within two ethologically relevant behaviors with opposing valance, prey capture and escape response, each of which are on average driven by distinct pre-motor neuronal ensembles ([Bollmann, 2019](#)) dependent on the size of the given sensory stimulus. Consistent with previous results ([Barker and Baier, 2015](#)), we found that the larvae's behavioral responses were highly variable across presentations of the same stimulus, with behavioral variability peaking in response to stimulus of an intermediate size between those which optimally drive attraction or avoidance behaviors. Given that we lack a mechanistic understanding of which specific neuronal circuits drive the observed variance across trials, we utilized fLFM to screen whole-brain dynamics at cellular resolution and identify neuronal ensembles which contribute to behavioral variability at the single-trial level.

We first asked whether the observed behavioral variability could be explained by noise or trial-to-trial deviations in encoding of the visual stimuli. We found that despite the variable responses of individual neurons to visual stimuli across trials, the population response faithfully and reliably encoded the visual stimulus in each trial. Further, we discovered that the visually-evoked neuronal activity patterns were nearly orthogonal to those dimensions that maximally covaried across stimulus repetitions, i.e. the noise modes, indicating that the source of behavioral variability was not unfaithful representation of the sensory inputs. Instead, we found that at least a third of the visually-evoked neurons contributed to noise modes that were correlated with various motor outputs. We ultimately identified two brain-wide neuronal populations which could predict the larvae's turn direction and responsiveness at the single-trial level. Surprisingly, a pre-stimulus bias component of these neuronal population activity could predict the larvae's trial-by-trial turn direction even before the onset of the stimulus presentation with an average accuracy of 77%, while after stimulus onset, a sharp ramping of the activity of this neuronal population approximately one second before movement initiation allowed for an increased turn direction prediction accuracy of 90%. Taken together, our data show that the larva's trial-by-trial decisions in response to any given stimulus are partially explained by pre-stimulus turn biases, and then ultimately driven by movement initiation dynamics which are broadcast within brain-wide neuronal ensembles.

In this context, the design of our behavioral paradigm has allowed us to gain insights into the nature of trial-to-trial neuronal and behavioral variability within and across two different ethologically relevant but highly divergent behaviors by tuning only a single parameter. This has allowed us to show that functionally different neuronal circuits have evolved to exhibit different



**Figure 5**

### Spontaneous turns are predictable from the same pre-motor neuronal population

**A.** Schematic of the approach to predict spontaneous turns. Models are fit to predict left or right turn direction during visual trials (top), as in [Figure 4](#). They are then tested on pre-motor periods one second before the spontaneous turn.

**B.** Spontaneous turns are predicted from visual-evoked model. Shown is the relationship between the spontaneous tail curvature and the predicted pre-motor turn direction using the visual-evoked model. Each dot is a single spontaneous turn and each color represents a different larva. They exhibit a significant correlation of  $r=0.41$ ,  $p<0.05$ .

**C.** Spontaneous turn classification accuracy. The mean cross-validated confusion matrix for spontaneous turn classification over  $n=5$  larvae. Spontaneous turns are predicted with an accuracy of  $70.2 \pm 6.0\%$  (mean  $\pm$  95% CI across  $n=5$  larvae).

levels of variability. Further, the differential interaction of such neuronal variability with the well-studied sensorimotor circuits may govern how the brain generates optimal behavior in different naturalistic contexts.

In addition, our study represents the first to our knowledge that identifies behavior-related noise correlations throughout an entire brain-wide sensorimotor decision-making circuit. The nearly orthogonal relationship between the stimulus decoding direction and trial-to-trial noise correlations is highly consistent with previous studies which found that noise correlations within single brain regions are organized so as to prevent additional information encoding channels within the same population from limiting sensory encoding capacity (Kafashan et al., 2021 [↗](#); Rummyantsev et al., 2020 [↗](#)). However, even a small degree of overlap between noise and sensory dimensions may decrease the information encoding capacity of the population (Moreno-Bote et al., 2014 [↗](#)). Why then might the brain favor a coding scheme in which populations mix information encoding with additional noise correlation structure?

Our data offer one possibility: that such additional correlations could encode internal or behavioral states that modulate sensorimotor transformations and behavioral output, enabling dynamic and flexible behavioral responses. Indeed, we found that a large fraction of the observed trial-to-trial neuronal variability was related to the larva's behavior: including activity correlated the instantaneous behavioral state of the animal as well as pre-stimulus turn biases that correlated with responsiveness and turn direction. We found that the largest trial-to-trial noise mode within the visually-evoked neurons was a subset of a brain-wide population correlated with the vigor of the larva's tail movements. This is conceptually similar to recent evidence in rodents that a large fraction of spontaneous neural activity reflects spontaneous and uninstructed movements (Musall et al., 2019 [↗](#); Stringer et al., 2019 [↗](#); Manley et al., 2024 [↗](#)).

Taken together, our study represents the first whole-brain confirmation that behaviorally relevant information is highly mixed throughout neuronal populations involved in processing sensory and other information, potentially enabling flexible and context-dependent behavior. Our observation of time-varying neuronal ensembles encoding a turn direction bias is reminiscent of the larval zebrafish's spontaneous swimming circuitry, which involves a central pattern generator (the ARTR) that biases the larva to swim in chains of turns in alternating directions (Dunn et al., 2016b [↗](#)); however, the neuronal populations we identified were not localized to the previously identified circuitry, and instead were distributed across the brain. While such spontaneous turn sequences are thought to underlie efficient exploration of local environments, our data could reflect a similar sensorimotor circuitry that injects variability as a learning mechanism to explore the space of possible sensorimotor consequences. Given the brain-wide distribution of the behavior-related ensembles within the variable visually-evoked behaviors studied here, we expect that they are not driven solely by a single central pattern generator as with ARTR and spontaneous turns, but instead involve the interaction of such pattern generators with the larva's internal state and recent experience. It has been shown that such hindbrain oscillatory activity integrates recent visual to enable efficient phototaxis (Wolf et al., 2017 [↗](#)), suggesting these pattern generators could act in concert with additional contextual populations. For example, it has been shown that hunger shifts decisions from avoidance to approach due to neuromodulation from the hypothalamic and serotonergic systems (Filosa et al., 2016 [↗](#)). Further, our observed dynamics could also reflect the foraging state of the larva, which is known to be encoded by an oscillatory, neuromodulatory network distributed across the brain (Marques et al., 2020 [↗](#)) and is a likely candidate to modulate a larva's response to prey-sized stimuli. We envision that utilization of the wide array of genetically labelled lines that have been associated with these sensorimotor and neuromodulatory circuits (Sammelhack et al., 2014 [↗](#); Barker and Baier, 2015 [↗](#); Marques et al., 2020 [↗](#); Okamoto et al., 2021 [↗](#)) could disentangle the specific contributions of the sensory, pre-motor, and internal state circuitry within the brain-wide populations observed here. Of particular interest will be how

the activity of these populations is organized over longer timescales and modulated by visual input or internal states related to hunger and foraging, which are critical to balance feeding behaviors and predator avoidance within the developing larvae.

## Methods

### fLFM setup

The sample is illuminated with an LED excitation light (470 nm, Thorlabs M470L4) via Köhler illumination through a standard GFP filter set (Thorlabs MDF-GFP) and a 20×/1.0-NA water immersion objective (Olympus XLUMPLFLN). In order to reduce the amount of excitation light reaching the eyes of the zebrafish, an aluminum mask is placed in the Köhler illumination path conjugate to the sample plane. The imaging path consisted of a f=180 mm Olympus-style tube lens (Thorlabs TTL180-A), a f=180 mm achromatic doublet (Thorlabs AC508-180-A-ML) for the Fourier lens, and a 13×13 mm microlens array with 1.5 mm pitch and r=13.9 mm radius of curvature (OKO Tech APO-GB-P1500-R13.9). The microlens array is mounted inside a five-axis kinematic mount (Thorlabs) to allow precise placement of the array orthogonal to the optical axis. A 5.5-megapixel sCMOS camera (Andor Zyla) is then positioned at the focus of the microlens array to capture the fLFM images.

The theoretical lateral resolution of an fLFM is limited by the numerical aperture  $NA_{ML}$  of the microlenses and the diffraction-limited spot size at the sensor. For an emission wavelength  $\lambda$ , Fourier lens focal length  $f_{FL}$ , objective magnification of  $M$ , and microlens diameter  $d_{MLA}$ , the Abbe diffraction limit when converted to object space and under the paraxial approximation is given by  $R_{xy} = \frac{\lambda f_{FL}}{d_{MLA} M}$ . The theoretical axial resolution for a microlens array with a radius  $d_{max}$  is given by  $R_z = \frac{d_{MLA} R_{xy}^2}{\lambda d_{max}}$  (Guo et al., 2019). Experimentally, we measured the lateral resolution of the fLFM as a function of distance from the native image plane (NIP) using a USAF target. The USAF target was positioned at various depths using an automated z-stage, and the slice of the reconstructed volume corresponding to that depth was analyzed. An element was considered resolved if the modulation transfer function (MTF) was greater than 30%.

### Zebrafish experiments

For head-fixed zebrafish experiments, n=13 *huc:h2b-gcamp6s* larvae with pan-neuronal and nuclear-localized GCaMP6s (NL-GCaMP6s) expression were imaged 6-9 days post fertilization (dpf). The day before experiments were performed, we immobilized larvae by embedding them in 2.5% agarose approximately 10 mm from the edge of a 35mm petri dish. We then removed the agarose around their eyes and tail to allow for visual stimulation and tail movement. Larvae were not fed after immobilization and were thus in a starved state on the day of the experiment.

The petri dish was covered in a rear projection film (Screen Solutions Int.) such that visual stimuli could be projected directly onto the dish. The projector (BenQ TH671ST) output was filtered with a 610 nm longpass filter (Schott RG610) such that only red light was displayed, removing crosstalk between the stimulation and GCaMP fluorescence. The tail was illuminated with an 850 nm LED and then monitored from below using a near infrared CMOS camera (Ximea MQ013RG-ON).

Visual stimulation was controlled by Stytra (Štíh et al., 2019). A short baseline period with no stimulation of 3 to 5 minutes was included at the start of each recording. To check for intact visual circuitry, drifting gratings with a period of 40° (visual angle) and speed of 40°/s to the left or right were shown three times every 30 minutes. Only larvae that displayed a robust optomotor response were further analyzed. The remaining trials consisted of drifting dots that began in the center of the visual field and moved to either the left or right at 30°/s. Dots were shown at maximum contrast against a dark background. Various diameters were shown from 0.3° to 44° visual angle,



with 3 to 6 different sizes used in a recording. Each trial lasted 3 seconds and the inter-trial interval was randomly distributed with a mean of 9 seconds and a standard deviation of 3 seconds. This inter-trial interval was chosen empirically to ensure that the visually-evoked activity from the previous trial was negligible during any given trial (see **Figure S3C** [↗](#)).

## Data acquisition

The fLFM camera was triggered at 5 or 10 Hz using a microcontroller (Adafruit Grand Central). The first trigger was used to initiate the tail behavior camera. The tail was monitored in real-time using Stytra ([Štíh et al., 2019](#) [↗](#)) at approximately 200 Hz. The recording duration lasted 1 to 2 hours.

## Data processing

Raw fLFM images were denoised using a custom-trained DeepInterpolation ([Lecoq et al., 2021](#) [↗](#)) model. DeepInterpolation is a self-supervised approach to denoising, which denoises the data by learning to predict a given frame from a set of frames before and after it. Time-varying signal can be distinguished from shot noise because shot noise is independent across frames, but signal is not. Therefore, only the signal is able to be predicted from adjacent frames. This has been shown to provide a highly effective and efficient denoising method ([Lecoq et al., 2021](#) [↗](#)). We trained a “unet\_single\_1024” model with 5 pre and post frames on a training set of  $n=20$  example recordings and validation set of  $n=5$  recordings. We ran the training until convergence of the validation loss. Denoised raw fLFM images were then reconstructed using the Richardson-Lucy deconvolution algorithm. We performed the deconvolution with an experimental PSF found by measuring the profile of a 1  $\mu\text{m}$  fluorescent bead through the axial depth in 2  $\mu\text{m}$  increments. The Richardson-Lucy algorithm took about 1 second per iteration on a GPU (NVIDIA TITAN V) and we used 12 to 20 iterations per frame. This resulted in a reconstructed volume of approximately  $760 \times 360 \times 280 \mu\text{m}^3$  with voxels of 2.4  $\mu\text{m}$  laterally and 4  $\mu\text{m}$  axially.

Neuronal ROIs and timeseries were then extracted using the CNMF-E ([Zhou et al., 2018](#) [↗](#)) variant of the CaImAn software package ([Giovannucci et al., 2019](#) [↗](#)), which is optimized for processing of one-photon imaging data. Each plane was analyzed individually and then the results were merged. To determine neuronal ROIs, the SNR threshold was set at 2, the spatial consistency (rval) threshold was set at 0.8, and the CNN-based classifier threshold was set at 0.9. Neuronal timeseries were not deconvolved. Background fluorescence was factored out using the CNMF-E ring model. The planes were then collated by merging any units with a correlation greater than 0.7, lateral distance less than 10  $\mu\text{m}$ , and axial distance less than 16  $\mu\text{m}$ . Overlapping units were merged by summing each timeseries weighted by its SNR. Effects of photobleaching were removed by normalizing by a fit to the mean signal  $f(t)$  across all neurons over the recording with a bi-exponential model  $\hat{f}(t) = ae^{bt} + ce^{dt}$ . Each neuronal timeseries was then z-scored before all following analyses.

For localization of neuronal ensembles to specific brain regions, each recording was aligned to the Z-Brain atlas ([Randlett et al., 2015](#) [↗](#)). Alignment was performed by manually marking a number of keypoints on the reference image and the standard deviation projection over time of the 3D fLFM recording, and then estimating the affine transformation between the reference and fLFM volume via least squares.

## Analysis of freely swimming behavior

For freely swimming experiments,  $n=9$  6-9 dpf *huc:h2b-gcamp6s* larvae were individually placed in a 90 mm petri dish with a shallow (~3 mm) amount of water. The dish was placed above a display (Apple iPad) with a 6 mm piece of clear acrylic in between. Black dots of sizes ranging from 1° to

30° were randomly shown at maximum contrast and moving across the dish in one direction at 30°/s. A camera (FLIR Grasshopper3 GS3-U3-41C6M-C) with a zoom lens (Navitar MVL7000) was mounted above to monitor the larvae's behavior at >100 Hz.

Each larva's movements were tracked using DeepLabCut (Mathis et al., 2018 [↗](#)), which was trained to track the position of the eyes, the swim bladder, and four points along the spline of the tail. Anipose was used to build a pipeline to process large numbers of videos (Karashchuk et al., 2020 [↗](#)). Movement bouts were extracted by applying a threshold to the velocity of the centroid of the tracked points. Only movement bouts in which the centroid was within approximately six body lengths of the visual stimulus were further analyzed. For each bout, we then calculated the mean change in distance to the visual stimulus, with positive values denoting the larvae moving further away from the visual stimulus.

## Analysis of head-fixed behavior

During head-fixed visual experiments, 20 points along the spline of the larva's tail were tracked in real time using Stytra (Štíh et al., 2019 [↗](#)). The overall tail curvature was then quantified by summing the angles between each segment of the tail tracking. For comparison with neural data, this was convolved with the NL-GCaMP6s kernel to define the tail direction kernel, while the GCaMP-kernel-convolved absolute tail curvature defined the tail vigor kernel. The angular velocity and angular acceleration were computed for the overall tail curvature summed over all segments. The NL-GCaMP6s kernel was estimated empirically by aligning and averaging a number of calcium events. This kernel corresponds to a half-rise time of 400 ms and half-decay time of 4910 ms. The use of a fixed calcium kernel does not account for any variability in the GCaMP response across cells, which could be due to differences such as cell type or expression level. Therefore, this analysis approach may not capture the full set of neurons which exhibit stimulus correlations but exhibit a different GCaMP response.

Bouts of movement were extracted by thresholding the absolute tail curvature with a value  $\sigma_{active}$  set to one standard deviation of the absolute tail curvature over time. To study the behavioral response to various stimuli, the mean tail curvature during active bouts was computed for each stimulus presentation period; trials in which there were no movement bouts were excluded. For turn direction analyses, trials with a mean tail curvature during bouts  $> \sigma_{active}$  were considered as left turn trials, whereas trials with a mean tail curvature during bouts  $< -\sigma_{active}$  were considered right turn trials.

## Visual stimulus decoding

For the decoding shown in **Figure 3B** [↗](#), the stimulus-evoked neural response was taken as the average response of each neuron during a single stimulus presentation. For each larva, a logistic regression model with lasso regularization was trained on all trials using 6-fold cross-validation with 3 repetitions. The lasso regularization parameter, which induces sparsity, was swept and the model with the highest average cross-validated accuracy was used. The confusion matrix, computed on the held-out testing trials, was reported as the average over all rounds of cross-validation.

## Identification of stimulus decoding and noise dimensions

The comparison of the neural dimensions encoding visual stimuli versus trial-to-trial noise was modeled after Rummyantsev et al. (2020) [↗](#). Partial least squares (PLS) regression was used to find a low-dimensional space that optimally predicted the visual stimuli, which we refer to as the visually-evoked neuronal activity patterns. To perform regression, a visual stimulus kernel was constructed by summing the timeseries of each individual stimulus type, weighted by the stimulus

size and negated for trials on the right visual field, thus providing a single response variable  $\bar{y}$  encoding both the location, size, and timing of all the stimulus presentations. This stimulus kernel was the convolved with the temporal response kernel of our calcium indicator (NL-GCaMP6s).

PLS regression identifies the normalized dimensions  $\bar{p}_i$  and  $\bar{q}_j$  that maximize the covariance between paired observations  $\bar{x}_i$  and  $\bar{y}_j$ , respectively. In our case, the visual stimulus is represented by a single variable  $\bar{y}$ , simplifying the problem to identifying the subspace of neural activity that optimally preserves information about the univariate visual stimulus (sometimes referred to as PLS1 regression). That is, the  $N \times T$  neural time series matrix  $X$  is reduced to a  $d \times T$  matrix spanned by a set of orthonormal vectors. PLS1 regression is performed as follows:

## PLS1 algorithm

Let  $X_i = X$  and  $y_i = \bar{y}$ . For  $i = 1 \dots d$ ,

1.  $p_i = X_i^T y_i / |X_i^T y_i|$
2.  $t_i = X_i p_i$
3.  $q_i = X_i^T y_i (t_i^T t_i)^{-1}$
4.  $X_{i+1} = X_i - t_i q_i^T$
5.  $c = y_i^T t_i (t_i^T t_i)^{-1}$  (note this is scalar)
6.  $y_{i+1} = y_i - c t_i$

The projections of the neural data  $\{p_i\}$  thus span a subspace that maximally preserves information about the visual stimulus  $\bar{y}$ . Stacking these projections into the  $N \times d$  matrix  $P$  that represents the transform from the whole-brain neural state space to the visually-evoked subspace, the optimal decoding direction is given by the linear least squares solution  $\bar{w}_{\text{opt}} = P(P^T X^T X P)^{-1} P^T X^T \bar{y}$ . The dimensionality  $d$  of PLS regression was optimized using 6-fold cross-validation with 3 repeats and choosing the dimensionality between  $d = 1$  and 20 with the lowest cross-validated mean squared error for each larva. Then,  $\bar{w}_{\text{opt}}$  was computed using all time points.

For each stimulus type, the noise covariance matrix  $C$  was computed in the low-dimensional PLS space, given that direct estimation of the noise covariances across many thousands of neurons would likely be unreliable. A noise covariance matrix was calculated separately for each stimulus, and then averaged across all stimuli. As before, the mean activity  $\mu_i$  for each neuron  $i$  was computed over each stimulus presentation period. The noise covariance then describes the correlated fluctuations  $\delta_i$  around this mean response for each pair of neurons  $i$  and  $j$ , where  $C_{ij} = \langle \delta_i \delta_j \rangle$ .

The noise modes  $\bar{e}_\alpha$  for  $\alpha = 1$  were subsequently identified by eigendecomposition of the mean noise covariance matrix across all stimuli,  $C = \sum \lambda_\alpha \bar{e}_\alpha \bar{e}_\alpha^T$ . The angle between the optimal stimulus decoding direction and the noise modes is thus given by  $\cos \theta_\alpha = (\bar{w}_{\text{opt}} \cdot \bar{e}_\alpha) / (|\bar{w}_{\text{opt}}| |\bar{e}_\alpha|)$

The significance of each neurons' contribution to a given neural population vector was evaluated by comparison to shuffled datasets. To created shuffled datasets, the ground truth stimulus presentation labels were randomly permuted. The same analysis above was performed on  $n=10$  shuffled datasets, to identify a distribution of shuffled coefficients for the optimal stimulus decoding direction, noise modes, and largest principal component. A neuron was then considered to be significantly contributing to a given population vector if its absolute coefficient was greater than three standard deviations above the mean of that of the shuffled datasets.

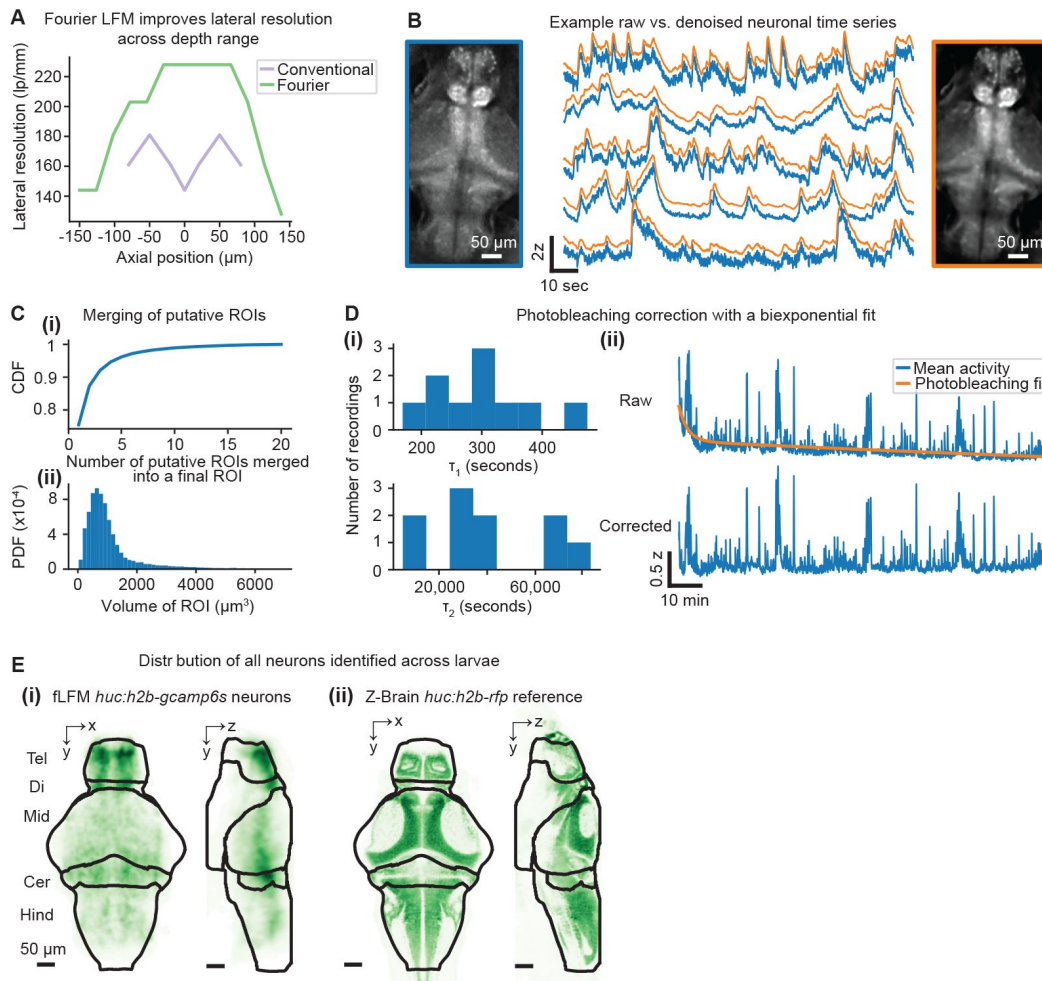
## Classification of single trial turn direction, responsiveness, and stimulus

To predict decisions from neural activity despite the variable reaction times, each trial was time-warped in order to align the stimulus onset and the movement initiation timepoints, following (Lin et al., 2020 [DOI](#)). Left and right turn trials were extracted as described previously. Response trials included both left and right turn trials (i.e., the absolute value of mean tail curvature  $> \sigma_{active}$ ), whereas nonresponse trials were motionless (absolute mean tail curvature  $< \sigma_{active}$ ). In particular, forward-motion trials were excluded from these analyses. For no response trials, a random reaction time was selected from the responsive trials and the corresponding time-warping was applied. Logistic regression binary classification models with lasso regularization were fit to classify left versus right turn trials or response versus no response trials from various time windows of the time-warped neural activity. Thus, time-warped neuronal activity was averaged over windows from various start and end timepoints to determine the predictability at each time window. The optimal predictability at any given timepoint was then considered as the best time window that ended at that timepoint. For each larvae, 6-fold cross-validation with 3 repetitions was utilized to determine the cross-validated F score, accuracy, and confusion matrix. For stimulus classification, logistic regression was similarly applied for multi-class classification using the one-versus-rest scheme. Overall classification performance was quantified using the mean F score across larvae and was compared to shuffled data in which the class labels had been randomly permuted.

Visualizations of single trial trajectories were made by projecting the time-warped trials onto the weights of an example, randomly chosen logistic regression model.

Significant neurons contributing to turn direction and responsiveness classification were identified by comparing a neuron's absolute coefficient across the multiple cross-validation rounds to that from models fit to shuffled class labels. Neurons were deemed to be significantly contributing using a one-tailed t-test between each neuron's coefficients and its coefficients estimated from shuffled data across the cross-validation rounds,  $p < 0.05$ . Similarly, neurons were considered to be significantly correlated with the turn direction bias projection and the responsiveness projection in **Figures S5F-G** [DOI](#) using a one-tailed t-test between each neuron's correlations and its correlations estimated from shuffled data across the cross-validation rounds.

## Supplementary figures



## Supplementary Figure S1

related to Figure 1 [↗](#), Fourier Light Field Microscopy (fLFM) provides a simple and cost-effective method for whole-brain imaging of larval zebrafish during behavior

**A.** fLFM versus cLFM lateral resolution across the axial imaging range. fLFM (green) exhibits higher lateral resolution and an expanded axial imaging range compared to cLFM (purple).

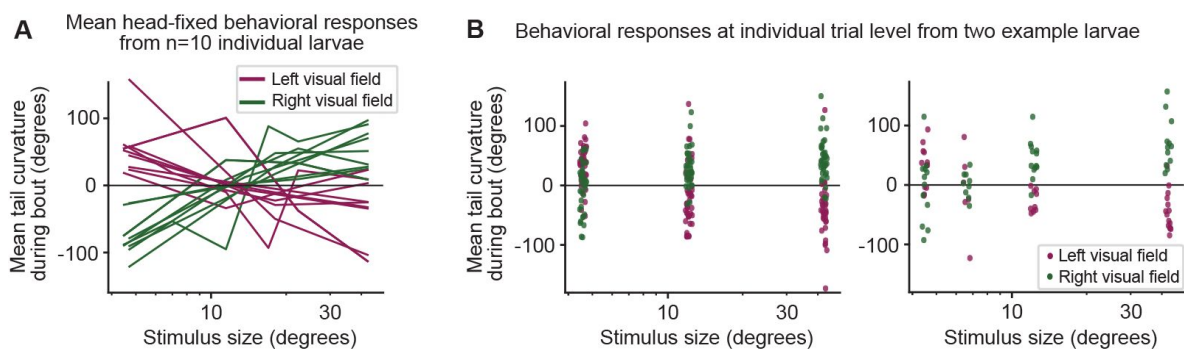
**B.** Examples of raw and denoised neuronal time series. Left: an example reconstructed plane from a raw fLFM image. Blue circle indicates the ROI whose activity is plotted in the middle. Middle: Raw (blue) and denoised (orange) time series of average fluorescence intensity from neurons detected in the plane shown in the images. Right: the corresponding reconstructed plane as on the left, but after denoising the fLFM image with DeepInterpolation.

**C.** Statistics of merging putative ROIs with high correlation and low spatial distance (see Methods for details) across  $n=10$  larvae. (i) The cumulative distribution function (CDF) of the number of putative ROIs which were merged to create a final neuronal ROI. (ii) The probability density function (PDF) of the volume of the final neuronal ROIs.

**D.** The effect of photobleaching on the fluorescence intensity was corrected with a biexponential fit. (i) The distribution of time constants resulting from biexponential fits to the mean neuronal signal for  $n=10$  larvae. Two time constants are fit for each recording, one on the order of minutes (top) and the other tending to be on the order of many hours (bottom). (ii) Examples of raw (top) and corrected (bottom) mean neuronal signals from one recording. The biexponential fit to the data is shown on the top in orange.

**E.** The distribution of neuronal ROIs across  $n=10$  larvae resembles the labeling from the Z-Brain reference image. (i) The distribution of identified neurons (labeled with *huc:h2b-gcamp6s*) across all recordings shown aligned to the Z-Brain atlas. Black borders mark the boundaries of the telencephalon (Tel), diencephalon (Di), midbrain (Mid), cerebellum (Cer), and the remaining hindbrain (Hind). Shown are the xy and yz projections of the kernel density estimate of the center of masses of each ROI. (ii) One xy and yz slice from the *huc:h2b-rfp* reference image to which each fLFM recording was aligned. Scale bars represent 50  $\mu\text{m}$ .



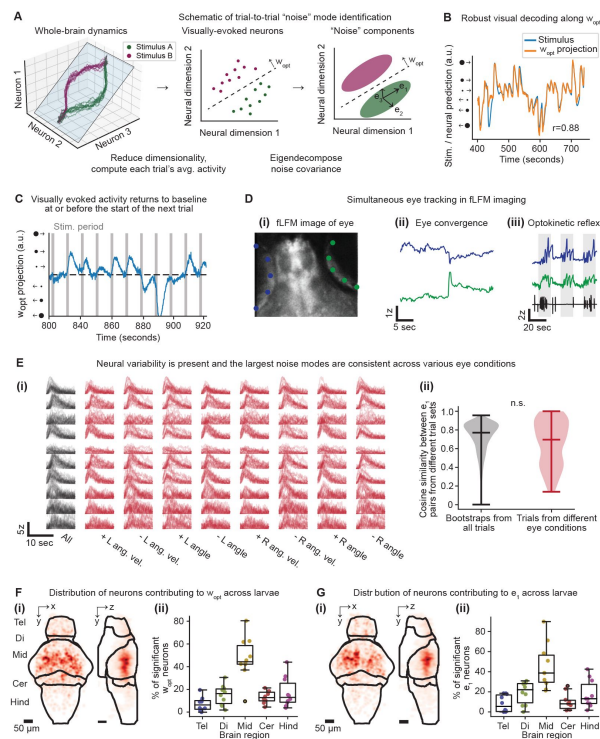


### Supplementary Figure S2

related to [Figure 2](#), zebrafish exhibit highly variable motor responses to visual stimuli

**A.** Inter-individual variability in the motor responses demonstrated by the mean head-fixed behavioral response from n=10 larvae, as shown in [Figure 2F](#) but at the individual larva level.

**B.** Intra-individual variability in the motor responses shown for two example larvae (left and right), similar to [Figure 2G](#). Each point represents the mean tail curvature during a single trial.



## Supplementary Figure S3

### related to Figure 3, trial-to-trial variability in visually-evoked neurons is largely orthogonal to visual decoding dimensions

**A.** Schematic of trial-to-trial noise mode identification. First, dimensionality reduction is applied to the whole-brain dynamics (left) to identify the neuronal activity patterns that optimally preserve visual information using partial least squares regression. Within these visually-evoked neurons, each trial's average activity during the visual presentation is computed (middle, each dot represents a single trial corresponding to the pink or green stimulus condition). Finally, the covariance across trials is decomposed to find the noise modes  $\vec{e}_\alpha$ , which represent orthogonal directions within the visually-evoked neurons and ranked in order of decreasing variance.

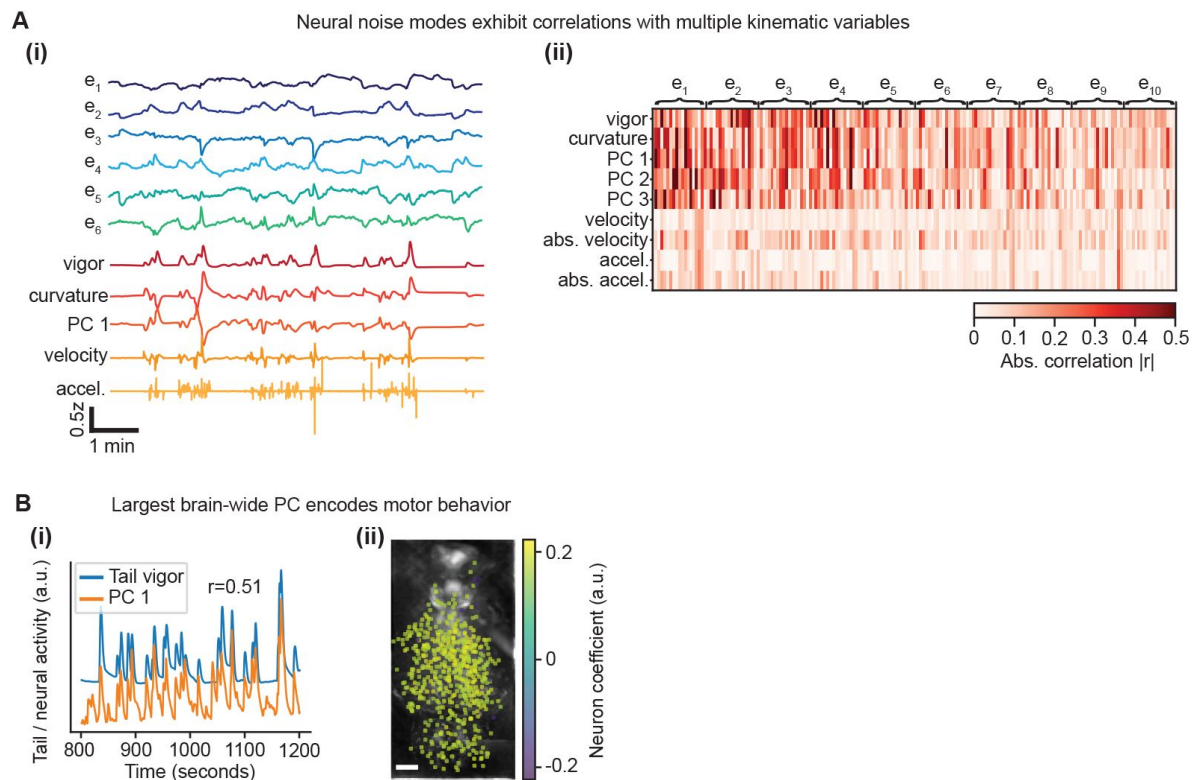
**B.**  $\vec{w}_{\text{opt}}$  robustly encodes the visual stimulus. Shown are example time traces of the stimulus kernel (blue, where the size of the stimulus is encoded by the magnitude and which visual field it is presented in is encoded by the sign) and the neural projection onto  $\vec{w}_{\text{opt}}$  (orange), which exhibit a correlation of  $r=0.88$ ,  $p<10^{-6}$ .

**C.** Visually-evoked activity returns to baseline at or before the start of the following stimulus presentation. The projection along the optimal stimulus decoding direction  $\vec{w}_{\text{opt}}$ , which demonstrates the dynamics of the visually-evoked population, is shown relative to the duration of the visual stimulus presentations (gray shaded bars).

**D.** Tracking of various eye movements in the raw fLFM images. (i) An example raw fLFM image of an eye, with keypoints tracked in DeepLabCut shown for the left (blue) and right (green) eyes. (ii) An example eye convergence event. (iii) An example of the optokinetic reflex. Saccades, in which both eyes exhibit motion in the same direction), are shown as the larva tracks a moving grating stimulus (gray shaded regions). The tail curvature is shown below in black, demonstrating the optomotor reflex as well.

**E.** Neural variability is consistent across various states of the eye and eye movements. (i) For one visual stimulus type, the activity of 10 neurons (rows) is shown during various subsets of trials (columns). While all of the trials for this stimulus are shown on the left in black, the remaining red traces exhibit various conditions for each eye (L or R): high (+) or low (–) angular velocity (ang. vel.), as well as larger, leftward looking eye angles (+ angle) or smaller, rightward looking eye angles (–angle). (ii) Comparison of the dominant noise mode estimates across trial subsets. A control distribution is shown on the left in black by performing bootstrapping with replacement to randomly select a subset of trials and then estimating  $\vec{e}_1$ . On the right in red, noise modes are compared from the different eye conditions in (i). The noise mode estimates are compared using the cosine similarity, where 1 indicates perfect alignment whereas 0 indicates orthogonality. The two distributions were not significantly different ( $p>0.05$ , Wilcoxon rank-sum test).

**F-G.** Distribution of neurons contributing to  $\vec{w}_{\text{opt}}$  (F) or  $\vec{e}_1$  (G) across  $n=10$  larvae. (i) The 3D distribution of significantly contributing neurons, shown via xy and yz projections relative to the Z-Brain atlas. (ii) Quantification of the percent of contributing neurons contained within each of the brain regions shown in (i).

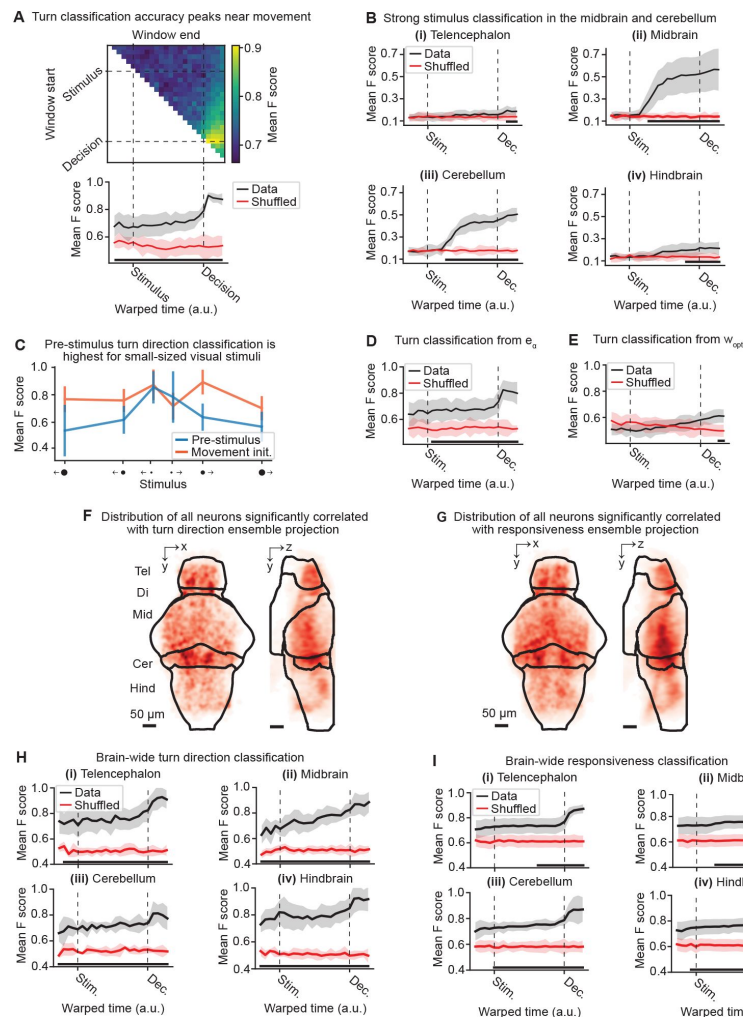


### Supplementary Figure S4

#### related to Figure 3, trial-to-trial variability in visually-evoked neurons is largely orthogonal to visual decoding dimensions

**A.** Multiple neural noise modes shown correlation with motor outputs. (i) Example timeseries of six neural noise modes  $\vec{e}_\alpha$  and various kinematic variables: tail vigor, tail curvature, the first principal component (PC 1) of the tail angles, tail angular velocity, and tail angular acceleration. Each of the kinematic variables is convolved with the NL-GCaMP6s response kernel. (ii) Absolute correlation matrix between the neural noise mode and kinematic timeseries for  $n=10$  larvae. Each row represents one motor variable. The columns are grouped by noise mode  $\vec{e}_\alpha$ . Within a noise mode group, each individual column represents one larva.

**B.** The largest brain-wide principal component (PC) encodes motor behavior. (i) Example time traces of the tail vigor (blue) and PC 1 (orange), which exhibit a correlation of  $r=0.51$ ,  $p<10^{-6}$ . (ii) The coefficients of the neurons contributing most to the PC, which are distributed across the brain and are nearly all positively correlated with tail vigor.



## Supplementary Figure S5

### related to Figure 4, pre-motor neuronal populations predictive of single-trial behavior

**A.** Classification of turn direction as a function of time window analyzed. The mean F score across  $n=7$  larvae is used to assess the performance of multi-class classification of the visual stimulus. Top: a time window is swept across various start timepoints (y-axis) and end timepoints (x-axis) to determine the classification performance across various time intervals of the neuronal dynamics. Bottom: Mean  $\pm$  95% confidence interval of the F score for the best time interval ending at the given timepoint (Data), compared to shuffled data in which the stimulus labels are randomized. The classification accuracy is significantly higher than shuffled data ( $p < 0.05$ , paired t-test) only after stimulus onset.

**B.** Stimulus classification from various brain regions. As in Figure 4B, except only performing classification using neurons in a single brain region. Strong decoding is only observed in the midbrain and cerebellum after stimulus onset.

**C.** Turn direction classification as a function of stimulus size. On average before stimulus onset (blue line), the F score is highest for the smallest size stimuli, whereas the F score is relatively constant at the time of movement initiation (orange line, mean  $\pm$  95% CI across  $n=5$  larvae).

**D.** As in A, except for classifying left or right turn direction from the projections onto the trial-to-trial noise modes  $\vec{e}_\alpha$ , as opposed to the whole-brain dynamics as used in Figure 4B.

**E.** As in A, except for classifying left or right turn direction from the projection onto the optimal stimulus decoding dimension  $\vec{w}_{opt}$ .

**F-G.** Distribution of all neurons correlated with the activity of the turn direction bias ensemble (F) or responsiveness ensemble (G). Significant correlations were determined by comparing the observed correlation to that of shuffled datasets where the neuron's timeseries was randomly circularly permuted.

**H-I.** Turn direction (H) and responsiveness (I) exhibit similar predictability across brain regions. As in panels 4D and 4C, respectively, expect only performing classification using neurons in a single brain region.

## Acknowledgements

We thank Q. Lin and S. Otero Coronel for discussions and feedback on the manuscript, T. Nöbauer and Y. Zhang for discussions regarding optical designs, and the members of the Vaziri Lab for discussions regarding the experiments and data analysis. We also thank A. Kaczynska, S. Campbell, and J. Hudspeth for housing zebrafish, and the Rockefeller University High Performance Computing Cluster for compute access. This work was supported in part by the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under award numbers 1RF1NS113251 and 1RF1NS110501 and the Kavli Foundation through the Kavli Neural Systems Institute.

## Additional information

### Author contributions

J.M. contributed to the project conceptualization, designed and built the imaging system, performed experiments, analyzed data, and wrote the manuscript. A.V. conceived, led and supervised the project, designed the imaging system and biological experiments, guided data analysis, and wrote the manuscript.

## Additional files

**Video 1.** Example target-directed and target-avoidance bouts in free behavior (related to Figure 2). On the left, an example target-directed bout showing a 7dpf larval zebrafish approaching a stimulus presented on a screen below the dish (see Figure 2A). On the right, a similar target-avoidance bout from the same larva. The videos are slowed down 10x. These videos correspond to the composite images shown in Figure 2B-C. [↗](#)



## References

- Afshar A, Santhanam G, Yu BM, Ryu SI, Sahani M, Shenoy KV (2011) **Single-Trial Neural Correlates of Arm Movement Preparation** *Neuron* **71**:555–564 <https://doi.org/10.1016/j.neuron.2011.05.047>Google Scholar
- Agard DA (1984) **Optical Sectioning Microscopy: Cellular Architecture in Three Dimensions** *Annu Rev Bioph Biom* **13**:191–219 <https://doi.org/10.1146/annurev.bb.13.060184.001203>Google Scholar
- Aimon S, Katsuki T, Jia T, Grosenick L, Broxton M, Deisseroth K, Sejnowski TJ, Greenspan RJ (2019) **Fast near-whole-brain imaging in adult *Drosophila* during responses to stimuli and behavior** *Plos Biol* **17**:e2006732 <https://doi.org/10.1371/journal.pbio.2006732>Google Scholar
- Barker AJ, Baier H. (2015) **Sensorimotor Decision Making in the Zebrafish Tectum** *Curr Biol* **25**:2804–2814 <https://doi.org/10.1016/j.cub.2015.09.055>Google Scholar
- Bartolo R, Saunders RC, Mitz AR, Averbeck BB (2020) **Information-Limiting Correlations in Large Neural Populations** *J Neurosci* **40**:1668–1678 <https://doi.org/10.1523/jneurosci.2072-19.2019>Google Scholar
- Bianco IH, Engert F. (2015) **Visuomotor Transformations Underlying Hunting Behavior in Zebrafish** *Curr Biol* **25**:831–846 <https://doi.org/10.1016/j.cub.2015.01.042>Google Scholar
- Bianco IH, Kampff AR, Engert F. (2011) **Prey Capture Behavior Evoked by Simple Visual Stimuli in Larval Zebrafish** *Frontiers Syst Neurosci* **5**:101 <https://doi.org/10.3389/fnsys.2011.00101>Google Scholar
- Bollmann JH (2019) **The Zebrafish Visual System: From Circuits to Behavior** *Annu Rev Vis Sc* **5**:269–293 <https://doi.org/10.1146/annurev-vision-091718-014723>Google Scholar
- Borla MA, Palecek B, Budick S, O'Malley DM (2002) **Prey Capture by Larval Zebrafish: Evidence for Fine Axial Motor Control** *Brain Behav Evol* **60**:207–229 <https://doi.org/10.1159/000066699>Google Scholar
- Broxton M, Grosenick L, Yang S, Cohen N, Andalman A, Deisseroth K, Levoy M. (2013) **Wave optics theory and 3-D deconvolution for the light field microscope** *Opt Express* **21**:25418–39 <https://doi.org/10.1364/oe.21.025418>Google Scholar
- Cafaro J, Rieke F (2010) **Noise correlations improve response fidelity and stimulus encoding** *Nature* **468**:964–967 <https://doi.org/10.1038/nature09570>Google Scholar
- Chen X, Mu Y, Hu Y, Kuan AT, Nikitchenko M, Randlett O, Chen AB, Gavornik JP, Sompolinsky H, Engert F, Ahrens MB (2018) **Brain-wide Organization of Neuronal Activity and Convergent Sensorimotor Transformations in Larval Zebrafish** *Neuron* **100**:876–890 <https://doi.org/10.1016/j.neuron.2018.09.042>Google Scholar

- Cohen MR, Kohn A. (2011) **Measuring and interpreting neuronal correlations** *Nat Neurosci* **14**:811–819 <https://doi.org/10.1038/nn.2842>Google Scholar
- Cohen MR, Maunsell JHR (2009) **Attention improves performance primarily by reducing interneuronal correlations** *Nat Neurosci* **12**:1594–1600 <https://doi.org/10.1038/nn.2439>Google Scholar
- Cohen N, Yang S, Andalman A, Broxton M, Grosenick L, Deisseroth K, Horowitz M, Levoy M. (2014) **Enhancing the performance of the light field microscope using wavefront coding** *Opt Express* **22**:24817–39 <https://doi.org/10.1364/oe.22.024817>Google Scholar
- Cong L, Wang Z, Chai Y, Hang W, Shang C, Yang W, Bai L, Du J, Wang K, Wen Q. (2017) **Rapid whole brain imaging of neural activity in freely behaving larval zebrafish (*Danio rerio*)** *eLife* **6**:e28158 <https://doi.org/10.7554/elife.28158>Google Scholar
- Dunn TW, Gebhardt C, Naumann EA, Riegler C, Ahrens MB, Engert F, Del Bene F. (2016a) **Neural Circuits Underlying Visually Evoked Escapes in Larval Zebrafish** *Neuron* **89**:613–628 <https://doi.org/10.1016/j.neuron.2015.12.021>Google Scholar
- Dunn TW, Mu Y, Narayan S, Randlett O, Naumann EA, Yang C-T, Schier AF, Freeman J, Engert F, Ahrens MB (2016b) **Brain-wide mapping of neural activity controlling zebrafish exploratory locomotion** *eLife* **5**:e12741 <https://doi.org/10.7554/elife.12741>Google Scholar
- Faisal AA, Selen LPJ, Wolpert DM (2008) **Noise in the nervous system** *Nat Rev Neurosci* **9**:292–303 <https://doi.org/10.1038/nrn2258>Google Scholar
- Filosa A, Barker AJ, Dal Maschio M, Baier H. (2016) **Feeding State Modulates Behavioral Choice and Processing of Prey Stimuli in the Zebrafish Tectum** *Neuron* **90**:596–608 <https://doi.org/10.1016/j.neuron.2016.03.014>Google Scholar
- Flavell SW, Gogolla N, Lovett-Barron M, Zelikowsky M. (2022) **The emergence and influence of internal states** *Neuron* **110**:2545–2570 <https://doi.org/10.1016/j.neuron.2022.04.030>Google Scholar
- Flavell SW, Pokala N, Macosko EZ, Albrecht DR, Larsch J, Bargmann CI (2013) **Serotonin and the Neuropeptide PDF Initiate and Extend Opposing Behavioral States in *C. elegans*** *Cell* **154**:1023–1035 <https://doi.org/10.1016/j.cell.2013.08.001>Google Scholar
- Förster D, Helmbrecht TO, Mearns DS, Jordan L, Mokayes N, Baier H. (2020) **Retinotectal circuitry of larval zebrafish is adapted to detection and pursuit of prey** *eLife* **9**:e58596 <https://doi.org/10.7554/elife.58596>Google Scholar
- Giovannucci A, Friedrich J, Gunn P, Kalfon J, Brown BL, Koay SA, Taxidis J, Najafi F, Gauthier JL, Zhou P, Khakh BS, Tank DW, Chklovskii DB, Pnevmatikakis EA (2019) **CaImAn an open source tool for scalable calcium imaging data analysis** *eLife* **8**:e38173 <https://doi.org/10.7554/elife.38173>Google Scholar
- Guo C, Liu W, Hua X, Li H, Jia S. (2019) **Fourier light-field microscopy** *Opt Express* **27**:25573 <https://doi.org/10.1364/oe.27.025573>Google Scholar

- Huang X, Lisberger SG (2009) **Noise correlations in cortical area MT and their potential impact on trial-by-trial variation in the direction and speed of smooth-pursuit eye movements** *J Neurophysiol* **101**:3012–30 <https://doi.org/10.1152/jn.00010.2009>Google Scholar
- Kafashan M, Jaffe AW, Chettih SN, Nogueira R, Arandia-Romero I, Harvey CD, Moreno-Bote R, Drugowitsch J. (2021) **Scaling of sensory information in large neural populations shows signatures of information-limiting correlations** *Nat Commun* **12**:473 <https://doi.org/10.1038/s41467-020-20722-y>Google Scholar
- Kanitscheider I, Coen-Cagli R, Pouget A. (2015) **Origin of information-limiting noise correlations** *Proc National Acad Sci* **112**:E6973–E6982 <https://doi.org/10.1073/pnas.1508738112>Google Scholar
- Kao JC, Nuyujukian P, Ryu SI, Churchland MM, Cunningham JP, Shenoy KV (2015) **Single-trial dynamics of motor cortex and their applications to brain-machine interfaces** *Nat Commun* **6**:7759 <https://doi.org/10.1038/ncomms8759>Google Scholar
- Karashchuk P, Rupp KL, Dickinson ES, Walling-Bell S, Sanders E, Azim E, Brunton BW, Tuthill JC (2020) **Anipose: A toolkit for robust markerless 3D pose estimation** *Cell Reports* **36**:109730 <https://doi.org/10.1016/j.celrep.2021.109730>Google Scholar
- Kaufman MT, Churchland MM, Ryu SI, Shenoy KV (2015) **Vacillation, indecision and hesitation in moment-by-moment decoding of monkey motor cortex** *eLife* **4**:e04677 <https://doi.org/10.7554/elife.04677>Google Scholar
- Kauvar IV, Machado TA, Yuen E, Kochalka J, Choi M, Allen WE, Wetzstein G, Deisseroth K. (2020) **Cortical Observation by Synchronous Multifocal Optical Sampling Reveals Widespread Population Encoding of Actions** *Neuron* **107**:351–367 <https://doi.org/10.1016/j.neuron.2020.04.023>Google Scholar
- Kim TH, Schnitzer MJ (2022) **Fluorescence imaging of large-scale neural ensemble dynamics** *Cell* **185**:9–41 <https://doi.org/10.1016/j.cell.2021.12.007>Google Scholar
- Kohn A, Coen-Cagli R, Kanitscheider I, Pouget A. (2015) **Correlations and Neuronal Population Information** *Annu Rev Neurosci* **39**:1–20 <https://doi.org/10.1146/annurev-neuro-070815-013851>Google Scholar
- Lecoq J, Oliver M, Siegle JH, Orlova N, Ledochowitsch P, Koch C. (2021) **Removing independent noise in systems neuroscience data using DeepInterpolation** *Nat Methods* **18**:1401–1408 <https://doi.org/10.1038/s41592-021-01285-2>Google Scholar
- Levoy M, Ng R, Adams A, Footer M, Horowitz M. (2006) **Light field microscopy** *Acm T Graphic* **25**:924 <https://doi.org/10.1145/1141911.1141976>Google Scholar
- Lin Q, Manley J, Helmreich M, Schlumm F, Li JM, Robson DN, Engert F, Schier A, Nöbauer T, Vaziri A. (2020) **Cerebellar Neurodynamics Predict Decision Timing and Outcome on the Single-Trial Level** *Cell* **180**:536–551 <https://doi.org/10.1016/j.cell.2019.12.018>Google Scholar
- Liu FL, Kuo G, Antipa N, Yanny K, Waller L. (2020) **Fourier DiffuserScope: Single-shot 3D Fourier light field microscopy with a diffuser** *Arxiv* <https://doi.org/10.1364/oe.400876>Google Scholar
- Llavador A, Sola-Pikabea J, Saavedra G, Javidi B, Martínez-Corral M. (2016) **Resolution improvements in integral microscopy with Fourier plane recording** *Opt Express* **24**:20792–

8 <https://doi.org/10.1364/oe.24.020792>Google Scholar

Machado TA, Kauvar IV, Deisseroth K. (2022) **Multiregion neuronal activity: the forest and the trees** *Nat Rev Neurosci* **23**:683–704 <https://doi.org/10.1038/s41583-022-00634-0>Google Scholar

Manley J, Lu S, Barber K, Demas J, Kim H, Meyer D, Traub FM, Vaziri A. (2024) **Simultaneous, cortex-wide dynamics of up to 1 million neurons reveal unbounded scaling of dimensionality with neuron number** *Neuron* **112**:1694–1709 <https://doi.org/10.1016/j.neuron.2024.02.011>Google Scholar

Marques JC, Li M, Schaak D, Robson DN, Li JM (2020) **Internal state dynamics shape brainwide activity and foraging behaviour** *Nature* **577**:239–243 <https://doi.org/10.1038/s41586-019-1858-z>Google Scholar

Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M. (2018) **DeepLabCut: markerless pose estimation of user-defined body parts with deep learning** *Nat Neurosci* **21**:1281–1289 <https://doi.org/10.1038/s41593-018-0209-y>Google Scholar

Montijn JS, Meijer GT, Lansink CS, Pennartz CMA (2016) **Population-Level Neural Codes Are Robust to Single-Neuron Variability from a Multidimensional Coding Perspective** *Cell Reports* **16**:2486–2498 <https://doi.org/10.1016/j.celrep.2016.07.065>Google Scholar

Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A. (2014) **Information-limiting correlations** *Nat Neurosci* **17**:1410–1417 <https://doi.org/10.1038/nn.3807>Google Scholar

Musall S, Kaufman MT, Juavinett AL, Gluf S, Churchland AK (2019) **Single-trial neural dynamics are dominated by richly varied movements** *Nat Neurosci* **22**:1677–1686 <https://doi.org/10.1038/s41593-019-0502-4>Google Scholar

Nöbauer T, Skocek O, Pernía-Andrade AJ, Weilguny L, Traub FM, Molodtsov MI, Vaziri A. (2017) **Video rate volumetric Ca<sup>2+</sup> imaging across cortex using seeded iterative demixing (SID) microscopy** *Nat Methods* **14**:811–818 <https://doi.org/10.1038/nmeth.4341>Google Scholar

Nöbauer T, Zhang Y, Kim H, Vaziri A. (2023) **Mesoscale volumetric light-field (MesoLF) imaging of neuroactivity across cortical areas at 18 Hz** *Nat Methods* **20**:600–609 <https://doi.org/10.1038/s41592-023-01789-z>Google Scholar

Okamoto H, Cherng B-W, Nakajo H, Chou M-Y, Kinoshita M. (2021) **Habenula as the experience-dependent controlling switchboard of behavior and attention in social conflict and learning** *Curr Opin Neurobiol* **68**:36–43 <https://doi.org/10.1016/j.conb.2020.12.005>Google Scholar

Oldfield CS, Grossrubatscher I, Chávez M, Hoagland A, Huth AR, Carroll EC, Prendergast A, Qu T, Gallant JL, Wyart C, Isacoff EY (2020) **Experience, circuit dynamics and forebrain recruitment in larval zebrafish prey capture** *eLife* **9**:e56619 <https://doi.org/10.7554/elife.56619>Google Scholar

Pandarinath C, Ames KC, Russo AA, Farshchian A, Miller LE, Dyer EL, Kao JC (2018) **Latent Factors and Dynamics in Motor Cortex and Their Application to Brain–Machine Interfaces** *J Neurosci* **38**:9390–9401 <https://doi.org/10.1523/jneurosci.1669-18.2018>Google Scholar

- Patterson BW, Abraham AO, MacIver MA, McLean DL (2013) **Visually guided gradation of prey capture movements in larval zebrafish** *J Exp Biology* **216**:3071–3083 <https://doi.org/10.1242/jeb.087742>Google Scholar
- Portugues R, Engert F. (2009) **The neural basis of visual behaviors in the larval zebrafish** *Curr Opin Neurobiol* **19**:644–647 <https://doi.org/10.1016/j.conb.2009.10.007>Google Scholar
- Prevedel R, Yoon Y-G, Hoffmann M, Pak N, Wetzstein G, Kato S, Schrödel T, Raskar R, Zimmer M, Boyden ES, Vaziri A. (2014) **Simultaneous whole-animal 3D imaging of neuronal activity using light-field microscopy** *Nat Methods* **11**:727–730 <https://doi.org/10.1038/nmeth.2964>Google Scholar
- Quicke P, Howe CL, Song P, Jadan HV, Song C, Knöpfel T, Neil M, Dragotti PL, Schultz SR, Foust AJ (2020) **Subcellular resolution 3D light field imaging with genetically encoded voltage indicators** *Biorxiv* <https://doi.org/10.1101/2020.05.22.108191>Google Scholar
- Randlett O, Wee CL, Naumann EA, Nnaemeka O, Schoppik D, Fitzgerald JE, Portugues R, Lacoste AM, Riegler C, Engert F, Schier AF (2015) **Whole-brain activity mapping onto a zebrafish brain atlas** *Nature methods* **12**:1039–1046 <https://doi.org/10.1038/nmeth.3581>Google Scholar
- Ruda K, Zylberberg J, Field GD (2020) **Ignoring correlated activity causes a failure of retinal population codes** *Nat Commun* **11**:4605 <https://doi.org/10.1038/s41467-020-18436-2>Google Scholar
- Rumyantsev OI, Lecoq JA, Hernandez O, Zhang Y, Savall J, Chrapkiewicz R, Li J, Zeng H, Ganguli S, Schnitzer MJ (2020) **Fundamental bounds on the fidelity of sensory cortical coding** *Nature* **580**:100–105 <https://doi.org/10.1038/s41586-020-2130-2>Google Scholar
- Scrofani G, Sola-Pikabea J, Llavador A, Sanchez-Ortiga E, Barreiro JC, Saavedra G, Garcia-Sucerquia J, Martínez-Corral M. (2017) **FIMic: design for ultimate 3D-integral microscopy of in-vivo biological samples** *Biomed Opt Express* **9**:335 <https://doi.org/10.1364/boe.9.000335>Google Scholar
- Semmelhack JL, Donovan JC, Thiele TR, Kuehn E, Laurell E, Baier H. (2014) **A dedicated visual pathway for prey detection in larval zebrafish** *eLife* **3**:e04878 <https://doi.org/10.7554/elife.04878>Google Scholar
- Shadlen MN, Newsome WT (1998) **The Variable Discharge of Cortical Neurons: Implications for Connectivity, Computation, and Information Coding** *J Neurosci* **18**:3870–3896 <https://doi.org/10.1523/jneurosci.18-10-03870.1998>Google Scholar
- Skocek O, Nöbauer T, Weilguny L, Traub FM, Xia CN, Molodtsov MI, Grama A, Yamagata M, Aharoni D, Cox DD, Golshani P, Vaziri A. (2018) **High-speed volumetric imaging of neuronal activity in freely moving rodents** *Nat Methods* **15**:429–432 <https://doi.org/10.1038/s41592-018-0008-0>Google Scholar
- Steinmetz NA, Zatzka-Haas P, Carandini M, Harris KD (2019) **Distributed coding of choice, action and engagement across the mouse brain** *Nature* **576**:266–273 <https://doi.org/10.1038/s41586-019-1787-x>Google Scholar
- Sternson SM (2020) **Exploring internal state-coding across the rodent brain** *Curr Opin Neurobiol* **65**:20–26 <https://doi.org/10.1016/j.conb.2020.08.009>Google Scholar



- Štíh V, Petrucco L, Kist AM, Portugues R. (2019) **Styttra: An open-source, integrated system for stimulation, tracking and closed-loop behavioral experiments** *Plos Comput Biol* **15**:e1006699 <https://doi.org/10.1371/journal.pcbi.1006699>Google Scholar
- Stringer C, Pachitariu M, Steinmetz N, Reddy CB, Carandini M, Harris KD (2019) **Spontaneous behaviors drive multidimensional, brainwide activity** *Science* **364**:eaav7893 <https://doi.org/10.1126/science.aav7893>Google Scholar
- Temizer I, Donovan JC, Baier H, Semmelhack JL (2015) **A Visual Pathway for Looming-Evoked Escape in Larval Zebrafish** *Curr Biol* **25**:1823–1834 <https://doi.org/10.1016/j.cub.2015.06.002>Google Scholar
- Torigoe M, Islam T, Kakinuma H, Fung CCA, Isomura T, Shimazaki H, Aoki T, Fukai T, Okamoto H. (2021) **Zebrafish capable of generating future state prediction error show improved active avoidance behavior in virtual reality** *Nat Commun* **12**:5712 <https://doi.org/10.1038/s41467-021-26010-7>Google Scholar
- Valente M, Pica G, Bondanelli G, Moroni M, Runyan CA, Morcos AS, Harvey CD, Panzeri S. (2021) **Correlations enhance the behavioral readout of neural population activity in association cortex** *Nat Neurosci* **24**:975–986 <https://doi.org/10.1038/s41593-021-00845-1>Google Scholar
- Vladimirov N, Mu Y, Kawashima T, Bennett DV, Yang C-T, Looger LL, Keller PJ, Freeman J, Ahrens MB (2014) **Light-sheet functional imaging in fictively behaving zebrafish** *Nat Methods* **11**:883–884 <https://doi.org/10.1038/nmeth.3040>Google Scholar
- Wei Z, Inagaki H, Li N, Svoboda K, Druckmann S. (2019) **An orderly single-trial organization of population dynamics in premotor cortex predicts behavioral variability** *Nat Commun* **10**:216 <https://doi.org/10.1038/s41467-018-08141-6>Google Scholar
- Weisenburger S, Vaziri A. (2018) **A Guide to Emerging Technologies for Large-Scale and Whole-Brain Optical Imaging of Neuronal Activity** *Annu Rev Neurosci* **41**:431–452 <https://doi.org/10.1146/annurev-neuro-072116-031458>Google Scholar
- Wolf S, Dubreuil AM, Bertoni T, Böhm UL, Bormuth V, Candelier R, Karpenko S, Hildebrand DGC, Bianco IH, Monasson R, Debrégeas G. (2017) **Sensorimotor computation underlying phototaxis in zebrafish** *Nat Commun* **8**:651 <https://doi.org/10.1038/s41467-017-00310-3>Google Scholar
- Yoon Y-G, Wang Z, Pak N, Park D, Dai P, Kang JS, Suk H-J, Symvoulidis P, Guner-Ataman B, Wang K, Boyden ES (2020) **Sparse decomposition light-field microscopy for high speed imaging of neuronal activity** *Optica* **7**:1457 <https://doi.org/10.1364/optica.392805>Google Scholar
- Zhou P, Resendez SL, Rodriguez-Romaguera J, Jimenez JC, Neufeld SQ, Giovannucci A, Friedrich J, Pnevmatikakis EA, Stuber GD, Hen R, Kheirbek MA, Sabatini BL, Kass RE, Paninski L. (2018) **Efficient and accurate extraction of in vivo calcium signals from microendoscopic video data** *eLife* **7**:e28728 <https://doi.org/10.7554/elife.28728>Google Scholar
- Zohary E, Shadlen MN, Newsome WT (1994) **Correlated neuronal discharge rate and its implications for psychophysical performance** *Nature* **370**:140–143 <https://doi.org/10.1038/370140a0>Google Scholar
- Zylberberg J, Cafaro J, Turner MH, Shea-Brown E, Rieke F. (2016) **Direction-Selective Circuits Shape Noise to Ensure a Precise Population Code** *Neuron* **89**:369–383 <https://doi.org/10.1016/j.neuron.2015.11.019>Google Scholar

Zylberberg J, Pouget A, Latham PE, Shea-Brown E. (2017) **Robust information propagation through noisy neural circuits** *Plos Comput Biol* **13**:e1005497 <https://doi.org/10.1371/journal.pcbi.1005497> [Google Scholar](#)

## Author information

### Jason Manley

Laboratory of Neurotechnology and Biophysics, The Rockefeller University, New York, United States, The Kavli Neural Systems Institute, The Rockefeller University, New York, United States

### Alipasha Vaziri

Laboratory of Neurotechnology and Biophysics, The Rockefeller University, New York, United States, The Kavli Neural Systems Institute, The Rockefeller University, New York, United States  
ORCID iD: [0000-0001-8346-7551](https://orcid.org/0000-0001-8346-7551)

**For correspondence:** [vaziri@rockefeller.edu](mailto:vaziri@rockefeller.edu)

## Editors

Reviewing Editor

### Carlos Brody

Princeton University, Howard Hughes Medical Institute, Princeton, United States of America

Senior Editor

### Joshua Gold

University of Pennsylvania, Philadelphia, United States of America

### Reviewer #1 (Public review):

Summary:

In this paper, Manley and Vaziri investigate whole-brain neural activity underlying behavioural variability in zebrafish larvae. They combine whole brain (single cell level) calcium imaging during the presentation of visual stimuli, triggering either approach or avoidance, and carry out whole brain population analyses to identify whole brain population patterns responsible for behavioural variability. They show that similar visual inputs can trigger large variability in behavioural responses. Though visual neurons are also variable across trials, they demonstrate that this neural variability does not degrade population stimulus decodability. Instead, they find that the neural variability across trials is in orthogonal population dimensions to stimulus encoding and is correlated with motor output (e.g. tail vigor). They then show that behavioural variability across trials is largely captured by a brain-wide population state prior to the trial beginning, which biases choice - especially on ambiguous stimulus trials. This study suggests that parts of stimulus-driven behaviour can be captured by brain-wide population states that bias choice, independently of stimulus encoding.

Comments on revisions:

The authors have revised their manuscript and provided novel analyses and figures, as well as additions to the text based on our reviewer comments.

As stated in my first review, the strength of the paper principally resides in the whole brain cellular level imaging - using a novel fourier light field microscopy (Flfm) method - in a well-known but variable behaviour.

Many of the authors' answers have provided additional support for their interpretations of results, but the new analysis in Figure 3g - further exploring the orthogonality of e1 and wopt - puts into question the interpretation of a key result: that e1 and wopt are orthogonal in a non-arbitrary way. This needs to be addressed. I have made suggestions below to address this:

Reviewer 3 had correctly highlighted the issue that in high-dimensional data, there is an increasingly high chance of two vectors being orthogonal. The authors address this by shuffling the stimulus labels. They then state (and provide a new panel g in Fig. 3) that the shuffled distribution is wider than the actual distribution, and state that a wilcoxon rank-sum test shows this is significant. Given the centrality of this claim, I would like the authors to clarify what exactly is being done here, as it is not clear to me how this conclusion can be drawn from this analysis:

In lines 449:453 the authors state:

'While it is possible to observe shuffled vectors which are nearly orthogonal to e1, the shuffled distribution spans a significantly greater range of angles than the observed data ( $p < 0.05$ , Wilcoxon rank-sum test), demonstrating that this orthogonality is not simply a consequence of analyzing multi-dimensional activity patterns.'

I don't understand how the authors arrive at the p-value using a rank-sum test here. (a) What is the n in this test? Is n the number of shuffles? If so, this violates the assumptions of the test (as n must be the number of independent samples and not the arbitrary number of shuffles). (b) If the shuffling was done once for each animal and compared with actual data with a rank-sum test, how likely is that shuffling result to happen in 10000 shuffle comparisons? I am highlighting this, as it looks from Figure 3g that the shuffled distribution is substantially overlapping with the actual data (i.e., not outside of the 95 percentile of the shuffled distribution), which would suggest that the angle found between e1 and wept could happen by chance.

I would also suggest the authors instead test whether e1 is consistently aligned with itself when calculated on separate held out data-sets (for example by bootstrapping 50-50 splits of the data). If they can show that there is a close alignment between independently calculated e1's across separate data sets (and do the same for wopt), and then show e1 and wopt are orthogonal, then that supports their statement that e1 and wopt are orthogonal in a meaningful way. Given that e1 captures tail vigor variability (and Wopt appears to not) then I would think this could be the case. But the current answer the authors have given is not supporting their statement.

<https://doi.org/10.7554/eLife.97014.2.sa2>

## Reviewer #2 (Public review):

This work by Manley and Vaziri identify brain networks that are associated with trial-to-trial variability during prey-capture and predator avoidance behaviors. However, mixing of signals across space and time make it difficult to interpret the data generated and relate the data to findings from prior work.

Comments on revisions:

In their response to prior reviewer comments, Manley and Vaziri have now provided helpful methodological clarity and additional analyses. The additional work makes clear that the

claims of variability and mixing of sensory, motor, and internal variables at the single-cell level are not well supported.

#### RESOLUTION

- The new information provided regarding resolution may not be very relevant as this was from an experiment in air. It would be much more informative to show how PSF degrades in the brain with depth.

#### DEPTH

- It is helpful to see the registered light-field and confocal images. Both appear to provide poor or little information in regions >200 below the surface (like the hypothalamus), making the claim that whole brain data is being collected at cellular resolution difficult to justify.

#### MERGING

- The typical soma at these ages has a radius of 2.5 microns, which corresponds to a volume of  $65 \text{ microns}^3$ . Given the close packing of most cells, this means that a typical ROI of  $750 \text{ microns}^3$  contains more than 10 neurons. Therefore, the authors should not claim they are reporting activity at cellular resolution.
- Furthermore, the fact that these ROIs contains tens of cells brings into question the degree of variability at the single-cell level. For example, if every cluster of 10 cells has one variable cell, then all clusters might be labeled as exhibiting variability even though only 10% of the cells show variability.

#### SLOW CALCIUM DYNAMICS

- Convolution/Deconvolution with the inappropriate kernel both have problems, some of which the authors have noted. However, by not deconvolving, the authors are significantly obscuring the interpretation of their data by mixing together signals across time.
- Also, the claim that "neurons highly tuned to a particular stimulus exhibited variability in their responses across multiple presentations of the same stimuli" should be clarified or qualified. It is not clear from what has been shown if the responses are indeed variable, or rather if there is additional activity (or apparent activity) occasionally present that shifts the pre-stimulus baseline around (for example, 3J suggests that in many cases the visual signal from the prior trial is still present when a new trial begins).
- Figure 3A should show when the stimulus occurs, should show some of the prestimulus period, and ideally be off-set corrected so all traces in a given panel start at the same y-value at the beginning of the stimulus period.

#### ORTHOGONALITY

- It is now clearer that the visual signal and noise vectors were determined for the entire time series with all trials. Therefore, the concern that sources of activation in advance of a given trial were being ignored is alleviated. The concern remains, however, that these sources are being properly accounted for given potential kernel variations and nonlinearity. Nonetheless, it is recognized that the GCaMP filtering most likely would lead to a decrease in the disparity between two populations.
- The authors' clarification that the analyzed ROIs consist of cell clusters raises the trivial possibility that the observed orthogonality between the visual signal and leading noise vectors is explained by noise simply reflecting the activation of different motor or motor-planning related neurons in an ROI, neurons that are separate from visually-encoding neurons in the same cluster.

#### SOURCES of VARIABILITY

- The data presented in Supplemental Figure 3Ei actually is suggestive that eye movements are a significant contributor to the reported variability. Notice how in (1 4) vs (1 5) and (4 7) vs (4 8) there is a notable difference in the distribution of responses. Adding eye kinematic variables to the analysis of Figure S4 could be clarifying.-

## Author response:

The following is the authors' response to the original reviews

### Public Reviews:

#### Reviewer #1 (Public Review):

##### Summary:

*In this paper, Manley and Vaziri investigate whole-brain neural activity underlying behavioural variability in zebrafish larvae. They combine whole brain (single cell level) calcium imaging during the presentation of visual stimuli, triggering either approach or avoidance, and carry out whole brain population analyses to identify whole brain population patterns responsible for behavioural variability. They show that similar visual inputs can trigger large variability in behavioural responses. Though visual neurons are also variable across trials, they demonstrate that this neural variability does not degrade population stimulus decodability. Instead, they find that the neural variability across trials is in orthogonal population dimensions to stimulus encoding and is correlated with motor output (e.g. tail vigor). They then show that behavioural variability across trials is largely captured by a brain-wide population state prior to the trial beginning, which biases choice - especially on ambiguous stimulus trials. This study suggests that parts of stimulus-driven behaviour can be captured by brain-wide population states that bias choice, independently of stimulus encoding.*

##### Strengths:

- The strength of the paper principally resides in the whole brain cellular level imaging in a well-known but variable behaviour.
- The analyses are reasonable and largely answer the questions the authors ask.
- Overall the conclusions are well warranted.

##### Weaknesses:

*A more in-depth exploration of some of the findings could be provided, such as:*

- Given that thousands of neurons are recorded across the brain a more detailed parcellation of where the neurons contribute to different population coding dimensions would be useful to better understand the circuits involved in different computations.

We thank the reviewer for noting the strengths of our study and agree that these findings have raised a number of additional avenues which we intend to explore in depth in future studies. In response to the reviewer's comment above, we have added a number of additional figure panels (new Figures S1E, S3F-G, 4I(i), 4K(i), and S5F-G) and updated panels (Figures 4I(ii) and 4K(ii) in the revised manuscript) to show a more detailed parcellation of the visually-evoked neurons, noise modes, turn direction bias population, and responsiveness bias population. To do so, we have aligned our recordings to the Z-Brain atlas (Randlett et al., 2015) as shown in new Figure S1E. In addition, we provided a more detailed parcellation of the neuronal ensembles by providing projections of the full 3D volume along the xy and yz axes, in addition to the unregistered xy projection shown in Figures 4H and 4J in the revised manuscript. We also found that the distribution of neurons across our *huc:h2b-gcamp6s*



recordings is very similar to the distribution of labeling in the *huc:h2b-rfp* reference image from the Z-Brain atlas (Figure S1E), which further supports our whole-brain imaging results.

Overall, we find that this more detailed quantification and visualization is consistent with our interpretations. In particular, we show that the optimal visual decoding population ( $w_{\text{opt}}$ ) and the largest noise mode (e1) are localized to the midbrain (Figures S3F-G). This is expected, as in Figure 3 we first extracted a low-dimensional subspace of whole-brain neural activity that optimally preserved visual information. Additionally, we provide new evidence that the populations correlated with the turn bias and responsiveness bias are distributed throughout the brain, including a relatively dense localization to the cerebellum, telencephalon, and dorsal diencephalon (habenula, new Figures 4H-K and S5F-G).

*- Given that the behaviour on average can be predicted by stimulus type, how does the stimulus override the brain-wide choice bias on some trials? In other words, a better link between the findings in Figures 2 and 3 would be useful for better understanding how the behaviour ultimately arises.*

We agree with the reviewer that one of the most fundamental questions that this study has raised is how the identified neuronal populations predictive of decision variables (which we describe as an internal “bias”) interact with the well-studied, visually-evoked circuitry. A major limitation of our study is that the slow dynamics of the NL-GCaMP6s prevent clearly distinguishing any potential difference in the onset time of various neurons during the short trials, which might provide clues into which neurons drive versus later reflect the motor output. However, given that these ensembles were also found to be correlated with spontaneous turns, our hypothesis is that these populations reflect brain-wide drives that enable efficient exploration of the local environment (Dunn et al. 2016, doi.org/10.7554/eLife.12741). Further, we suspect that a sufficiently strong stimulus drive (e.g., large, looming stimuli) overrides these ongoing biases, which would explain the higher average pre-stimulus predictability in trials with small to intermediate-sized stimuli. An important follow-up line of experimentation could involve comparing the neuronal dynamics of specific components of the visual circuitry at distinct internal bias states, ideally utilizing emerging voltage indicators to maximize spatiotemporal specificity. For example, what is the difference between trials with a large looming stimulus in the left visual fields when the turn direction bias indicates a leftward versus rightward drive?

*- What other motor outputs do the noise dimensions correlate with?*

To better demonstrate the relationship between neural noise modes and motor activity that we described, we have provided a more detailed correlation analysis in new Figure S4A. We extracted additional features related to the larva’s tail kinematics, including tail vigor, curvature, principal components of curvature, angular velocity, and angular acceleration (S4A(i)). Some of these behavioral features were correlated with one another; for example, in the example traces, PC1 appears to capture nearly the same behavioral feature as tail vigor. The largest noise modes showed stronger correlations with motor output than the smaller noise modes, which is reminiscent recent work in the mouse showing that some of the neural dimensions with highest variance were correlated with various behavioral features (Musall et al. 2019; Stringer et al. 2019; Manley et al. 2024). We anticipate additional motor outputs would exhibit correlations with neural noise modes, such as pectoral fin movements (not possible to capture in our preparation due to immobilization) and eye movements.

*The dataset that the authors have collected is immensely valuable to the field, and the initial insights they have drawn are interesting and provide a good starting ground for a more expanded understanding of why a particular action is determined outside of the parameters experimenters set for their subjects.*

We thank the reviewer for noting the value of our dataset and look forward to future efforts motivated by the observations in our study.

**Reviewer #2 (Public Review):**

*Overview*

*In this work, Manley and Vaziri investigate the neural basis for variability in the way an animal responds to visual stimuli evoking prey-capture or predator-avoidance decisions. This is an interesting problem and the authors have generated a potentially rich and relevant data set. To do so, the authors deployed Fourier light field microscopy (Flfm) of larval zebrafish, improving upon prior designs and image processing schemes to enable volumetric imaging of calcium signals in the brain at up to 10 Hz. They then examined associations between neural activity and tail movement to identify populations primarily related to the visual stimulus, responsiveness, or turn direction - moreover, they found that the activity of the latter two populations appears to predict upcoming responsiveness or turn direction even before the stimulus is presented. While these findings may be valuable for future more mechanistic studies, issues with resolution, rigor of analysis, clarity of presentation, and depth of connection to the prior literature significantly dampen enthusiasm.*

*Imaging*

*- Resolution: It is difficult to tell from the displayed images how good the imaging resolution is in the brain. Given scattering and lensing, it is important for data interpretation to have an understanding of how much PSF degrades with depth.*

We thank the reviewer for their comments and agree that the dependence of the PSF and resolution as a function of depth is an important consideration in light field imaging. To quantify this, we measured the lateral resolution of the fLFM as a function of distance from the native image plane (NIP) using a USAF target. The USAF target was positioned at various depths using an automated z-stage, and the slice of the reconstructed volume corresponding to that depth was analyzed. An element was considered resolved if the modulation transfer function (MTF) was greater than 30%.

In new Figure S1A, we plot the resolution measurements of the fLFM as compared to the conventional LFM (Prevedel et al., 2014), which shows the increase in resolution across the axial extent of imaging. In particular, the fLFM does not exhibit the dramatic drop in lateral resolution near the NIP which is seen in conventional LFM. In addition, the expanded range of high-resolution imaging motivates our increase from an axial range of 200 microns in previous studies to 280 microns in this study.

*- Depth: In the methods it is indicated that the imaging depth was 280 microns, but from the images of Figure 1 it appears data was collected only up to 150 microns. This suggests regions like the hypothalamus, which may be important for controlling variation in internal states relevant to the behaviors being studied, were not included.*

The full axial range of imaging was 280 microns, i.e. spanning from 140 microns below to 140 microns above the native imaging plane. After aligning our recordings to the Z-Brain dataset, we have compared the 3D distribution of neurons in our data (new Figure S1E(i)) to the labeling of the reference brain (Figure S1E(ii)). This provides evidence that our imaging preparation largely captures the labeling seen in a dense, high-resolution reference image within the indicated 280 microns range.

*- fLfm data processing: It is important for data interpretation that the authors are clearer about how the raw images were processed. The de-noising process specifically needs to be explained in greater detail. What are the characteristics of the noise being removed? How is time-varying signal being distinguished from noise? Please provide a supplemental with images and algorithm specifics for each key step.*

We thank the reviewer for their comment. To address the reviewer's point regarding the data processing pipeline utilized in our study, in our revised manuscript we have added a number of additional figure panels in Figure S1B-E to quantify and describe the various steps of the pipeline in greater depth.

First, the raw fLFM images are denoised. The denoising approach utilized in the fLFM data processing pipeline is not novel, but rather a custom-trained variant of Lecoq et al.'s (2021) DeepInterpolation method. In our original manuscript, we also described the specific architecture and parameters utilized to train our specific variation of DeepInterpolation model. To make this procedure clearer, we have added the following details to the methods:

"DeepInterpolation is a self-supervised approach to denoising, which denoises the data by learning to predict a given frame from a set of frames before and after it. Time-varying signal can be distinguished from shot noise because shot noise is independent across frames, but signal is not. Therefore, only the signal is able to be predicted from adjacent frames. This has been shown to provide a highly effective and efficient denoising method (Lecoq et al., 2021)."

Therefore, time-varying signal is distinguished from noise based on the correlations of pixel intensity across consecutive imaging frames. To better visualize this process, in new Figure S1B we show example images and fluorescence traces before and after denoising.

*- Merging: It is noted that nearby pixels with a correlation greater than 0.7 were merged. Why was this done? Is this largely due to cross-contamination due to a drop in resolution? How common was this occurrence? What was the distribution of pixel volumes after aggregation? Should we interpret this to mean that a 'neuron' in this data set is really a small cluster of 10-20 neurons? This of course has great bearing on how we think about variability in the response shown later.*

First, to be clear, nearby pixels were not merged; instead neuronal ROIs identified by CNMF-E were merged, as we had described: "the CNMF-E algorithm was applied to each plane in parallel, after which the putative neuronal ROIs from each plane were collated and duplicate neurons across planes were merged." If this merging was not performed, the number of neurons would be overestimated due to the relatively dense 3D reconstruction with voxels of 4  $\mu$ m axially. Therefore, this merging is a requisite component of the pipeline to avoid double counting of neurons, regardless of the resolution of the data.

However, we agree with the reviewer that the practical consequences of this merging were not previously described in sufficient detail. Therefore, in our revision we have added additional quantification of the two critical components of the merging procedure: the number of putative neuronal ROIs merged and the volume of the final 3D neuronal ROIs, which demonstrate that a neuron in our data should not be interpreted as a cluster of 10-20 neurons.

In new Figure S1C(i), we summarize the rate of occurrence of merging by assessing the number of putative 2D ROIs which were merged to form each final 3D neuronal ROI. Across  $n=10$  recordings, approximately 75% of the final 3D neuronal ROIs involved no merging at all, and few instances involved merging more than 5 putative ROIs. Next, in Figure S1C(ii), we quantify the volume of the final 3D ROIs. To do so, we counted the number of voxels contributing to each final 3D neuronal ROI and multiplied that by the volume of a single

voxel ( $2.4 \times 2.4 \times 4 \mu\text{m}^3$ ). The majority of neurons had a volume of less than  $1000 \mu\text{m}^3$ , which corresponds to a spherical volume with a radius of roughly 6.2  $\mu\text{m}$ . In summary, both the merging statistics and volume distribution demonstrate that few neuronal ROIs could be consistent with “a small cluster of 10-20 neurons”.

- *Bleaching: Please give the time constants used in the fit for assessing bleaching.*

As described in the Methods, the photobleaching correction was performed by fitting a bi-exponential function to the mean fluorescence across all neurons. We have provided the time constants determined by these fits for  $n=10$  recordings in new Figure S1D(i). In addition, we provided an example of raw mean activity, the corresponding bi-exponential fit, and the mean activity after correction in Figure S1D(ii). These data demonstrate that the dominant photobleaching effect is a steep decrease in mean signal at the beginning of the recording (represented by the estimated time constant  $\tau_1$ ), followed by a slow decay ( $\tau_2$ ).

#### Analysis

- *Slow calcium dynamics: It does not appear that the authors properly account for the slow dynamics of calcium-sensing in their analysis. Nuclear-localized GCaMP6s will likely have a kernel with a multiple-second decay time constant for many of the cells being studied. The value used needs to be given and the authors should account for variability in this kernel time across cell types. Moreover, by not deconvolving their signals, the authors allow for contamination of their signal at any given time with a signal from multiple seconds prior. For example, in Figure 4A (left turns), it appears that much of the activity in the first half of the time-warped stimulus window began before stimulus presentation - without properly accounting for the kernel, we don't know if the stimulus-associated activity reported is really stimulus-associated firing or a mix of stimulus and pre-stimulus firing. This also suggests that in some cases the signals from the prior trial may contaminate the current trial.*

We would like to respond to each of the points raised here by the reviewer individually.

(1) *“It does not appear that the authors properly account for the slow dynamics of calcium-sensing in their analysis. Nuclear-localized GCaMP6s will likely have a kernel with a multiple-second decay time constant for many of the cells being studied. The value used needs to be given...”*

We disagree with the reviewer’s claim that the slow dynamics of the calcium indicator GCaMP were not accounted for. While we did not deconvolve the neuronal traces with the GCaMP response kernel, in every step in which we correlated neural activity with sensory or motor variables, we convolved the stimulus or motor timeseries with the GCaMP kernel, as described in the Methods. Therefore, the expected delay and smoothing effects were accounted for when analyzing the correlation structure between neural and behavioral or stimulus variables, as well as during our various classification approaches. To better describe this, we have added the following description of the kernel to our Methods:

“The NL-GCaMP6s kernel was estimated empirically by aligning and averaging a number of calcium events. This kernel corresponds to a half-rise time of 400 ms and half-decay time of 4910 ms.”

This approach accounts for the GCaMP kernel when relating the neuronal dynamics to stimuli and behavior, while avoiding any artifacts that could be introduced from improper deconvolution or other corrections directly to the calcium dynamics. Deconvolution of calcium imaging data, and in particular nuclear-localized (NL) GCaMP6s, is not always a robust procedure. In particular, GCaMP6s has a much more nonlinear response profile than

newer GCaMP variants such as jGCaMP8 (Zhang et al. 2023, doi:10.1038/s41586-023-05828-9), as the reviewer notes later in their comments. The nuclear-localized nature of the indicator used in our study also provides an additional nonlinear effect. Accounting for a nonlinear relationship between calcium concentration and fluorescence readout is significantly more difficult because such nonlinearities remove the guarantee that the optimization approaches generally used in deconvolution will converge to global extrema. This means that deconvolution assuming nonlinearities is far less robust than deconvolution using the linear approximation (Vogelstein et al. 2010, doi: 10.1152/jn.01073.2009). Therefore, we argue that we are not currently aware of any appropriate methods for deconvolving our NL-GCaMP6s data, and take a more conservative approach in our study.

We also argue that the natural smoothness of calcium imaging data is important for the analyses utilized in our study (Shen et al., 2022, doi:10.1016/j.jneumeth.2021.109431). Even if our data were deconvolved in order to estimate spike trains or more point-like activity patterns, such data are generally smoothed (e.g., by estimating firing rates) before dimensionality reduction, which is a core component of our neuronal population analyses. Further, Wei et al. (2020, doi:10.1371/journal.pcbi.1008198) showed in detail that deconvolved calcium data resulted in less accurate population decoding, whereas binned electrophysiological data and raw calcium data were equally accurate. When using other techniques, such as clustering of neuronal activity patterns (a method we do not employ in this study), spike and deconvolved calcium data were instead shown to be more accurate than raw calcium data. Therefore, we do not believe deconvolution of the neuronal traces is appropriate in this case without a better understanding of the NL-GCaMP6s response, and do not rely on the properties of deconvolution for our analyses. Still, we agree with the reviewer that one must be mindful of the GCaMP kernel when analyzing and interpreting these data, and therefore have noted the delayed and slow kinematics of the NL-GCaMP within our manuscript, for example: “To visualize the neuronal activity during a given trial while accounting for the delay and kinematics of the nuclear-localized GCaMP (NL-GCaMP) sensor, a duration of approximately 15 seconds is extracted beginning at the onset of the 3-second visual stimulus period.”

| (2) “... and the authors should account for variability in this kernel time across cell types.”

In addition to the points raised above, we are not aware of any deconvolution procedures which have successfully shown the ability to account for variability in the response kernel across cell types in whole-brain imaging data when cell type is unknown *a priori*. Pachitariu et al. (2018, doi:10.1523/JNEUROSCI.3339-17.2018) showed that the best deconvolution procedures for calcium imaging data rely on a simple algorithm with a fixed kernel. Further, more complicated approaches either utilize either explicit priors about the calcium kernel or learn implicit priors using supervised learning, neither of which we would be able to confirm are appropriate for our dataset without ground truth electrophysiological spike data.

However, we agree with the reviewer that we must interpret the data while being mindful that there could be variability in this kernel across neurons, which is not accounted for in our fixed calcium kernel. We have added the following sentence to our revised manuscript to highlight this limitation:

“The used of a fixed calcium kernel does not account for any variability in the GCaMP response across cells, which could be due to differences such as cell type or expression level. Therefore, this analysis approach may not capture the full set of neurons which exhibit stimulus correlations but exhibit a different GCaMP response.”

| (3) “without properly accounting for the kernel, we don't know if the stimulus-associated activity reported is really stimulus-associated firing or a mix of stimulus and pre-stimulus firing”

While we agree with the reviewer that the slow dynamics of the indicator will cause a delay and smoothing of the signal over time, we would like to point out that this effect is highly directional. In particular, we can be confident that pre-stimulus activity is not contaminated by the stimulus given the data we describe in the next point regarding the timing of visual stimuli relative to the GCaMP kernel. The reviewer is correct that post-stimulus firing can be mixed with pre-stimulus firing due to the GCaMP kernel. However, our key claims in Figure 4 center around turn direction and responsiveness biases, which are present even before the onset of the stimulus. Still, we have highlighted this delay and smoothing to readers in the updated version of our manuscript.

(4) *"This also suggests that in some cases the signals from the prior trial may contaminate the current trial"*

We have carefully chosen the inter-stimulus interval for maximum efficiency of stimulation, while ensuring that contamination from the previous stimulus is negligible. The inter-stimulus interval was chosen by empirically analyzing preliminary data of visual stimulation with our preparation. New Figure S3C shows the delay and slow kinematics due to our indicator; indeed, visually-evoked activity peaks after the end of the short stimulus period. Importantly, however, the visually-evoked activity is at or near baseline at the start of the next trial.

Finally, we would like to note that our stimulation protocol is randomized, as described in the Methods. Therefore, the previous stimulus has no correlation with the current stimulus, which would prevent any contamination from providing predictive power that could be identified by our visual decoding methods.

- *Partial Least Squares (PLS) regression: The steps taken to identify stimulus coding and noise dimensions are not sufficiently clear. Please provide a mathematical description.*

We have updated the Results and Methods sections of our revised manuscript to describe in more mathematical detail the approach taken to identify the relevant dimensions of neuronal activity:

"The comparison of the neural dimensions encoding visual stimuli versus trial-to-trial noise was modeled after Rumyantsev et al. (2020). Partial least squares (PLS) regression was used to find a low-dimensional space that optimally predicted the visual stimuli, which we refer to as the visually-evoked neuronal activity patterns. To perform regression, a visual stimulus kernel was constructed by summing the timeseries of each individual stimulus type, weighted by the stimulus size and negated for trials on the right visual field, thus providing a single

response variable  $\bar{y}$  encoding both the location, size, and timing of all the stimulus

presentations. This stimulus kernel was the convolved with the temporal response kernel of our calcium indicator (NL-GCaMP6s).

PLS regression identifies the normalized dimensions  $\bar{p}_i$  and  $\bar{q}_j$  that maximize the covariance

between paired observations  $\bar{x}_i$  and  $\bar{y}_j$ , respectively. In our case, the visual stimulus is

represented by a single variable  $\bar{y}$ , simplifying the problem to identifying the subspace of neural activity that optimally preserves information about the visual stimulus (sometimes



referred to as PLS1 regression). That is, the  $N \times T$  neural time series matrix  $X$  is reduced to a  $d \times T$  matrix spanned by a set of orthonormal vectors. PLS1 regression is performed as follows:

PLS1 algorithm

Let  $X_i = X$  and  $y_i = \vec{y}$ . For  $i = 1 \dots d$ ,

$$(1) \quad p_i = X_i^T y_i / |X_i^T y_i|$$

$$(2) \quad t_i = X_i p_i$$

$$(3) \quad q_i = X_i^T y_i (t_i^T t_i)^{-1}$$

$$(4) \quad X_{i+1} = X_i - t_i q_i^T$$

$$(5) \quad c = y_i^T t_i (t_i^T t_i)^{-1} \text{ (note this is scalar)}$$

$$(6) \quad y_{i+1} = y_i - c_i t_i$$

The projections of the neural data  $\{p_i\}$  thus span a subspace that maximally preserves information about the visual stimulus  $\vec{y}$ . Stacking these projections into the  $N \times d$  matrix  $P$

that represents the transform from the whole-brain neural state space to the visually-evoked subspace, the optimal decoding direction is given by the linear least squares solution

$\vec{w}_{\text{opt}} = P(P^T X^T X P)^{-1} P^T X^T \vec{y}$ . The dimensionality  $d$  of PLS regression was optimized using 6-fold cross-validation with 3 repeats and choosing the dimensionality between  $d = 1$  and 20 with the lowest cross-validated mean squared error for each larva. Then,  $\vec{w}_{\text{opt}}$  was computed using all time points.

For each stimulus type, the noise covariance matrix was computed in the low-dimensional PLS space, given that direct estimation of the noise covariances across many thousands of neurons would likely be unreliable. A noise covariance matrix was calculated separately for each stimulus, and then averaged across all stimuli. As before, the mean activity  $\mu_i$  for each neuron was computed over each stimulus presentation period. The noise covariance then describes the correlated fluctuations  $\delta_i$  around this mean response for each pair of neurons  $i$  and  $j$ , where  $c_{ij} = \langle \delta_i \delta_j \rangle$

The noise modes  $\vec{e}_\alpha$  for  $\alpha = 1 \dots d$  were subsequently identified by eigendecomposition of the

mean noise covariance matrix across all stimuli,  $c = \sum \lambda_\alpha \vec{e}_\alpha \vec{e}_\alpha^T$ . The angle between the optimal

stimulus decoding direction and the noise modes is thus given by  $\cos \theta_{\alpha} = (\vec{w}_{\text{opt}} \cdot \vec{e}_{\alpha}) / (|\vec{w}_{\text{opt}}| |\vec{e}_{\alpha}|)$ .

- No response: It is not clear from the methods description if cases where the animal has no tail response are being lumped with cases where the animal decides to swim forward and thus has a large absolute but small mean tail curvature. These should be treated separately.

We thank the reviewer for raising the potential for this confusion and agree that forward-motion trials should not be treated the same as motionless trials. While these types of trial were indeed treated separately in our original manuscript, we have updated the Methods section of our revised manuscript to make this clear:

“Left and right turn trials were extracted as described previously. Response trials included both left and right turn trials (i.e., the absolute value of mean tail curvature  $> \sigma_{\text{active}}$ ), whereas nonresponse trials were motionless (absolute mean tail curvature  $< \sigma_{\text{active}}$ ). In particular, forward-motion trials were excluded from these analyses.”

While our study has focused specifically on left and right turns, we hypothesize that the responsiveness bias ensemble may also be involved in forward movements and look forward to future work exploring the relationship between whole-brain dynamics and the full range of motor outputs.

- Behavioral variability: Related to Figure 2, within- and across-subject variability are confounded. Please disambiguate. It may also be informative on a per-fish basis to examine associations between reaction time and body movement.

The reviewer is correct that our previously reported summary statistics in Figure 2D-F were aggregated across trials from multiple larvae. Following the reviewer’s suggestion to make the magnitudes of across-larvae and within-larva variability clear, in our revised manuscript we have added two additional figure panels to Figure S2.

New Figure S2A highlights the across-larvae variability in mean head-directed behavioral responses to stimuli of various sizes. Overall, the relationship between stimulus size and the mean tail curvature across trials is largely consistent across larvae; however, the crossing-over point between leftward (positive curvature) and rightward (negative curvature) turns for a given side of the visual field exhibits some variability across larvae.

New Figure S2B shows examples of within-larva variability by plotting the mean tail curvature during single trials for two example larvae. Consistent with Figure 2G which also demonstrates within-larva variability, responses to a given stimulus are variable across trials in both examples. However, this degree of within-larva variability can appear different across larvae. For example, the larva shown on the left of Figure S2B exhibits greater overlap between responses to stimuli presented on opposite visual fields, whereas the larva shown on the right exhibits greater distinction between responses.

- Data presentation clarity: All figure panels need scale bars - for example, in Figure 3A there is no indication of timescale (or time of stimulus presentation). Figure 3I should also show the time series of the  $w_{\text{opt}}$  projection.

We appreciate the reviewer’s attention to detail in this regard. We have added scalebars to Figures 3A, 3H-I, S4B(ii), 4H, 4J in the revised manuscript, and all new figure panels where relevant. In addition, the caption of Figure 3A has been updated to include a description of the time period plotted relative to the onset of the visual stimulus.

Additionally, we appreciate the reviewer's idea to show  $w_{opt}$  in Figure 3J of the revised manuscript (previously Figure 3I). This clearly shows that the visual decoding project is inactive during the short baseline period before visual stimulation begins, whereas the noise mode is correlated with motor output throughout the recording.

*- Pixel locations: Given the poor quality of the brain images, it is difficult to tell the location of highlighted pixels relative to brain anatomy. In addition, given that the midbrain consists of much more than the tectum, it is not appropriate to put all highlighted pixels from the midbrain under the category of tectum. To aid in data interpretation and better connect this work with the literature, it is recommended that the authors register their data sets to standard brain atlases and determine if there is any clustering of relevant pixels in regions previously associated with prey-capture or predator-avoidance behavior.*

We agree with the reviewer that registration of our datasets to a standard brain atlas is a highly useful addition. While the dense, pan-neuronal labeling makes the isolation of highly specific circuit components difficult, we have shown in more detail the specific brain regions contributing to these populations by aligning our recordings to the Z-Brain atlas (Randlett et al., 2015) as shown in new Figures S1E, S3F-G, 4I, 4K, and S5F-G. In addition, we provided a more detailed parcellation of the neuronal ensembles by providing projections of the full 3D volume along the xy and yz axes, in addition to the unregistered xy projection shown in new Figures 4H and 4J. We also found that the distribution of neurons in our huc:H2B-GCaMP6s recordings is very similar to the distribution of labeling in the huc:H2B-RFP reference image from the Z-Brain atlas (new Figure S1E), which further supports our whole-brain imaging results.

Overall, we find that this more detailed quantification and visualization is consistent with the interpretations in the previous version of our manuscript. In particular, we show that optimal visual decoding population ( $w_{opt}$ ) and largest noise mode ( $e_1$ ) are localized to the midbrain (new Figures S3F-G), which is expected since in Figure 3 we first extracted a low-dimensional subspace of whole-brain neural activity that optimally preserved visual information. Additionally, we provide additional evidence that the populations correlated with the turn bias and responsiveness bias are distributed throughout the brain, including a relatively dense localization to the cerebellum, telencephalon, and dorsal diencephalon (habenula, new Figures 4H-K and S5F-G).

Finally, the reviewer is correct that our original label of "tectum" was a misnomer; the region analyzed corresponded to the midbrain, including the tegmentum, torus longitudinalis, and torus semicircularis in addition to the tectum. We have updated the brain regions shown and labels throughout the manuscript.

#### *Interpretation*

*-  $w_{opt}$  and  $e_1$  orthogonality: The statement that these two vectors, determined from analysis of the fluorescence data, are orthogonal, actually brings into question the idea that true signal and leading noise vectors in firing-rate state-space are orthogonal. First, the current analysis is confounding signals across different time periods - one could assume linearity all the way through the transformations, but this would only work if earlier sources of activation were being accounted for. Second, the transformation between firing rate and fluorescence is most likely not linear for GCaMP6s in most of the cells recorded. Thus, one would expect a change in the relationship between these vectors as one maps from fluorescence to firing rate.*

Unfortunately, we are not entirely sure we have understood the reviewer's argument. We are assuming that the reviewer's first sentence is suggesting that the observation of orthogonality in the neural state space measured in calcium imaging precludes the possibility ("actually brings into question", as the reviewer states) that the same neural ensembles could be orthogonal in firing rate state space measured by electrophysiological data. If this is the reviewer's conjecture, we respectfully disagree with it. Consider a toy example of a neural network containing  $N$  ensembles of neurons, where the neurons within an ensemble all fire simultaneously, and two populations never fire at the same time. As long as the "switching" of firing between ensembles is not fast relative to the resolution of the GCaMP kernel, the largest principal components would represent orthogonal dimensions differentiating the various ensembles, both when observing firing rates or observing timeseries convolved by the GCaMP kernel. This is a simple example where the observed orthogonality would appear similar in both calcium imaging and electrophysiological data, demonstrating that we should not allow conclusions from fluorescence data to "bring into question" that the same result could be observed in firing rate data.

We also disagree with the reviewer's argument that we are "confounding signals across time periods". Indeed, we must interpret the data in light of the GCaMP response kernel. However, all of the analyses presented here are performed on instantaneous measurements of population activity patterns. These activity patterns do represent a smoothed, likely nonlinear integration of recent neuronal activity, but unless the variability in the GCaMP response kernel (discussed above) is widely different across these populations (which has not been observed in the literature), we do not expect that the GCaMP transformations would artificially induce orthogonality in our analysis approach. Such smoothing operations tend to instead increase correlations across neurons and population decoding approaches generally benefit from this smoothness, as we have argued above. However, a much more problematic situation would be if we were comparing the activity of two neuronal populations at different points in time (which we do not include in this study), in which case the nonlinearities could overaccentuate orthogonality between non-time-matched activity patterns.

Finally, we agree with the reviewer that the transformation between firing rate and fluorescence is very likely nonlinear and that these vectors of population activity do not perfectly represent what would be observed if one had access to whole-brain, cellular-resolution electrophysiology spike data. However, similar observations regarding the brain-wide, distributed encoding of behavior have been confirmed across recording modalities in the mouse (Stringer et al., 2019; Steinmetz et al., 2019), where large-scale electrophysiology utilizing highly invasive probes (e.g., Neuropixels) is more feasible than in the larval zebrafish. With the advent of whole-brain voltage imaging in the larval zebrafish, we expect any differences between calcium and voltage dynamics will be better understood, yet such techniques will likely continue to suffer to some extent from the nonlinearities described here.

*- Sources of variability: The authors do not take into account a fairly obvious source of variability in trial-to-trial response - eye position. We know that prey capture responsiveness is dependent on eye position during stimulus (see Figure 4 of PMID: 22203793). We also expect that neurons fairly early in the visual pathway with relatively narrow receptive fields will show variable responses to visual stimuli as the degree of overlap with the receptive field varies with eye movement. There can also be small eye-tracking movements ahead of the decision to engage in prey capture (Figure 1D, PMID: 31591961) that can serve as a drive to initiate movements in a particular direction. Given these possibilities indicating that the behavioral measure of interest is gaze, and the fact that eye movements were apparently monitored, it is surprising that the authors did not include eye movements in the analysis and interpretation of their data.*

We agree with the reviewer that eye movements, such as saccades and convergence, are important motor outputs that are well-known to play a role in the sequence of motor actions during prey capture and other behaviors. Therefore, we have added the following new eye tracking results to our revised manuscript:

“In order to confirm that the observed neural variability in the visually-evoked populations was not predominantly due to eye movements, such as saccades or convergence, we tracked the angle of each eye. We utilized DeepLabCut, a deep learning tool for animal pose estimation (Mathis et al., 2018), to track keypoints on the eye which are visible in the raw fLFM images, including the retina and pigmentation (Figure S3D(i)). This approach enabled identification of various eye movements, such as convergence and the optokinetic reflex (Figure S3D(ii-iii)). Next, we extracted a number of various eye states, including those based on position (more leftward vs. rightward angles) and speed (high angular velocity vs. low or no motion). Figure S3E(i) provides example stimulus response profiles across trials of the same visual stimulus in each of these eye states, similar to a single column of traces in Figure 3A broken out into more detail. These data demonstrate that the magnitude and temporal dynamics of the stimulus-evoked responses show apparently similar levels of variability across eye states. If neural variability was driven by eye movement during the stimulus presentation, for example, one would expect to see much more variability during the high angular velocity trials than low, which is not apparent. Next, we asked whether the dominant neural noise modes vary across eye states, which would suggest that the geometry of neuronal variability is influenced by eye movements or states. To do so, the dominant noise modes were estimated in each of the individual eye conditions, as well as bootstrapped trials from across all eye conditions. The similarity of these noise modes estimated from different eye conditions (Figure S3E(ii), right)) was not significantly different from the similarity of noise modes estimated from bootstrapped random samples across all eye conditions (Figure S3E(ii), left)). Therefore, while movements of the eye likely contribute to aspects of the observed neural variability, they do not dominate the observed neural variability here, particularly given our observation that the largest noise mode represents a considerable fraction of the observed neural variance (Figure 3E).”

While these results provide an important control in our study, we anticipate further study of the relationship between eye movements or states, visually-evoked neural activity, and neural noise modes would identify the additional neural ensembles which are correlated with and drive this additional motor output.

#### **Reviewer #3 (Public Review):**

##### *Summary:*

*In this study, Manley and Vaziri designed and built a Fourier light-field microscope (fLFM) inspired by previous implementations but improved and exclusively from commercially available components so others can more easily reproduce the design. They combined this with the design of novel algorithms to efficiently extract whole-brain activity from larval zebrafish brains.*

*This new microscope was applied to the question of the origin of behavioral variability. In an assay in which larval zebrafish are exposed to visual dots of various sizes, the fish respond by turning left or right or not responding at all. Neural activity was decomposed into an activity that encodes the stimulus reliably across trials, a 'noise' mode that varies across trials, and a mode that predicts tail movements. A series of analyses showed that trial-to-trial variability was largely orthogonal to activity patterns that encoded the stimulus and that these noise modes were related to the larvae's behavior.*

*To identify the origins of behavioral variability, classifiers were fit to the neural data to predict whether the larvae turned left or right or did not respond. A set of neurons that were highly distributed across the brain could be used to classify and predict behavior. These neurons could also predict spontaneous behavior that was not induced by stimuli above chance levels. The work concludes with findings on the distributed nature of single-trial decision-making and behavioral variability.*

*Strengths:*

*The design of the new fLFM microscope is a significant advance in light-field and computational microscopy, and the open-source design and software are promising to bring this technology into the hands of many neuroscientists.*

*The study addresses a series of important questions in systems neuroscience related to sensory coding, trial-to-trial variability in sensory responses, and trial-to-trial variability in behavior. The study combines microscopy, behavior, dynamics, and analysis and produces a well-integrated analysis of brain dynamics for visual processing and behavior. The analyses are generally thoughtful and of high quality. This study also produces many follow-up questions and opportunities, such as using the methods to look at individual brain regions more carefully, applying multiple stimuli, investigating finer tail movements and how these are encoded in the brain, and the connectivity that gives rise to the observed activity. Answering questions about variability in neural activity in the entire brain and its relationship to behavior is important to neuroscience and this study has done that to an interesting and rigorous degree.*

*Points of improvement and weaknesses:*

*The results on noise modes may be a bit less surprising than they are portrayed. The orthogonality between neural activity patterns encoding the sensory stimulus and the noise modes should be interpreted within the confounds of orthogonality in high-dimensional spaces. In higher dimensional spaces, it becomes more likely that two random vectors are almost orthogonal. Since the neural activity measurements performed in this study are quite high dimensional, a more explicit discussion is warranted about the small chance that the modes are not almost orthogonal.*

We agree with the reviewer that orthogonality is less “surprising” in high-dimensional spaces, and we have added this important point of interpretation to our revised manuscript. Still, it is important to remember that while the full neural state space is very high-dimensional (we record that activity of up to tens of thousands of neurons simultaneously), our analyses regarding the relationship between the trial-to-trial noise modes and decoding dimensions were performed in a low-dimensional subspace (up to 20 dimensions) identified by PLS regression to that optimally preserved visual information. This is a key step in our analysis which serves two purposes: 1. it removes some of the confound described the reviewer regarding the dimensionality of the neural state space analyzed; and 2. it ensures that the noise modes we analyze are even relevant to sensorimotor processing. It would certainly not be surprising or interesting if we identified a neural dimension outside the midbrain which was orthogonal to the optimal visual decoding dimension.

Regardless, in order to better control for this confound, we estimated the distribution of angles between random vectors in this subspace. As we describe in the revised manuscript:

“However, in high-dimensional spaces, it becomes increasingly common that two random vectors could appear orthogonal. While this is particularly a concern when analyzing a neural state space spanned by tens of thousands of neurons, our application of PLS regression to identify a low-dimensional subspace of relevant neuronal activity partially mitigates this concern. In order to control for this confound, we compared the angles



between  $w_{\text{opt}}$  and  $e_1$  across larvae to that computed with shuffled versions of  $w_{\text{opt,shuff}}$  estimated by randomly shuffling the stimulus labels before identifying the optimal decoding direction. While it is possible to observe shuffled vectors which are nearly orthogonal to  $e_1$ , the shuffled distribution spans a significantly greater range of angles than the observed data, demonstrating that this orthogonality is not simply a consequence of analyzing multi-dimensional activity patterns.”

*The conclusion that sparsely distributed sets of neurons produce behavioral variability needs more investigation because the way the results are shown could lead to some misinterpretations. The prediction of behavior from classifiers applied to neural activity is interesting, but the results are insufficiently presented for two reasons.*

*(1) The neurons that contribute to the classifiers (Figures 4H and J) form a sufficient set of neurons that predict behavior, but this does not mean that neurons outside of that set cannot be used to predict behavior. Lasso regularization was used to create the classifiers and this induces sparsity. This means that if many neurons predict behavior but they do so similarly, the classifier may select only a few of them. This is not a problem in itself but it means that the distributions of neurons across the brain (Figures 4H and J) may appear sparser and more distributed than the full set of neurons that contribute to producing the behavior. This ought to be discussed better to avoid misinterpretation of the brain distribution results, and an alternative analysis that avoids the confound could help clarify.*

We thank the reviewer for raising this point, which we agree should be discussed in the manuscript. Lasso regularization was a key ingredient in our analysis; l2 regularization alone was not sufficient to prevent overfitting to the training trials, particularly when decoding turn direction and responsiveness. Previous studies have also found that sparse subsets of neurons better predict behavior than single neuron or non-sparse populations, for example Scholz et al. (2018).

While showing l2 regularization would not be a fair comparison given the poor performance of the l2-regularized classifiers, we opted to identify a potentially “fuller” set of neurons correlated with these biases based on the correlation between each neuron’s activity over the recording and the projection along the turn direction or responsiveness dimension identified using l1 regularization. This procedure has the potential to identify all neurons correlated with the final ensemble dynamics, rather than just a “sufficient set” for lasso regression. In new Figures S5F-G, we show the 3D distribution of all neurons significantly correlated with these biases, which appear similar to those in Figures 4H-K and widely distributed across practically the entire labeled area of the brain.

*(2) The distribution of neurons is shown in an overly coarse manner in only a flattened brain seen from the top, and the brain is divided into four coarse regions (telencephalon, tectum, cerebellum, hindbrain). This makes it difficult to assess where the neurons are and whether those four coarse divisions are representative or whether the neurons are in other non-labeled deeper regions. For these two reasons, some of the statements about the distribution of neurons across the brain would benefit from a more thorough investigation.*

We agree with the reviewer that a more thorough description and visualization of these distributed populations is warranted.

While the dense, pan-neuronal labeling makes the isolation of highly specific circuit components difficult, we have shown in more detail the specific brain regions contributing to these populations by aligning our recordings to the Z-Brain atlas (Randlett et al., 2015) as shown in new Figures S1E, S3F-G, 4I, 4K, and S5F-G. In addition, we provided a more detailed

parcellation of the neuronal ensembles by providing projections of the full 3D volume along the xy and yz axes, in addition to the unregistered xy projection shown in new Figures 4H and 4J. We also found that the distribution of neurons in our huc:H2B-GCaMP6s recordings is very similar to the distribution of labeling in the huc:H2B-RFP reference image from the Z-Brain atlas (new Figure S1E), which further supports our whole-brain imaging results.

Overall, we find that this more detailed quantification and visualization is consistent with the interpretations in the previous version of our manuscript. In particular, we show that optimal visual decoding population ( $w_{opt}$ ) and largest noise mode (e1) are localized to the midbrain (new Figures S3F-G), which is expected since in Figure 3 we first extracted a low-dimensional subspace of whole-brain neural activity that optimally preserved visual information. Additionally, we provide additional evidence that the populations correlated with the turn bias and responsiveness bias are distributed throughout the brain, including a relatively dense localization to the cerebellum, telencephalon, and dorsal diencephalon (habenula, new Figures 4H-K and S5F-G).

### **Recommendations for the authors:**

#### **Reviewer #1 (Recommendations For The Authors):**

*In addition to the overall strengths and weaknesses above, I have a few specific comments that I think could improve the study:*

*(1) In lines 334-335 you write that 'We proceeded to build various logistic regression classifiers to decode'. Do you mean you tested this with other classifier types as well (e.g. SVM, Naive Bayes) or do you mean various because you trained the classifier described in the methods on each animal? This is not clear. If it is the first, more information is needed about what other classifiers you used.*

We appreciate the reviewer raising this point of clarification. Here, we simply meant that we fit the multiclass logistic regression classifier in the one-vs-rest scheme. In this sense, a single multiclass logistic regression classifier was fit for each larva. We have updated our revised manuscript with this clarification: “The visual stimuli were decoded using a one-versus-rest, multiclass logistic regression classifier with lasso regularization.”

*(2) In Figure 3 you train the decoder on all visually responsive cells identified across the brain. Does this reliability of stimulus decoding also hold for neurons sampled from specific brain regions? For example, does this reliable decoding come from stronger and more reliable responses in the optic tectum, whereas stimulus decodability is not as good in visual encoding neurons identified in other structures?*

In new Figure S5B, we show the performance of stimulus decoding from various brain regions. We find that stimulus classification is possible from the midbrain and cerebellum, very poor from the hindbrain, and not possible from the telencephalon during the period between stimulus onset and the decision.

*(3) In relation to point 2, it would be good to show in which brain areas the visually responsive neurons are located, and maybe the average coefficients per brain area. Plots like Figures 3G, and H would benefit from a quantification into areas. Similarly, a parcellation into more specific brain areas in Figure 4 would also be valuable.*

In addition to providing a more detailed parcellation of the turn direction and responsiveness bias populations in Figure 4, we have provided a similar visualization and quantification of the optimal stimulus decoding population and the dominant noise mode in new Figures S3F-G, respectively.

(4) In Figure 3f, it is not clear to me how this shows that  $w_{opt}$  and  $e1$  are orthogonal. They appear correlated.

The orthogonality we quantify is related to the pattern of coefficients across neurons, not necessarily the timeseries of their projections. The slight shift in the noise mode activations as you move from stimuli on the left visual field to the right actually comes from the motor outputs. Large left stimuli tend to evoke a rightward turn and vice versa, and the example noise mode shown encodes the directionality and vigor of tail movements, resulting in the slight shifts observed.

(5) I think the wording of this conclusion is too strong for the results and a bit illogical:

*'Thus, our data suggest that the neural dynamics underlying single-trial action selection are the result of a widely-distributed circuit that contains subpopulations encoding internal time-varying biases related to both the larva's responsiveness and turn direction, yet distinct from the sensory encoding circuitry.'*

*If that is the case, how is it even possible that the larvae can do a visually guided behaviour?*

*Especially given Suppl Fig 4C it would be more appropriate to say something along the lines of: 'When stimuli are highly ambiguous, single trial action selection is dominated by widely-distributed circuit that contains subpopulations encoding internal time-varying biases related to both the larva's responsiveness and turn direction, that encode choice distinctly from the sensory encoding circuitry'.*

We appreciate the reviewer's suggestion and have re-worded this line in the discussion in order to clarify that these time-varying biases are predominant in the case of ambiguous stimuli, as shown in Figure S5C in our revised manuscript (corresponding to Figure S4C in our original submission).

(6) Line 599: typo: trial-to-trail

We thank the reviewer for noting this error, which has been corrected in the revised text of the manuscript.

<https://doi.org/10.7554/eLife.97014.2.sa0>