

University of California  
Santa Barbara

# **The Visual Influence on Zebrafish Larval Locomotion through Experimental–Computational Methods**

A senior thesis submitted in partial satisfaction  
of the requirements for the degree

Bachelors of Science  
in  
Physics

by

Eric Jiawei Zhu

Committee in charge:

Professor Sebastian Streichan, Principle Investigator  
Pieter Derksen, Graduate Student

June 2025



The senior thesis of Eric Jiawei Zhu is approved.

---

Professor Sebastian Streichan

June 2025



## Acknowledgements

I want to express my deepest gratitude to all those who have supported and guided me throughout my research experience as an undergraduate. First and foremost, I thank Professor Sebastian Streichan for his mentorship and vision for this project. His insightful guidance, rigorous scientific standards, and encouragement have been instrumental in shaping the direction of this project and my growth as a researcher.

I am grateful to Pieter Derksen, whose constructive feedback was invaluable at every stage of this work. Pieter's research expertise and willingness to discuss ideas at length greatly enhanced the quality and depth of my findings, effectively bridging the gap between biology and physics.

Additional thanks to Dr. Susie Wopat for her generous assistance with the experimental setup, the handling and maintenance of mutant zebrafish lines, and for patiently teaching me the intricacies of zebrafish husbandry.

Finally, I acknowledge the broader support of the Streichan Lab and the Animal Resource Center's staff, whose collaboration made this journey both productive and enjoyable.



## Abstract

The Visual Influence on Zebrafish Larval Locomotion through  
Experimental–Computational Methods

by

Eric Jiawei Zhu

Larval zebrafish integrate hydrodynamic and visual cues to navigate their environment, yet the developmental onset and mechanistic role of vision in shaping spatial exploration remain poorly understood. Here, we combine quantitative behavioral experimentation with computational modeling to understand how visual feedback influences boundary interactions in zebrafish larvae. We recorded individual trajectories of wild-type larvae in circular Petri dish arenas under different boundary conditions at various points in development. Using idtracker.ai for high-precision tracking, we transformed positions and velocities into radial and angular distributions. We employed kernel-smoothed two-sample Kolmogorov–Smirnov tests to detect shifts in spatial occupancy over development.

At 7 days post-fertilization (dpf), larvae exhibited uniform distributions across all arena types. By 14dpf, wild-type fish in clear arenas showed a significant migration toward the perimeter, an effect that intensified by 21dpf, where fish circled the boundary in sustained patterns. In half-sanded arenas, wild-type larvae developed strong symmetry-breaking, preferentially occupying the side where reflection was visible. In contrast, blind mutants maintained uniform distributions at all stages, indicating that visual input is necessary for boundary bias.

To examine underlying mechanisms, we extracted age-dependent speed distributions and implemented three simulation conditions: (1) a Brownian random walk model, (2) a probabilistic turning model featuring a tunable weight function over distance, and (3) a reflection-guided model using Snell’s law to sample turning directions toward perceived reflected light. The physics-based turning model accurately recapitulated edge preference

in sanded arenas and the symmetry-breaking observed in half-sanded arenas, validating the hypothesis that vision-driven reflection sampling governs boundary navigation.

Our integrated pipeline combining controlled genetic and arena manipulations with rigorous statistical analysis and flexible computational models provides a framework for understanding sensorimotor integration in zebrafish and other small organisms. Through a simple, low-parameter model, we can recapitulate the various symmetry-breaking events that occur during the development of the zebrafish visual system. These findings explicate the developmental role of visual cues in spatial behavior and offer a modular platform for future studies linking neural circuitry to emergent movement patterns.

# Contents

|  |           |
|--|-----------|
| <b>Abstract</b>                                  | <b>iv</b> |
| <b>1 Introduction</b>                            | <b>1</b>  |
| 1.1 Animal Behavior . . . . .                    | 1         |
| 1.2 Physics-Based Approach . . . . .             | 2         |
| 1.3 Zebrafish: A Model Organism . . . . .        | 4         |
| 1.4 Larval Development . . . . .                 | 6         |
| 1.5 Zebrafish Visual System . . . . .            | 7         |
| 1.6 Analysis of Existing Models . . . . .        | 8         |
| 1.7 Experimental Setup . . . . .                 | 9         |
| <b>2 Results</b>                                 | <b>12</b> |
| 2.1 Initial Studies . . . . .                    | 12        |
| 2.2 Blind Fish . . . . .                         | 16        |
| 2.3 Simulation . . . . .                         | 20        |
| <b>3 Discussion</b>                              | <b>28</b> |
| 3.1 Vision Drives Larval Locomotion . . . . .    | 28        |
| 3.2 Future Work . . . . .                        | 30        |
| 3.3 Conclusion . . . . .                         | 31        |
| <b>A Methods</b>                                 | <b>32</b> |
| A.1 Zebrafish Handling and Maintenance . . . . . | 32        |
| A.2 Injections . . . . .                         | 33        |
| A.3 Experimentation . . . . .                    | 33        |
| A.4 Tracking . . . . .                           | 34        |
| A.5 Statistical Analysis . . . . .               | 35        |
| A.6 Simulation . . . . .                         | 38        |
| <b>Bibliography</b>                              | <b>44</b> |



# Chapter 1

## Introduction

### 1.1 Animal Behavior

Studying animal behavior helps us understand how living organisms interact with their environment and with one another. By observing animals closely, we gain valuable insights into how they navigate challenges, such as finding food, evading predators, and adapting to changing climates. Through their behaviors, animals provide clues about sensory processes and cognitive abilities that dictate how they perceive the world. Research on animal behavior sheds light on the development of decision-making processes in animals, enabling them to balance risks and benefits in everyday survival. It helps us uncover how animals communicate through sounds, gestures, chemical signals, and complex social cues [1]. Additionally, examining animal interactions provides a deeper understanding of social structures, revealing why some species thrive alone while others depend heavily on family units or large groups.

Behavioral research offers valuable insights that can significantly inform human societies, providing us with profound lessons on cooperation, competition, and coexistence. By studying the adaptability and resilience of various animal species, we can gain a deeper understanding of how these traits contribute to their survival and thriving in diverse environments.

Ultimately, by integrating these lessons into our environmental efforts, we can foster healthier ecosystems that benefit all living organisms. This holistic view not only enriches our knowledge but also empowers us to take informed actions that support both wildlife and human communities, ensuring a sustainable future for generations to come.

## 1.2 Physics-Based Approach

However, there are still many gaps in our understanding of animal behavior. A physics-based approach to studying animal behavior is beneficial and essential for advancing our understanding of how living organisms interact within their environments, providing an additional lens of exploration. This methodology transcends mere descriptive analysis, offering concrete, quantitative rules and patterns that govern animal movements and interactions. Using mathematical frameworks and fundamental physical principles, we can develop robust models that accurately predict animal behaviors across a wide array of ecological contexts [2, 3].

Understanding group behavior through the lens of physics enables us to conceptualize these dynamics as a physical dynamical system, allowing for the derivation of generalized laws that govern collective actions in animal groups. This shift is crucial, as it facilitates precise hypothesis testing via controlled experiments and computational simulations, turning behavior into a measurable and predictable phenomenon. As a result, we unlock more profound insights into the intricate ways living organisms interact with their surroundings. This approach offers a significant advantage over traditional qualitative analyses, which are often hampered by subjective visual interpretations and speculative assumptions [4].

One of the most profound benefits of applying physics to the study of animal group behavior is the systematic reduction of anthropocentric biases. Physics-based methodologies rely on empirically derived forces, interactions, and measured data, allowing us to understand complex animal behaviors without projecting human-like characteristics.

For instance, the Vicsek model employs simple force-based interaction rules to elucidate collective phenomena, such as bird flocking, demonstrating how individual physical movements shape large-scale group dynamics [5]. Hydrodynamic models based on the Navier-Stokes equations similarly clarify how aquatic organisms, such as fish, interact with their fluid environments and fellow group members [6].

Furthermore, physics-inspired frameworks comparable to this one extend their predictive capability to various organisms, encompassing insect swarming, bird flocking, and mammal herding [7, 8]. By integrating physics with biological observations, we enrich our understanding and create avenues for systematic experimentation, allowing conditions to be varied rigorously while robustly quantifying how animals respond, individually and collectively, within their environments. A physics-based lens is imperative for studying animal behavior, as it ultimately transforms our investigation into a realm of measurable predictions, illuminating the complexities of life itself.

Beyond informing the biological sphere, these systems also serve as a testing ground for understanding new concepts in physics, specifically in the rapidly growing field of active matter physics. In this realm, we view animals like larval zebrafish as examples of active material, collections of self-driven agents that consume energy to move, interact, and adapt. Unlike traditional passive materials, which respond predictably to external forces, active matter is inherently out of equilibrium, often displaying emergent phenomena such as swarming, flocking, or spontaneous symmetry breaking.

Active matter physics is a theory-driven discipline that seeks to develop minimal models capturing the essence of these complex, non-equilibrium dynamics. However, these models require experimental validation to remain grounded in reality. Biological organisms offer this opportunity: their movements and interactions are physically constrained, but not rigidly so, allowing researchers to test and refine theories in a setting that combines tunable control with ecological richness.

Zebrafish, in particular, embody this intersection. They are intelligent, deformable bodies that exert forces on their environment while also responding to sensory inputs.

As such, they operate at the interface between active particle physics and embodied cognition. Studying them allows us to explore how internal drives and environmental feedback shape collective motion and decision-making. In doing so, these biological systems not only enrich our understanding of behavior but also inform and challenge the foundational assumptions of active matter physics, offering insight into how local rules give rise to global organization in complex, driven systems.

Building on these insights, researchers are also exploring the implications of these universal principles for various fields beyond biology. For instance, understanding the dynamics of animal movement can enhance robotic design, enabling engineers to mimic natural behaviors to create more efficient and adaptive machines. Additionally, these findings can inform conservation strategies by shedding light on how animals interact with their environments and with one another, ultimately aiding in the preservation of biodiversity.

By bridging physics, biology, and technology, we can gain deeper insights into the intricate patterns of life, leading to innovative solutions that benefit both nature and human development. This interdisciplinary collaboration enhances our comprehension of biological phenomena and paves the way for advancements that resonate across multiple domains.

### 1.3 Zebrafish: A Model Organism

To rigorously test these principles, we turn to tractable model organisms that combine experimental control with well-characterized genetics and neurobiology. By starting with a system that allows us to manipulate the organism through genetic tools and the environment through controlled tank setups or stimuli, we can directly compare observed behavior with our physics-driven conclusions.

Zebrafish have emerged as an exceptionally valuable model organism for studying animal behavior, due to their unique combination of features that facilitate research across

genetics and developmental biology. One of their most distinguishing characteristics is their rapid development; from embryo to free-swimming larva in just a matter of days, this accelerated life cycle allows us to observe critical developmental stages in a fraction of the time required by other species [9].

Moreover, the transparency of zebrafish larvae is a significant advantage. This transparency enables scientists to visualize internal processes and dynamics without the need for invasive techniques, making it possible to record high-resolution movement data and even capture neural activity in real-time. Such capabilities enable us to better understand the relationships between physical interactions and biological responses, effectively bridging the gap between theoretical predictions and empirical observations.

Another benefit of zebrafish is their genetic makeup. These creatures possess a relatively simple yet sophisticated nervous system, with essential similarities to those of higher vertebrates. This complexity places them in an optimal position where understandable structures facilitate the manifestation of advanced behaviors and rapid responses. These behaviors make them particularly apt for studying behavioral responses to various stimuli or interventions.

Additionally, zebrafish are prolific spawners. With just a few carefully planned crosses, we can obtain hundreds of embryos at once, providing a rich resource for experimentation. This high yield maximizes research efficiency and supports a wide array of experimental designs, from developmental biology to pharmacological testing and beyond [10].

The combination of rapid development, transparency, and ease of breeding makes zebrafish an ideal model organism for unraveling the complexities of behavior and development. Their utility in research continues to grow, providing critical insights that can be applied to understanding more complex systems, including human biology.

## 1.4 Larval Development

Our research delves deeply into the vital early stages of zebrafish development, focusing on the larval phase that spans from embryonic hatching at 3 days post-fertilization (dpf) to the significant transition into the juvenile stage by 21dpf. This larval period is not just a growth phase; it is a crucial time for observing how foundational developmental processes shape behavior. Young zebrafish display complex behaviors such as schooling, foraging, and responding to environmental stimuli [9].

Studying larval zebrafish offers unique advantages. One remarkable aspect is our ability to introduce genetic mutations via embryonic injections, which allows us to dissect the effects of specific traits on behavioral development. For example, mutations that lead to blindness can be employed to investigate how the absence of specific sensory inputs influences movement and social interactions. Given the extensive behavioral research conducted with zebrafish as a model organism, we understand how various mutations impact vital functions such as movement and sensory processing. The advent of CRISPR/Cas9 technology further enriches our approach, providing tools to knock out specific genes effectively. This capability grants us unprecedented access to studying the onset and consequences of genetic mutations concerning zebrafish behavior, especially when compared to control wild-type zebrafish [11].

Throughout the larval stage, zebrafish rely significantly on their lateral line system for movement, particularly until around 5dpf, when their physical development is still in its infancy. At this juncture, while their eyes are developed, they do not yet depend on their vision for navigation. As foraging behavior emerges, zebrafish increase their movement and engagement with their environment. Their motivation to move is primarily driven by foraging needs in controlled experimental settings free from predation threats. This safe environment fosters rapid development of their locomotor skills, coinciding with the maturation of their visual systems [12].

## 1.5 Zebrafish Visual System

By 3dpf, the zebrafish visual system has made remarkable strides, forming five main cell layers within the retina, which initiate basic visual responses, such as startle reactions and reflexive eye movements. As the days progress, the critical cell types within the retina mature, segregating various visual signals, including small moving targets and looming threats. By 7dpf, the complex neural pathways begin to facilitate sophisticated actions, such as prey capture, escape responses, and phototaxis. This rapid development of the larval visual system illustrates how zebrafish quickly establish the neural circuitry necessary for turning visual input into coordinated movements [13, 14].

The anatomical positioning of zebrafish eyes on the sides of their heads grants them an expansive field of view, critical for both foraging and predator avoidance. Studies of zebrafish have profoundly enhanced our understanding of visual processing in the brain, the functionality of vision-related genes, and the mechanisms underlying eye diseases. Thus, the larval zebrafish stage not only represents a critical window for exploring behavioral development but also offers rich insights into the complexities of sensory perception and adaptation in a living system.

While zebrafish are developing their complete visual systems, they are already exhibiting complex swimming behaviors dependent on their lateral line's ability to sense the hydrodynamic environment effectively. Their swimming involves generating undulatory waves along their bodies, producing thrust and turning essential for various activities, including foraging and predator evasion [15]. Studies have shown that during swimming, zebrafish create specific flow patterns in the water, such as vortex rings, which are crucial for propulsion and stability. Body curvature and tail beat frequency influence the hydrodynamic interactions, which change as the fish grows, affecting swimming efficiency and maneuverability. Additionally, zebrafish adjust their group behaviors in response to water flow, where in some cases, zebrafish present in slow water flow can lead to less cohesive group dynamics and more aggressive interactions. These results underscore the importance of hydrodynamic forces in understanding changes in zebrafish behavior [16].

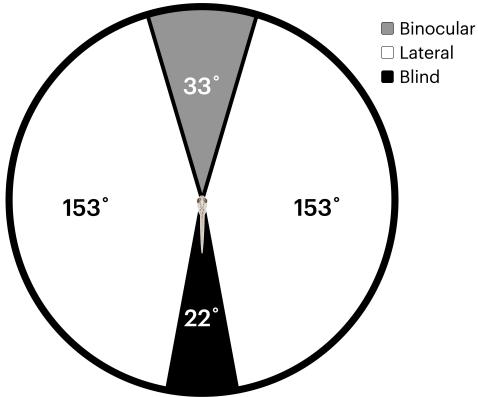


Figure 1.1: Zebrafish exhibit strong lateral vision, with predominantly lateral vision on both sides, covering  $306^\circ$  of their total visual field. In front, there is a  $33^\circ$  window of binocular vision, while they only have a slim  $22^\circ$  blind spot behind them.

Towards the end of the larval stage, zebrafish begin using their eyes as part of the sensory input that coordinates their movement. While previous research has explored aspects of zebrafish locomotion, a significant gap remains in understanding how hydrodynamic sensing through the lateral line and optical sensing through the eyes interact to influence movement dynamics. Our work aims to fill this void by leveraging these sensory behaviors to produce an effective model of basic zebrafish movement patterns, ultimately contributing to a more comprehensive understanding of their sensory integration.

## 1.6 Analysis of Existing Models

Researchers have developed various models to understand, simulate, and predict zebrafish behavior, enhancing our understanding of their complex actions. We aim to model previous data-driven workflows that utilize exploratory data analysis of position distributions and inter-point correlations to identify key movement features and detect sudden relocations. This structure builds more realistic step-length models by estimating correlated, non-Gaussian noise through copulas and kernel density estimates. Insights

from this exploratory data analysis inform a Langevin-style model for individual paths, which can be extended to simulate various animals [4].

Another study on zebrafish classified each bout into a small set of types and represented larval behavior as a time-ordered sequence of labeled actions. The study then developed probabilistic models that predict which bout type occurs and how long the fish rests beforehand, taking into account the larva's hunger level, recent actions, and the position and characteristics of nearby prey. The models were validated by demonstrating that they accurately fit test data and generate realistic virtual fish trajectories in simulated environments. The simulations reproduced behavioral patterns across multiple timescales and revealed apparent shifts in actions, with selection rules favoring food-seeking when hungry and risk-averse movements when satiated [17].

Researchers have developed mathematical models to simulate the social behavior of zebrafish in response to psychoactive substances, helping to understand how these compounds affect group dynamics [18]. These modeling efforts are crucial for advancing our knowledge of zebrafish behavior and have applications in fields such as neuroscience, pharmacology, and robotics[19]. However, there are many gaps in understanding the development of zebrafish and its integration of the visual system into movement. We look to build upon these models, leveraging the understanding of the visual system, and we set out to simultaneously leverage these factors through a simulation. We aim to isolate movement behavior that is solely dependent on our controlled factors, thereby limiting potential influence from fluctuations in water temperature, hunger, and time of day [20].

## 1.7 Experimental Setup

We combine knowledge of zebrafish development and behavior with experimental data to inform our predictive model of movement patterns across the larval development timeframe. We can modify factors such as obstacles, movement conditions, or social

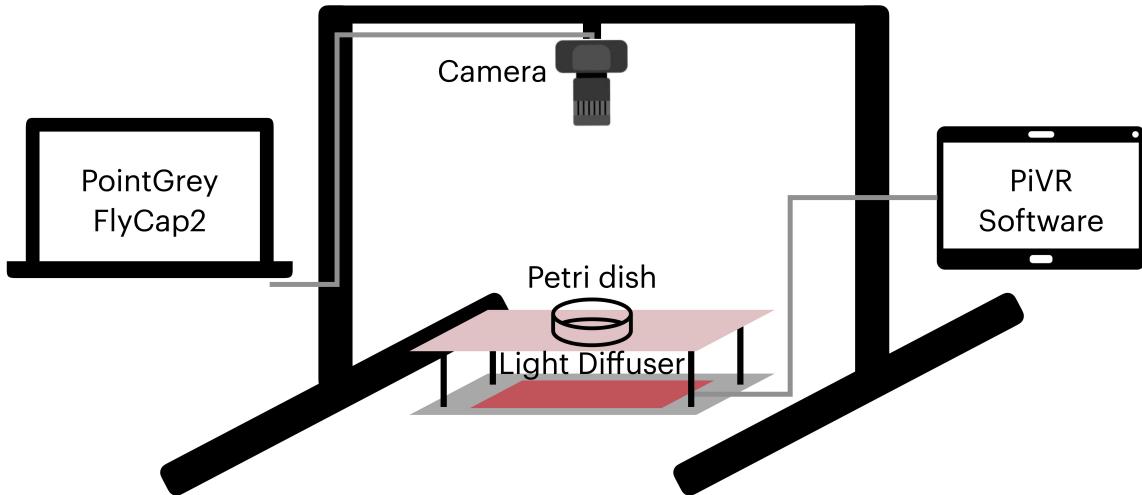


Figure 1.2: Our experimental setup consists of running the PiVR software through the tablet to control the power of the red light emitted through the ceramic plates. We placed a Petri dish on the top ceramic plate and recorded from above using a Teledyne FLIR Grasshopper3 USB3 camera with a Zeiss Milvus 85mm f/1.4 lens. We connected the camera setup to the PointGrey FlyCapture2 SDK, where AVI file-format videos were recorded and saved onto a hard drive.

influences through computer simulations to observe the effects [8, 17]. We aim to develop the simplest prediction framework to minimize the potential for overfitting our collected data. Simulations also allow us to isolate variables that are difficult to control in live experiments, such as external water flow or social influences. We strive to systematically investigate the contributions of vision and hydrodynamic sensing to behavior by adjusting parameters within a computational model.

For this, we constructed our experimental setup to utilize the PiVR software, which provides even light from below, much stronger than that of the surrounding room [21]. Previous experiments have confirmed that zebrafish are not sensitive to red light in the same way they might be towards lower wavelengths; therefore, increasing illumination for our recordings would have minimal impact on their behavior. Thus, we placed our Petri dishes, which will house the zebrafish, directly in the center of the PiVR's ceramic

plates, allowing us a consistent framework for collecting the data (Figure 1.2). Similarly, we can work our simulations around this framework, implementing them with circular boundary conditions.

Our work on zebrafish utilizes a physics-based approach to study their behavior, providing clear insights into their interactions and responses to environmental changes. By applying quantitative models to zebrafish behavior, we aim to unravel the complex systems that govern their social interactions, revealing how these fish adapt to varying ecological conditions.

# Chapter 2

## Results

### 2.1 Initial Studies

To investigate how larval zebrafish interact with boundaries of different visual properties, we examined their swimming patterns in three types of tanks. For all our tests, we used a 100mm Petri dish, approximately twenty times the length of a larval zebrafish body, providing sufficient room for movement. We use the Petri dish as is for the first condition, serving as the control condition. We placed zebrafish in a dish with sanded borders for the second condition to block reflections and produce a physical visual barrier. In our final condition, we sanded down just half of the dish to create a condition that breaks symmetry and test whether any asymmetric behavior ensues.

Across all trials and visual conditions, we visualize their spatial distributions to analyze age-dependent behavior shifts and consider the differences in each visual condition at any given age (Figure 2.1). We begin by comparing each visual condition across larval development at 7, 14, and 21dpf to identify any notable differences in spatial distributions or symmetry-breaking in the half-sanded condition.

At 7dpf, the heatmaps indicated a relatively uniform occupancy throughout the dish, with no pronounced clustering in the center or along the edges. By 14dpf, a subtle yet noticeable change emerged in the heatmaps, with several bins near the perimeter

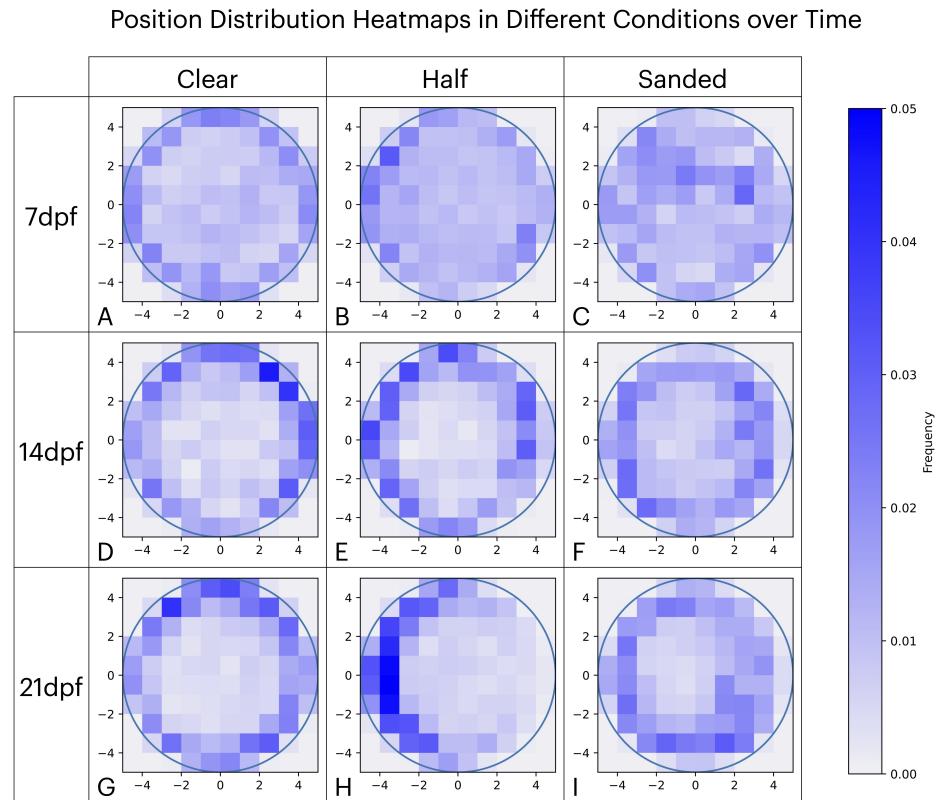


Figure 2.1: From the iteratively tracked videos, we can plot the spatial distribution of each condition (clear, sanded, and half-sanded) across various time points (7, 14, and 21dpf). We sand the half-dish on the right half of the dish as represented by the right half of all three plots (B, E, H). We scaled all plots to the same dimensions (10 bins by 10 bins) with a shared frequency color bar to label distribution intensity. For the sanded dishes (C, F, I), there is no significant change in behavior. In contrast, across the clear dishes (A, D, G), a shift towards the boundary occurs over time. The behavior at 7dpf (A, B, C) is comparable across all three conditions. The main conclusion of these results is evident in the half-sanded dish (B, E, H), where, by 21dpf, a clear symmetry breaking occurs, resulting in a strongly skewed distribution away from the sanded side.

exhibiting a higher frequency of occupation. The overall shift towards the boundary was consistent across all samples, directing a change toward edge exploration in the clear dish. Despite this emerging trend, some zebrafish continued to traverse or linger in central regions, rather than a complete shift towards the boundaries. Looking specifically at the symmetry-breaking case of the half-sanded dish, we notice that the clear side consists of a similar distribution to the default clear dish. In contrast, the sanded side portrays a similar distribution to the fully sanded dish, with a stronger distribution near the boundaries overall.

At 21dpf, the zebrafish continues to prefer the transparent boundary edges. The spatial distribution shown on the heatmaps depicts a strong preference towards the perimeter, and the contrast between edge and center squares increased. Many zebrafish in the clear dish begin to traverse around the edges of the boundary, following the circular shape of the edges. At some points, they choose to turn around and circle the other way, making minimal movements within the central regions of the dish. Even when these fish enter this region, they often traverse quickly before returning to the boundary, as exemplified by examining single zebrafish position plots (Figure 2.7). This pattern suggests that the older larvae spent much of their time near the clear, unsanded boundaries of the dish, with correspondingly reduced occupancy in the middle of the tank.

For any two distributions, we quantitatively compare them pairwise, utilizing the 2-sample Kolmogorov-Smirnov test (KS-test) to determine the similarity between both the radial and angular distributions. We then produce plots of the p-value and D-statistic distributions (see Appendix A for additional details). Using the KS-test to output p-values and D-statistics, we can compare the radial and angular distributions in all three conditions at 7dpf and confirm that there are no statistically significant differences in the overall position distributions (Figure 2.2). Similarly, for 21dpf, we confirm the statistically significant radial distributions between clear and sanded dishes by conducting the KS-test on pairwise distribution comparisons.

Meanwhile, as expected, we can confirm that the angular distributions are similar,

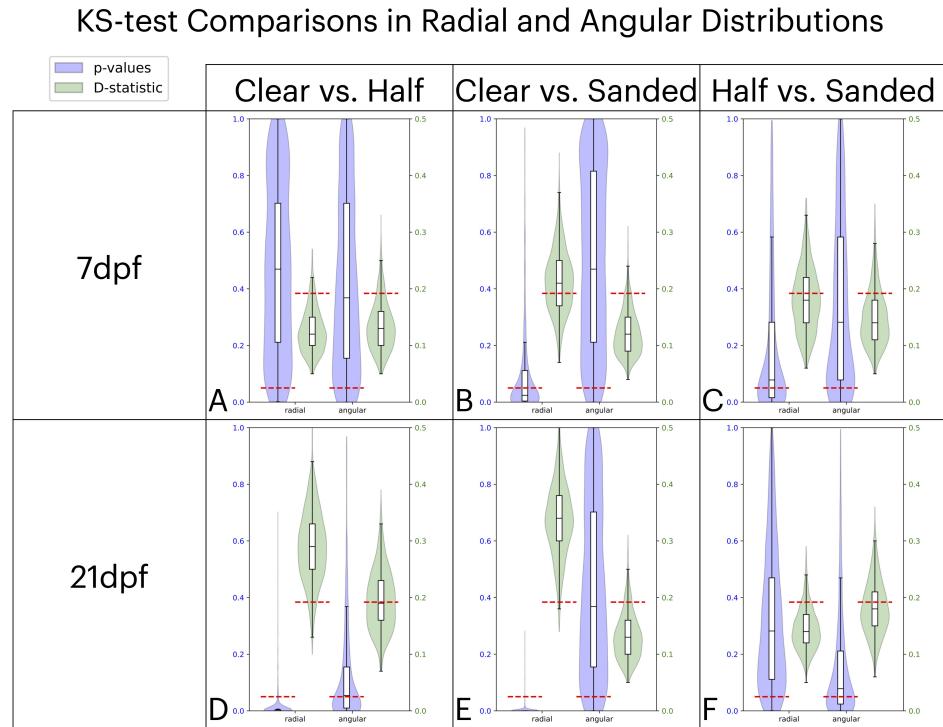


Figure 2.2: Given all of our radial and angular distributions from our experimental data across our three dish conditions at 7 (A, B, C) and 21dpf (D, E, F), we can pairwise compare these distributions using the Kolmogorov-Smirnov test for comparing sample distributions. For statistically significantly different distributions, the median p-value should be less than the critical value for p-values (lower dashed line), and the median D-statistic should be above the critical value for D-statistics (higher dashed line). Using the KS-test helps us qualitatively visualize the similarities and differences in distributions. In this case, we compare pairwise within the three conditions for 7dpf and 21dpf by comparing the clear dish with the half-sanded dish (A, D), the clear dish with the sanded dish (B, E), and the half-sanded dish with the sanded dish (C, F).

which helps us conclude that random factors, such as uneven lighting conditions, do not contribute to the unusual symmetry-breaking. Similarly, the KS-test confirms statistically different angular distributions for the half-sanded symmetry-breaking case (Figure 2.2).

Once again, looking specifically at the half-sanded dish, there is now a strong bias towards the default, unsanded side. The distribution shifts in phases initially at 14dpf towards the boundary, and then by 21dpf, there is another shift towards the unsanded side of the dish. We confirmed this by iterating through individual zebrafish's paths, where each zebrafish maintains similar behavior. The KS-test comparison between time points reveals a change in the radial distribution between 7 and 14dpf, followed by an angular distribution change between 14 and 21dpf.

The zebrafish's radial distribution in the sanded dish remains unchanged as it matures, suggesting that the development of its visual system does not affect how the zebrafish maneuvers in the sanded dish. We can confirm these conclusions by pairing the KS-test between the two sample distributions at the different time points (Figure 2.3).

In contrast, the clear dish distributions shift towards the boundary from 7dpf to 14 and 21dpf. This preference towards the boundary has mainly developed by 14dpf, with no drastic shift from 14 to 21dpf (Figure 2.3). Thus, there are apparent differences in the clear and sanded dishes associated with the parallel development of the visual system.

## 2.2 Blind Fish

To determine whether this symmetry-breaking results from the visual system, we obtained CRISPR/Cas9 knockouts of the *TYR* gene, which produce zebrafish with impaired vision. As a result, the factors that drive the visual system to play a role in movement are no longer present. In all three dishes, the hydrodynamic properties that the zebrafish experiences in each environment are the same, which is unlikely to contribute to the symmetry-breaking. We then collected the same data to directly compare the wild-type

### KS-test Comparisons in Radial and Angular Distributions

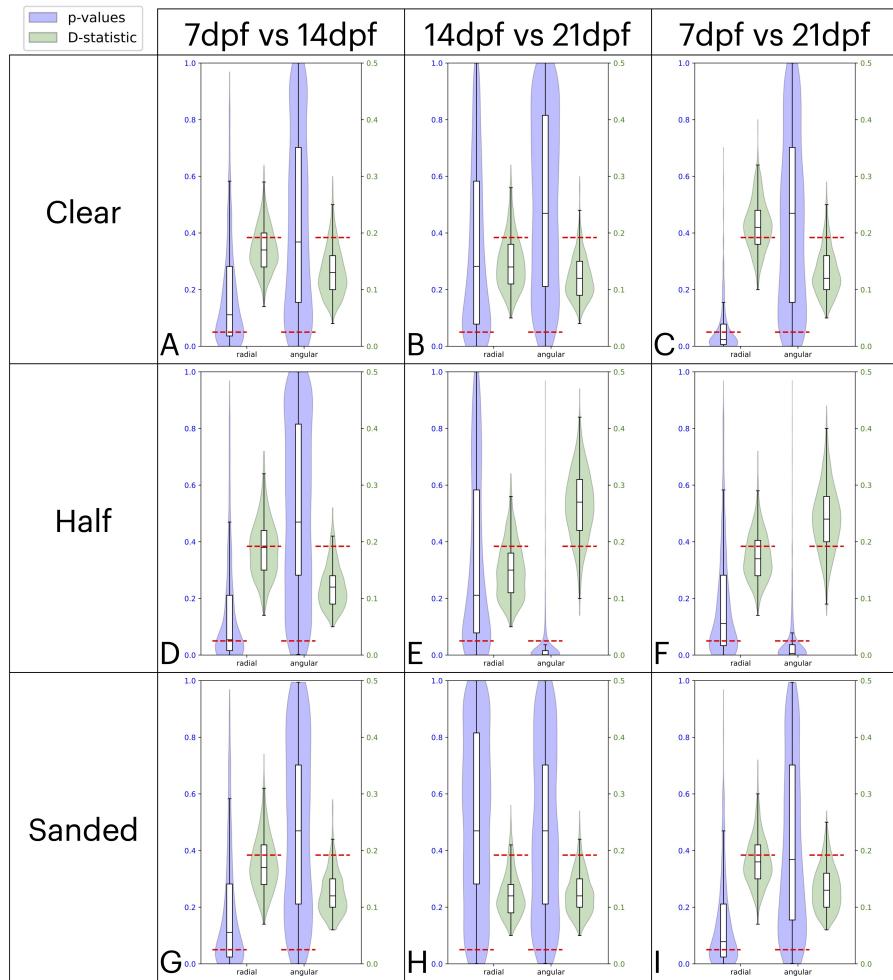


Figure 2.3: Given all of our radial and angular distributions from our experimental data across our three dish conditions at 7, 14, and 21dpf, we can pairwise compare these distributions using the Kolmogorov-Smirnov test for comparing sample distributions. Thus for each of the clear (A, B, C), half-sanded (D, E, F), and sanded (G, H, I) conditions, we analyze the similarities or differences in distribution between 7 and 14dpf (A, D, H), 14 and 21dpf (B, E, H), and 7 and 21dpf (C, F, I). For statistically significantly different distributions, the median p-value should be less than the critical value for p-values (lower dashed line), and the median D-statistic should be above the critical value for D-statistics (higher dashed line). Using the KS-test helps us qualitatively visualize the similarities and differences in distributions.

Blind Position Distribution Heatmaps in Different Conditions over Time

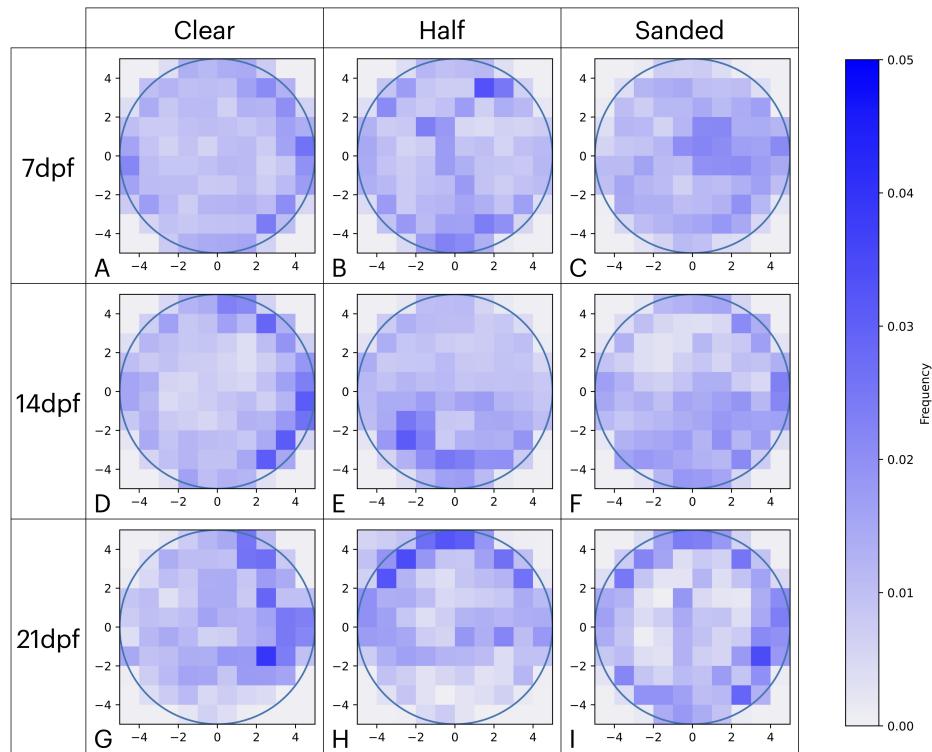


Figure 2.4: From the iteratively tracked videos of CRISPR/Cas9 tyr-knockout zebrafish, we can make the same spatial distribution plots of each condition (clear, sanded, and half-sanded) across various time points (7, 14, and 21dpf). Again, we only sand the half-sanded dish on the right half of the dish, as reflected in all three plots. In this case, there are no significant changes in behavior across all plots. The sample size of this dataset is smaller, resulting in some random distribution; however, there is no consistent behavior across the time points and different conditions.

results. After collecting five iterations of the same data, we produced the same data visualizations (Figure 2.4). The lack of complete symmetry in the positional distributions is likely due to the smaller sample size and the slower movement speed exhibited by the mutated fish.

A slower movement speed within tyr-knockout zebrafish can lead to a singular zebrafish dominating a spatial distribution. As a result, the zebrafish would cover less ground in the same time frame, equivalent to collecting a much smaller sample of the wild-type zebrafish. Thus, there is far less symmetry within the distributions despite a comparable number of iterations. However, the previous data is still statistically significant given the extreme symmetry-breaking in wild-type fish. These differences in symmetry would be insignificant when considering the half-sanded dish. There is no uneven distribution between the left and right sides, while the physical boundary condition does not directly correlate to the lack of symmetry. By iterating through individual zebrafish's paths again, we notice that specific fish that move minimally throughout the recording cause some of the peaks in the spatial distribution, leading to spikes in the heatmaps. No consistent symmetry-breaking or significant behavioral observation occurs consistently from fish to fish, in contrast to the behavior of the wild-type fish.

When comparing the sanded-dish distributions of both the blind fish and the control fish directly, we observe similar distributions with minimal changes between the different time points. In contrast, the preference for the boundary is no longer present in the clear dish, indicating some reliance on the visual system in zebrafish. When looking at the 21dpf clear-dish distribution, there is a shift away from the boundary, which is the opposite effect of the wild-type fish. There is no development or change in behavior when comparing dishes and timepoints, since previous studies have asserted that zebrafish do not strongly rely on the visual system at 7dpf, the spatial distributions across the board are the most symmetric and similar to that of the wild-type zebrafish, potentially pointing towards the fact that the factors driving movement at 7dpf across wild-type and tyr-knockout zebrafish are predominantly the same. By 21dpf, the expected reliance on

the visual system in movement is no longer present, which could lead to slower movement speeds due to a greater reliance on hydrodynamic sensing. In contrast, wild-type zebrafish exhibit an increasing speed distribution as they age.

## 2.3 Simulation

We set out to utilize the data we collected to produce accurate simulations of the zebrafish. We simulated zebrafish movement, starting with the basic Vicsek bird flocking model and adapting it to fish behavior. We introduced a Markov model that alternates between "moving" and "resting" states, based on experimental data, to capture zebrafish behavior more accurately. Our flexible simulations allowed us to test different boundary conditions, comparing them to a baseline Brownian motion model with simple wall interactions, which served as the starting point for zebrafish.

### 2.3.1 Brownian Random Walk Model

We utilize the physical results to inform the parameters we embed into our simulations. For this, we extract the speed distributions from all the data at different ages, where wild-type zebrafish begin to move faster as they age. From this, we start with simple simulations, using a Brownian random walk model. Once the Brownian zebrafish reaches a point where it cannot complete a random walk, it remains idle in that location until it randomly picks a possible movement direction at a future time point. We can then compare these distributions to the 7dpf wild-type data to see if the random walk accurately represents the zebrafish's initial movement patterns when in its most immature state. In addition, we hope to compare this baseline to that of the sanded distribution from 7 to 21dpf, which we concluded does not change statistically significantly across our experimental results.

Similarly, we generated speed histograms from the blind fish data to produce a similar Brownian random walk across the time points (Figure 2.5). This information helps us

Simulations for Brownian Random Walk Model Using Given Speed Distributions

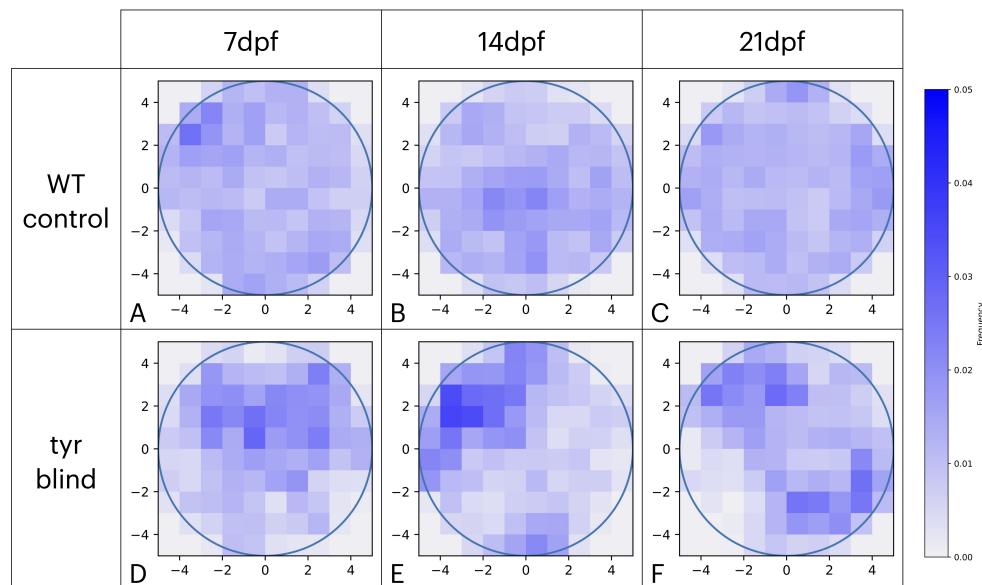


Figure 2.5: Given our experimental data, we extracted the speed distributions at each time point for the control wild-type fish (A, B, C) and the tyr-knockout blind fish (D, E, F). We use those speed distributions to produce heatmaps of the position distributions from a Brownian random walk simulation. There is less symmetry due to slower speeds from the tyr-knockout blind distributions, leading to more random behavior across ten 900 timepoint iterations.

test the hydrodynamic behavior of the zebrafish, examining how much they are impacted by the approaching boundary conditions, which drives a turn away from the boundary in their behavior, primarily through the lateral line. Upon comparing the distributions, we found that using the speed distributions of the blind fish to complete a random walk yields similar levels of asymmetry in the distribution. Meanwhile, when we use the wild-type speed distributions to simulate the same random walk, we see a more symmetric distribution across the 10 iterations. This simulation serves as an effective baseline for modeling behavior in the sanded condition, while simultaneously confirming some of the randomness in distribution observed in the physical blind fish data.

### 2.3.2 Single Parameter Probabilistic Turning Model

An alternative model we utilize to develop the rest of our simulations starts with the assumption that the zebrafish moves forward with some noise in orientation in the sanded visual condition. We implemented boundary interactions, initially treating the fish as bouncing particles and refining the model with a gradual turning mechanism. The main parameter we embed in our simulations is the turning parameter, which we can iterate through different values to determine the turning probability distribution, or effectively, how close to the boundary a zebrafish begins to consider the boundary. Most importantly, we aim to replicate the apparent symmetry breaking in the half-sanded simulation. Meanwhile, we use the turning probability function at any given moment to determine if the zebrafish decides to turn (see Appendix A Methods for details).

Once the simulated zebrafish completes a turn, we randomly select a turn angle in the direction of least resistance, turning towards the equivalent of a bounce off the wall. We can use this condition to tune our probability function and noise ratio to reproduce the distributions of the sanded dish. We introduce a noise ratio of 0.3, which means that we add a random noise vector proportional to the noise ratio of the velocity vector to the velocity at every time. This random noise vector enables the simulated zebrafish to exhibit random movements that parallel the experimental zebrafish behavior. From this,

we can scan various turning probability function constants to determine what closely matches the zebrafish swimming patterns. We did not want to overfit the simulation to the experimental data, focusing on a rough comparison between the sanded experiment and simulation. As a result, we settled on a turning probability function constant of 10, which means that the zebrafish only chooses to turn when it is, for the most part, within 2 centimeters of the boundary.

We will use the turning constant determined by our sanded dish condition simulation to prevent additional overfitting from custom turning parameters across all our simulations, both in time and visual conditions. After completing these comparisons for the sanded condition, we generate the simulated data and can visualize it similarly to the physical data (Figure 2.6).

### 2.3.3 Reflection-Following Probabilistic Turning Model

Thus, we can utilize the sanded dish simulation as the backbone for the sanded half of the half-sanded dish simulation. To complete this, we utilized the simulation of the default clear dish, which randomly selects a reflection location to sample a turning direction. We will continue to use the same turning parameter and noise ratio for all the simulations. For transparent tank walls, we modeled zebrafish responses to reflections using Snell’s law to randomly turn towards borders where light is reflected (see Appendix A Methods for details). We primarily focus on the 21dpf distributions for these simulations because those fish in different visual conditions exhibit the most substantial differences in behavior. An important note about the clear dish is that it operates similarly to the sanded dish, provided the zebrafish sees no reflection anywhere in the tank, which we base on the hydrodynamic similarities in all these experimental conditions. We then run the same iterations as the sanded dish using the equivalent distributions to produce the same plots (Figure 2.7).

Finally, we can use the clear, unsanded dish simulation to produce the half dish by limiting where the zebrafish can see reflection, dropping any point on the right half of

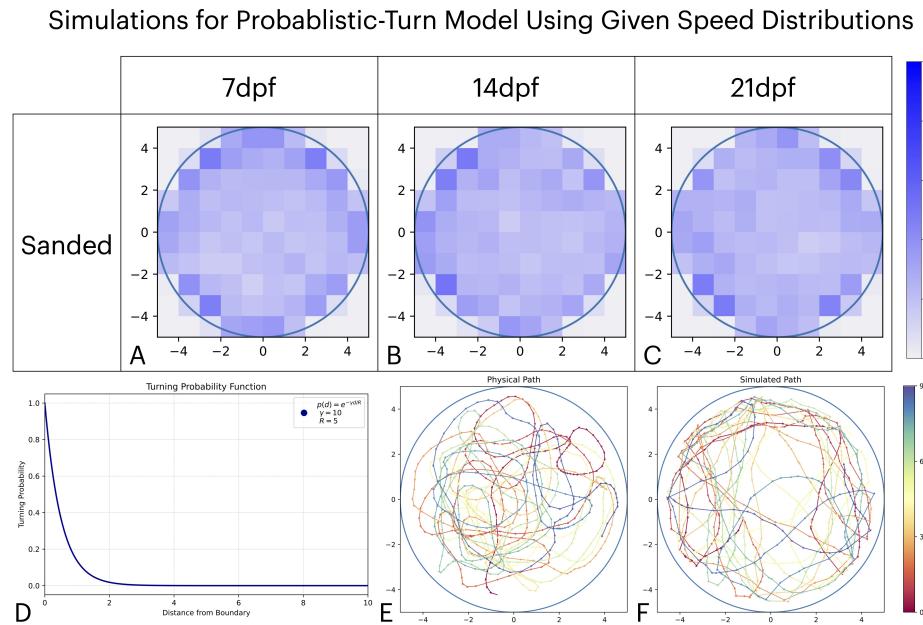


Figure 2.6: We simulate our sanded dish using the speed distributions from the respective experimental data at each time point. We then apply our turning probability function, which has an exponential turning parameter of 10 (D), determining the likelihood of turning based on the straight distance to the boundary. Combining the depicted function with normal movement and some noise level, we can produce simulated paths of a single zebrafish (F). We can compare it with an arbitrary experimental zebrafish (E) to see that our simulation produces development movement patterns, albeit with sharper turns. From this, we take ten 900 timepoint observations for each stage of development to again produce 2D spatial distribution heatmaps (A, B, C) representing our simulated sanded dish.

the dish as a possible point of reflection. This combination effectively creates a sanded dish movement pattern on one half and a clear dish movement pattern on the other half, hopefully resulting in a symmetry-breaking effect.

Under these conditions, we can compare the radial and angular distributions with those of the physical data to confirm our ability to replicate the symmetry breaking that occurs by 21dpf. We observe similar position distributions to those in the physical data, using the same turning constant across speed distributions. Compared to the Brownian motion simulation, specifically for the sanding distribution, the turning parameter-based distributions are more statistically similar to the physical data than the Brownian walk model, pointing towards an improvement in simulating the physical data.

Examining the heatmap distributions, we observe that our simulation accurately reproduces the general patterns observed in the experimental data at 21dpf, including both the zebrafish's preference for the border in the clear dish and the strong symmetry-breaking that occurs in the half-sanded dish. Going further, we can again compare the paths of singular zebrafish, extracting an arbitrary simulated and experimental fish. We observe a similar pattern of movement among all the fish. Still, the simulated data is not as smooth for the clear and half-sanded dishes as the experimental data, identical to our earlier conclusions from the sanded dish (Figure 2.7).

As we did with the experimental data, we can utilize the KS-test to compare our simulated data with the corresponding experimental data. We hope that we see that the distribution of the p-values and D-statistic is not statistically significant. From the KS-test, our sanded and clear simulations are strongly correlated, meaning the simulation effectively matches the experimental data. For the half-sanded condition, the radial distributions are still statistically significantly different. Still, more importantly, we can replicate the angular distribution to be statistically indistinguishable from the experimental data (Figure 2.8). Thus, through our noise-embedded turning probability function model, which incorporates turning towards the direction where the zebrafish sees reflected light on the boundary, we successfully replicated each of our three tested

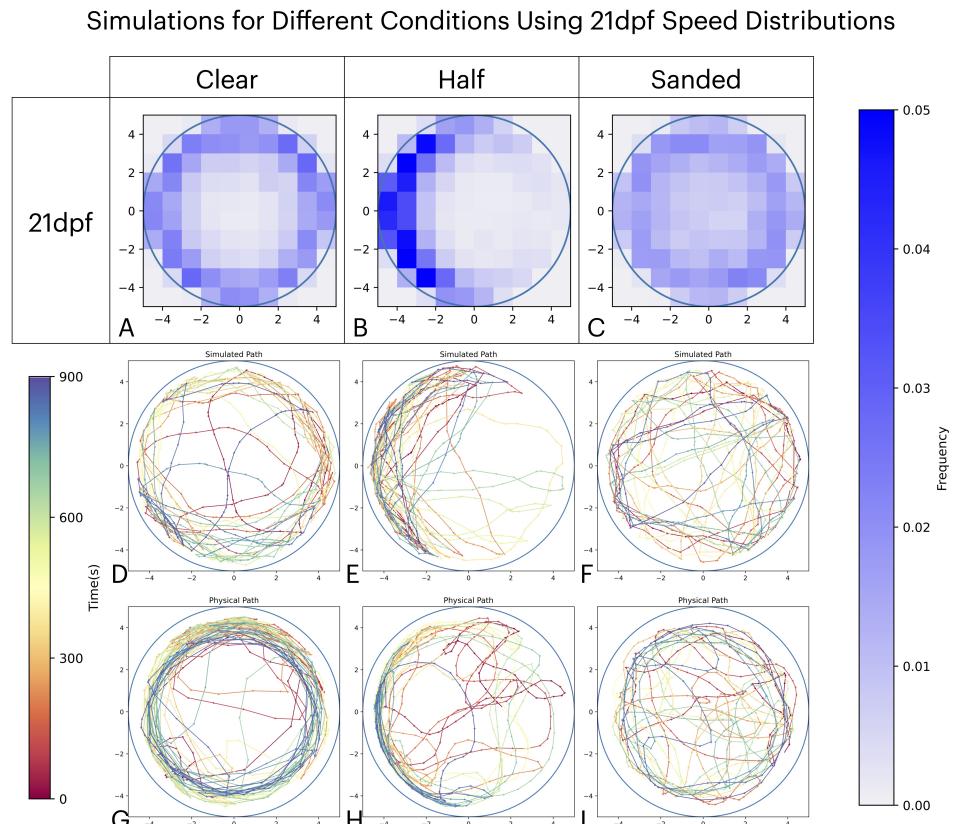


Figure 2.7: After simulating 10 iterations of 900 time points, we can produce the equivalent two-dimensional heatmaps for each simulation condition (A, B, C). We also include an arbitrary simulation of each condition, including the simulated zebrafish's path (D, E, F). Directly below, the bottom row includes a physical zebrafish swimming path of a single fish (G, H, I). Overall, our simulated distributions accurately reflect the overall behavior exhibited by the physical data, including both the symmetry-breaking and preference for a clear boundary. Again, the individual paths reflect the physical paths, but they are less smoothed out.

KS-test Comparisons Between Experiment and Simulation

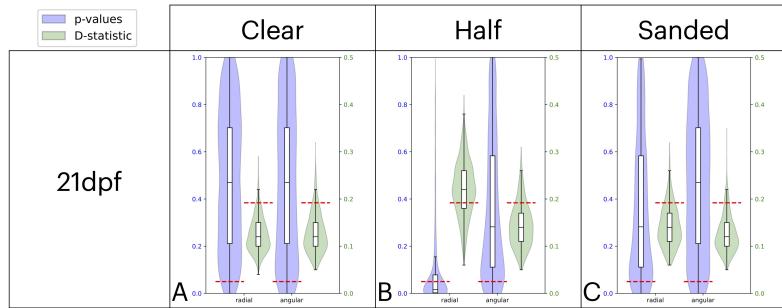


Figure 2.8: We run the experimental and simulation radial and angular distribution through the KS-test comparison for our data at 21dpf. For each condition, it outputs the distribution of p-values and the D-statistics. For statistically significantly different distributions, the median p-value should be less than the critical value for p-values (dashed line), and the median D-statistic should be above the critical value for D-statistics (dashed line). Both distributions are effectively the same as the experimental data for the clear dish (A) and sanded dish (C). The radial distributions are statistically significantly different for the half-sanded dish (B), but the main symmetry-breaking reflected in the angular distribution is effectively statistically similar.

conditions at 21dpf using a tunable turning parameter and a noise ratio factor.

# Chapter 3

## Discussion

### 3.1 Vision Drives Larval Locomotion

Our results indicate that larval zebrafish develop a progressively stronger reliance on vision, which becomes evident by 14dpf and becomes prevalent by 21dpf. Specifically, in the half-sanded experimental condition, our key finding is a distinct behavioral pattern characterized by an initial approach toward the boundary, followed by pronounced symmetry-breaking behavior, favoring movement toward the clear side of the dish. Our experiment setup intentionally sets up hydrodynamic similarity across all three experimental conditions: fully clear, fully sanded, and half-sanded. It becomes evident that hydrodynamic factors alone cannot account for the observed directional preferences, as we isolate the impact of the visual system development.

Thus, our findings suggest that visual sensory cues are crucial in guiding zebrafish locomotor decisions at these developmental stages. The preference for visually unobstructed spaces highlights the significant impact of visual stimuli on behavioral outcomes. Additionally, our data indicate that visual reliance is not merely present but strengthens substantially within the observed developmental window, implying an adaptive integration of sensory information critical for efficient navigation and environmental interaction in zebrafish larvae.

Previous studies have shown that movement is predominantly dictated by hydrodynamic and visual sensing [12, 13, 14, 15]. Thus, we tested the reliance on visual sensing in movement through two paths, which, in unison, provide us with a deeper understanding: testing blind fish and simulating our experimental fish using vision-based hypotheses.

Our results indicate that tyr-knockout mutant zebrafish with impaired vision do not exhibit the same symmetry-breaking development, suggesting that the visual system drives the boundary exploration behavior. By simulating zebrafish that operate as active Brownian particles, we can model these tyr-knockout mutant zebrafish, revealing the minimal decision-making involved in the movement system driven by hydrodynamic sensing.

The zebrafish visual system processes reflected light from the edges of the Petri dish, viewing a lack of a physical boundary and instead perceiving a boundary that is further than reality, leading the zebrafish to swim towards greater distances. This behavior may be because our zebrafish are never under the threat of predatory behavior; they can remain in a more exploratory state, where they are willing to explore areas where they perceive there is more space to explore, while maintaining relatively simple movement patterns when first hatched. This exploratory behavior could be a driving factor of the visual system having a positive affinity towards reflected light, rather than being more attracted to perceived physical boundaries. Thus, under this assumption, the zebrafish would effectively move as a weighted random walk, shifting towards this exploratory behavior closer to the boundary. When the zebrafish begin to see reflected light on the physical boundaries, it can drive the symmetry-breaking we see in the half-sanded dish.

We conclude that the development of the visual system influences the zebrafish's preference for the clear dish, where reflections can be perceived when the fish is near the boundary. These perceived reflections likely lead zebrafish to move close to the clear boundary, following the perceived reflected light. Especially closer to 21dpf, the circling of the boundary exhibited in the clear dish indicates the preference towards maintaining movement along the boundary (Figure 2.7). This contrasts with the movement we would

expect from a more random walk-based behavior, characterized by random movements on and off the boundary. Compared to our simulations with the reflection-based hypothesis, the smoother physical movement patterns display a more apparent circular motion around the borders. Even as we overlay physical data where the zebrafish can see reflection, it continues to move towards the next part of the boundary where it can see reflection. By applying the same overlay to the half-sanded dish, we can observe the zebrafish turning around near the boundary crossover point between the sanded and unsanded areas. Due to the zebrafish's larger field of vision, it can also sense reflected light coming from behind it. In the half-sanded case, we utilize our hypothesis within the simulation that, once moving forward towards reflected light is no longer possible, the zebrafish can still turn towards reflected light behind it, thereby continuing the exploratory process of moving towards the perceived boundaries that are farther away.

## 3.2 Future Work

Future work to confirm our conclusion involves understanding the neurodynamic implications of the visual system on movement patterns, as well as studying precisely what the zebrafish see and how their brains process this information [22]. We can further expand our experimental setup to implement new boundary conditions by placing the dish in a larger container filled with water, covering the edges with foil, or adding lenses and mirrors. This would allow us to hone in on what the zebrafish is seeing and how it might be driving movement behaviors. Through these studies, we can continue to utilize non-invasive methods to understand the function of the visual system in zebrafish [23]. We would leverage existing work to understand the neural implications of the visual system in driving zebrafish behavior, which would be critical in concluding the symmetry-breaking experienced through our studies [24, 25]. By studying zebrafish under new conditions, drawing on existing methods to investigate neural activity in zebrafish, we can place zebrafish under various visual boundary conditions and lighting conditions to understand

the decision-making process during the development of larval zebrafish [26]. On the simulation front, as more data is collected, we can employ machine learning approaches to create more robust models while still avoiding overfitting, thereby accurately modeling the overarching movement patterns [14].

### 3.3 Conclusion

Ultimately, we successfully developed an integrated experimental and computational pipeline for analyzing zebrafish behavior, providing a robust foundation that future work can expand on to investigate increasingly complex systems. On the experimental side, we have established a versatile imaging setup capable of recording zebrafish swimming behavior under various environmental conditions. Its simplicity and modularity allow rapid adjustments, such as changing camera lenses or swapping tanks to explore different experimental scenarios. As an initial validation, we investigated how water depth influences group behavior by substituting a Petri dish with a similarly sized beaker, requiring minimal modification to our existing apparatus. Complementing this, on the computational end, we have incorporated a Vicsek-based flocking model within our simulations, enabling us to scale and precisely simulate collective behavior and boundary interactions. We achieve direct quantitative comparisons by extracting comparable positional and velocity data from experimental observations and simulations, firmly bridging empirical results with theoretical models. Consequently, our comprehensive pipeline streamlines current analyses and provides a flexible, extensible platform to address future questions in zebrafish behavioral research, ultimately facilitating a deeper understanding of complex biological phenomena.

# Appendix A

## Methods

### A.1 Zebrafish Handling and Maintenance

We performed all zebrafish studies in accordance with IACUC-approved protocols at the Animal Resource Center (ARC) within the University of California, Santa Barbara. All of our experiments used wild-type AB zebrafish bred at the facility. Zebrafish were reared within the optimal window for maximum embryo production, between 3 and 18 months, euthanized, and replaced with younger batches when they reached 18 months of age. There is a decline in embryo production [27].

The embryos were raised in incubators at 28°C until 5 days after fertilization (dpf), from which we transferred the zebrafish to drip flow tanks placed on the rack. During incubation, the embryos were monitored until hatching, and any unfertilized embryos were removed to prevent contamination of the egg water. Larval zebrafish were housed in tanks of 8 fish to provide the fish with the potential for developing social cues without overcrowding.

We performed experiments at 7, 14, and 21dpf. Upon completion of the experiments at 21dpf, the zebrafish were euthanized via immersion in an MS-222 bath and remained in the immersion solution for at least 10 minutes until cessation of opercular (gill) movement was confirmed. The fish were then disposed of in a freezer as a secondary confirmation

of death. All zebrafish, larval and adult, are periodically monitored to ensure the overall health of the zebrafish facility.

## A.2 Injections

To utilize blind fish in experiments, some embryos were injected with a CRISPR/Cas9 tyr knockout via a microinjection pump system at the one-cell stage [11]. The Cas9 solution was mixed 1:1 with 1  $\mu$ L sgRNA for the tyr gene for each of the four guides. 0.67  $\mu$ L of each guide was then combined to generate the final guides. A 0.05% phenol red dye was added to visualize successful injections. The Cas9 protein sgRNA mixture was incubated at 37°C for 5 minutes. We performed microinjection by injecting 0.5-1 nL of the mixture into the cytoplasm of 1-cell stage embryos. Screening for successful injection was performed at 1dpf and 5dpf. At each data collection time, we selected only fish with clear depigmentation for experimentation. In addition, there is the potential for an impaired spinal cord if zebrafish receive an injection of an excess amount, which leads to minimal movements and earlier deaths, so for our experimentation, we prioritized fish that exhibited consistent movement.

## A.3 Experimentation

We used custom-designed experimental arenas with different boundary conditions, utilizing a Petri dish with a diameter of 10 cm and a height of 2 cm. We chose this size as an apt size for studying zebrafish larval development, as it is approximately 20 times the body length at this point, providing enough room to move without being too expansive. Dishes were either left as is, or half or fully sanded down along the border. Experiments were conducted using an overhead Teledyne FLIR Grasshopper3 USB3 camera with a Zeiss Milvus 85mm f/1.4 lens and illuminated from below using a modified version of the Louis Lab's PiVR setup, which emits uniform bright red light through ceramic plates

[21] (Figure 1.2). In our experimentation, we used our software for collection rather than the camera and software embedded within the PiVR software. Studies have shown that uniform red light does not impact the movement of larval zebrafish, despite potential long-term impacts on development [28]. However, this does not affect our experimental setup, as we are focused on short-term movement patterns, and euthanize the zebrafish upon completion of the experiment. Videos were captured at one frame per second (fps) for 15 minutes using a camera connected to a computer running the PointGrey FlyCapture SDK software. Experiments consisted of acquiring repeated data sets in a wide range of conditions, from varying ages (7, 14, and 21dpf), tanks, water depth (150mL, 250mL, 350mL), number of fish (1, 5, 10, 15, 25, 50, 100), and mutations (tyr and wild-type control).

## A.4 Tracking

For tracking, we uploaded our AVI videos to idtracker.ai to separate fish trajectories from the background. idtracker.ai is a machine learning software designed to track individual animals within small or large groups (up to 100 individuals) in video recordings, even when animals are unmarked and frequently cross paths or touch each other. The model uses two convolutional neural networks (CNNs): one for detecting crossings or touching events, and another for reliably identifying each animal across video frames.

The process begins with a preprocessing stage, where the software segments animals from the background into discrete image regions or 'blobs.' A dedicated CNN, termed the "crossing detector," classifies these blobs as either individual animals or crossings (instances of animals touching or overlapping each other).

Following this step, the identification network assigns identities to individual animals. Training this network requires constructing a dataset from video segments ("global fragments") where animals are distinctly separated. Sometimes, the software rejects videos if the network cannot train and track the animals reliably across frames [29].

The software would automatically track videos, including errors or uncertainties, from which we manually validated each video to adjust for errors, gaps, jumps, or duplicates. We then analyzed the processed and tracked videos via the trajectorytools package in Python to open the tracked .npy files to produce a working numpy array of positions and velocities. We then used these values, combined with the measured size of the dish, to normalize all values so that they could be accurately output in centimeters for positions and centimeters per second for velocities. From this point, we utilize the working positions and velocities for statistical analysis.

## A.5 Statistical Analysis

We converted the array of positions and velocities for statistical analysis into a usable data frame of all the potentially useful information. For each condition, we experimented by taking multiple different videos, so we merged all of these positions and velocities into one array per condition. We converted the positions and velocities into polar coordinates and extracted the corresponding vectors and angles, as shown in Figure A.2. We then conducted an exploratory data analysis to compare these variables across different conditions. One key plot we produced was the 2D heatmaps for each experimental condition. These plots provide us with vital information to drive the direction of our research. Upon finalizing the data used for our work, we compared the distributions of different conditions using the Kolmogorov-Smirnov test.

To prevent overfitting due to the large sample size, we applied Kernel Density Estimation (KDE) to smooth out the distributions, thereby limiting the impact of spikes in the experimental data. Gaussian KDE, we place a smooth Gaussian "bump" of width  $h$  at each data point and sum them to estimate the underlying density. Scott's rule gives a data-driven choice of this bandwidth.

$$h = \sigma n^{-1/(d+4)} \tag{A.1}$$

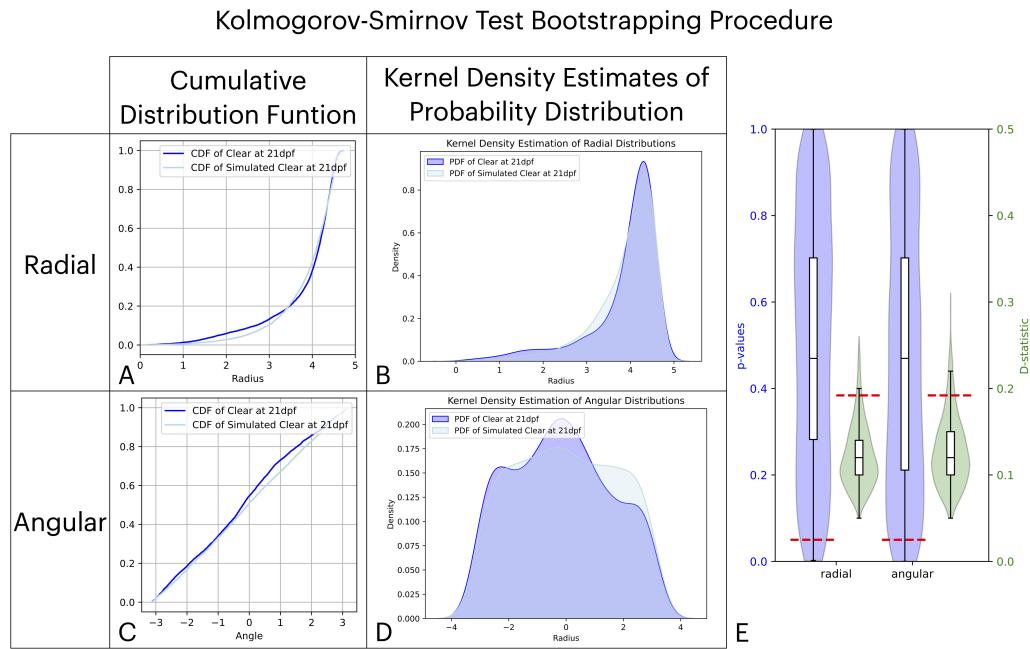


Figure A.1: Sample procedure for the Kolmogorov-Smirnov test, which we utilize to compare two arbitrary radial and angular distributions. We compare the 21dpf clear dish simulation and experimental data in this example. The first column represents the complete cumulative distribution function of the radial and angular values (A, C). Next, we apply kernel density estimation to smooth out the distributions (B, D), removing any sharp peaks caused by the randomness of the data. These two smoothed distributions are then bootstrapped to smaller sample sizes, and the Kolmogorov-Smirnov test is run. The final column plots the distribution of radial and angular p-values and D-statistic values using a violin plot and box plot to determine whether the distributions are statistically significant (E). In this case, radial and angular distributions are not statistically significantly different.

where  $\sigma$  is the sample standard deviation,  $n$  the number of observations, and  $d$  the data dimensionality. By automatically scaling with  $n$  and  $d$ , Scott's estimate balances bias and variance, minimizing the mean integrated squared error without requiring costly cross-validation. We then bootstrapped down to the optimal sample size for the KS-test, taking samples of 100 and comparing the distributions. The results of the iterations provided a distribution of p-values for the comparison between the sample distributions, based on 1000 bootstrapped comparisons. We used the results to make a statistical comparison between the different conditions. We can also consider the D-statistic, a measure of the maximum gap between the distributions, which provides an additional metric to assess the statistical significance between two sample distributions. We can make decisions on statistical significance using the p-value of 0.05. Thus, we calculate the D-statistic critical value using the Kolmogorov-Smirnov cumulative density function:

$$F_K(x) = 1 - 2 \sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2x^2} \quad (\text{A.2})$$

Thus, we can find the critical D-statistic by taking

$$\lambda_{0.05} = F_K^{-1}(0.95) \approx 1.3581 \quad (\text{A.3})$$

and finding the critical value to be

$$D_{\text{crit}}(0.05) = \lambda_{0.05} \sqrt{\frac{n_1 + n_2}{n_1 n_2}} = \lambda_{0.05} \sqrt{\frac{2}{n}} \quad (\text{A.4})$$

which yields a D-statistic critical value of 0.192. The KS-test provides us with the most informed comparison between the two different experimental conditions, as well as a comparison between experimental and simulated conditions. We utilize both the p-value and D-statistic critical values to determine whether distributions are statistically different. It's important to note that for statistically different distributions, the p-values are distributed below the p-value of 0.05, and the D-statistics are distributed above the

D-statistic critical point of 0.192.

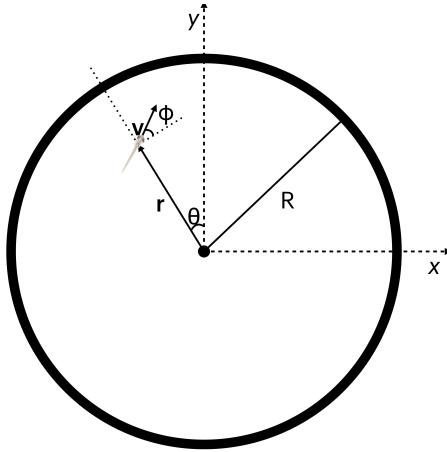


Figure A.2: For both our experimental and simulation data, we define the position as the radial vector from the center of the dish. We set our x and y axes from the default Cartesian coordinates from the top-down point of view. The polar angle  $\theta$  is the angle between the position vector and the y-axis. Meanwhile, the angle  $\phi$  is the angle between the velocity vector and the tangent vector. We utilize the  $\phi$  angle to represent the alignment of the fish with the wall, with small angles representing when the fish swims parallel to the wall.

In addition, we can take any simulation or experimental video and take the position and velocity values to determine at what point that specific fish is seeing a reflection. From that, we can make a visual to test our hypothesis by overlaying the actual video file and the points on the boundary where the zebrafish could see reflected light. From that, we can visualize whether each zebrafish exhibits the following behavior towards the points of reflection or if it randomly moves away from these points of reflected light.

## A.6 Simulation

We utilize our experimental data for all simulations to extract speed distributions at any time point during zebrafish development. We started with the most baseline model for any active particle that we will compare to both our experimental data and to our complex simulations by applying a 2-dimensional Brownian random walk with bounces

off the wall in addition to a Brownian random walk with a particle being stuck at the wall until it randomly chooses a direction it is capable of moving in without hitting a wall. For this, we can run our simulations in parallel with our physical data, running 10 iterations of 900 time points (equivalent to 15 minutes at one frame per second from our experimental data). From this, we can extract the positions and velocities from the simulation through an .npy file, just as they are output through idtracker.ai, which allows us to run the same statistical analysis tests on our simulated data and compare them directly with the physical data.

For our simulations of zebrafish movement, we started with the most elementary Vicsek bird flocking model, which aligns with the nearest neighbors within a specified radius [5]. We updated velocities according to the following equation:

$$v_a = \frac{1}{Nv} \sum_{i=1}^N | v_i | \quad (\text{A.5})$$

A key difference in our simulations from most Vicsek-based models was the application of constraints specific to a circular tank. Next, we introduced a Markov model approach to accurately represent zebrafish movement in pulses, between "moving" and "resting" states. For this, we referenced our preliminary data to inform the frequency of these movement bouts and establish some minimal randomized movement in the "rest" states. When we resorted to capturing single zebrafish data at one frame per second, we simplified our model to skip neighbor interactions (since there are none) and permanently remain in the "moving" Markov state. However, our simulations have the adaptability to include multiple fish traveling through Markov states, which can be explored in future work at higher frame rates.

Next, we implemented boundary interactions that reflected our hypothesis for how the zebrafish interacted with the boundary in our different conditions. We used the starting point of our turning mechanism as a bounce off the wall, treating the fish as a ball.

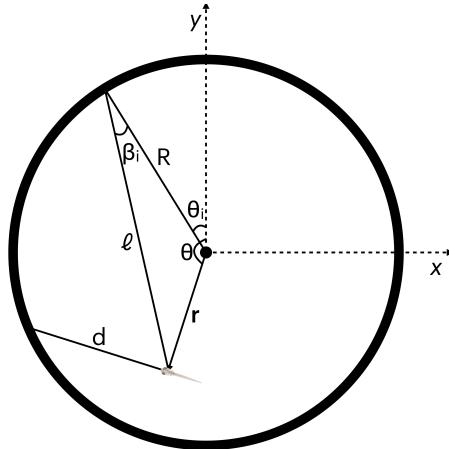


Figure A.3: For the zebrafish at an arbitrary time point, we define  $d$  as the distance to the boundary traveling directly forward. For reflection calculation, we define an arbitrary boundary point by the polar angle  $\theta_i$ , where the angle of incidence from that point of view is  $\beta_i$ . We utilize that angle to calculate where the zebrafish can see reflected light.

From that, we implemented a gradual turning mechanism representative of the sanded tank, turning towards the direction of least resistance (based on the same theory of bouncing off the wall), but making the turn in randomly selected angle increments. Given our ability to sample from the speed distribution, we could also adapt this to sample from the turn angle distribution. We now use a Gaussian distribution about a 0.75 radian turn with a standard deviation of 0.5 radians. After determining the initial direction the zebrafish wants to move, the turning probability function can override this decision to make a turn. Through a customizable turning probability function,

$$p(d) = e^{-\gamma d/R} \quad (\text{A.6})$$

We set any arbitrary function as the probability of turning at any given distance from the border. We take this probability value and run a random selection of 0 and 1 using the weight from the turning probability function. In this case, we define it as our tunable turning parameter, which varies the frequency at which the zebrafish turns. We calculated the distance to the border as the distance to hit the wall if traveling in a straight line.

Based on this, we can change the simulation to a series of different turning functions. For us, we tuned this turning probability function to roughly match the distributions of the sanded dish condition.

To model the clear side of the tank, we utilize the hypothesis that the zebrafish follows the reflection, so we take any given orientation to calculate exactly where the zebrafish can see the reflection around the tank. Given the zebrafish's arbitrary position and velocity, we utilized Snell's law to calculate at what points on the boundary light was refracted or reflected. If

$$n_{\text{water}} \sin(\beta_i) \geq n_{\text{air}} \text{ for arbitrary } \theta_i \quad (\text{A.7})$$

We define the relationship between the angle of incidence  $\beta_i$  and the angle  $\theta_i$ , representing a set of points along the boundary for which the zebrafish could be looking. Using the law of sines, we can find that

$$\frac{\sin(\theta - \theta_i)}{l} = \frac{\sin(\beta_i)}{r} \quad (\text{A.8})$$

which allows us, in conjunction with factoring in the blind spot, to determine whether the zebrafish would potentially see reflected light on the boundaries. We then took the subset of points on the boundary where zebrafish can perceive reflected light with preference in the direction of least resistance. We applied a uniform probability distribution to allow the zebrafish to randomly select a direction with reflected light, towards which the zebrafish would turn. Through Snell's law calculations, the zebrafish can only see reflections near the boundaries; otherwise, the sanded and clear dish simulations are the same when away from the boundary. Now, with these two key boundary conditions, we could run simulations in all three physical conditions.

The half-sanded dish implements both by removing the possibility of seeing reflection on half the dish and completely ignoring the possibility of reflection when near the sanded half of the boundary.

For all simulations, we can either randomly set speed values or utilize the experimental speed data to sample speed values. Embedded within our simulation, we have an embedded noise ratio, which adds the following vector to the speed:

$$\text{noise} = (\text{noise ratio} \cdot \text{speed})(\cos \theta \hat{\mathbf{x}} + \sin \theta \hat{\mathbf{y}}) \quad (\text{A.9})$$

For a random angle  $\theta$ . This noise vector is applied both in movement and in turning. Thus, when the turning probability function value samples a 0, representing no turn, we calculate where the zebrafish moves based on position  $\mathbf{p}$  and velocity  $\mathbf{v}$

$$\begin{aligned} \mathbf{p}(t+1) &= \mathbf{p}(t) + \left( s_t \frac{\mathbf{v}(t)}{|\mathbf{v}(t)|} + \text{noise} \right) \\ \mathbf{v}(t+1) &= s_t \frac{\mathbf{v}(t)}{|\mathbf{v}(t)|} + \text{noise} \end{aligned} \quad (\text{A.10})$$

Similarly, when we sample a 1 out of the probability function, the noise is embedded into a turn such that the position and velocity follow the following functions for when it completes a turn from the given turn vector, driven by either reflection or the turn gaussian:

$$\begin{aligned} \mathbf{p}(t+1) &= \mathbf{p}(t) + \left( s_t \left( \frac{\mathbf{v}(t)}{|\mathbf{v}(t)|} + \text{turn} \right) + \text{noise} \right) \\ \mathbf{v}(t+1) &= s_t \left( \frac{\mathbf{v}(t)}{|\mathbf{v}(t)|} + \text{turn} \right) + \text{noise} \end{aligned} \quad (\text{A.11})$$

This set of equations provides the complete basis for understanding the movement of our single zebrafish simulations.

Initially, we set up our simulation so that, similar to the Brownian motion model, the zebrafish would remain still if the desired movement location is outside the boundary. However, we decided to change our model so that the zebrafish will iterate in place until a plausible movement angle and speed are determined, based on the notion that the physical zebrafish has some mechanism embedded that prevents it from colliding with

the walls.

With these simulations, we can extract all the position and velocity values in the same manner as we would for physical data, which allows us to directly make comparisons as referenced in Appendix A.5: Statistical Analysis.

# Bibliography

- [1] E. A. Hebets, A. B. Barron, C. N. Balakrishnan, M. E. Hauber, P. H. Mason, and K. L. Hoke, *A systems approach to animal communication*, *Proceedings of the Royal Society B: Biological Sciences* **283** (Mar., 2016) 20152889.
- [2] D. A. Wijeyakulasuriya, E. W. Eisenhauer, B. A. Shaby, and E. M. Hanks, *Machine learning for modeling animal movement*, *PLOS ONE* **15** (July, 2020) e0235750.
- [3] J. Xia and X. Li, *Bifurcation analysis in a discrete predator-prey model with herd behaviour and group defense*, *Electronic Research Archive* **31** (2023), no. 8 4484–4506.
- [4] D. J. Butts, N. E. Thompson, S. A. Christensen, D. M. Williams, and M. S. Murillo, *Data-driven agent-based model building for animal movement through Exploratory Data Analysis*, *Ecological Modelling* **470** (Aug., 2022) 110001.
- [5] T. Vicsek, A. Czirok, E. Ben-Jacob, I. Cohen, and O. Sochet, *Novel type of phase transition in a system of self-driven particles*, *Physical Review Letters* **75** (Aug., 1995) 1226–1229. arXiv:cond-mat/0611743.
- [6] M. Bergmann and A. Iollo, *Modeling and simulation of fish-like swimming*, *Journal of Computational Physics* **230** (Jan., 2011) 329–348.
- [7] V. Ajraldi, M. Pittavino, and E. Venturino, *Modeling herd behavior in population systems*, *Nonlinear Analysis: Real World Applications* **12** (Aug., 2011) 2319–2338.
- [8] W. Li, D. Wolinski, J. Pettré, and M. C. Lin, *Biologically-Inspired Visual Simulation of Insect Swarms*, *Computer Graphics Forum* **34** (May, 2015) 425–434.
- [9] M. B. Orger and G. G. De Polavieja, *Zebrafish Behavior: Opportunities and Challenges*, *Annual Review of Neuroscience* **40** (July, 2017) 125–147.
- [10] S. M. Bugel, R. L. Tanguay, and A. Planchart, *Zebrafish: A Marvel of High-Throughput Biology for 21st Century Toxicology*, *Current Environmental Health Reports* **1** (Dec., 2014) 341–352.

- [11] R. S. Wu, I. I. Lam, H. Clay, D. N. Duong, R. C. Deo, and S. R. Coughlin, *A Rapid Method for Directed Gene Knockout for Screening in G0 Zebrafish*, *Developmental Cell* **46** (July, 2018) 112–125.e4.
- [12] V. C. Fleisch and S. C. Neuhauss, *Visual Behavior in Zebrafish*, *Zebrafish* **3** (June, 2006) 191–201.
- [13] J. H. Bollmann, *The Zebrafish Visual System: From Circuits to Behavior*, *Annual Review of Vision Science* **5** (Sept., 2019) 269–293.
- [14] H. Baier and E. K. Scott, *The Visual Systems of Zebrafish*, *Annual Review of Neuroscience* **47** (Aug., 2024) 255–276.
- [15] R. Thandiackal and G. V. Lauder, *How zebrafish turn: analysis of pressure force dynamics and mechanical work*, *Journal of Experimental Biology* (Jan., 2020) jeb.223230.
- [16] P. S. Suriyampola, D. J. Sykes, A. Khemka, D. S. Shelton, A. Bhat, and E. P. Martins, *Water flow impacts group behavior in zebrafish (Danio rerio)*, *Behavioral Ecology* **28** (2017), no. 1 94–100.
- [17] R. E. Johnson, S. Linderman, T. Panier, C. L. Wee, E. Song, K. J. Herrera, A. Miller, and F. Engert, *Probabilistic Models of Larval Zebrafish Behavior Reveal Structure on Many Scales*, *Current Biology* **30** (Jan., 2020) 70–82.e4.
- [18] M. Tuqan and M. Porfiri, *Mathematical Modeling of Zebrafish Social Behavior in Response to Acute Caffeine Administration*, *Frontiers in Applied Mathematics and Statistics* **7** (Oct., 2021) 751351.
- [19] P. De Lellis, E. Cadolini, A. Croce, Y. Yang, M. D. Bernardo, and M. Porfiri, *Model-Based Feedback Control of Live Zebrafish Behavior via Interaction With a Robotic Replica*, *IEEE Transactions on Robotics* **36** (Feb., 2020) 28–41.
- [20] G. Le Goc, J. Lafaye, S. Karpenko, V. Bormuth, R. Candelier, and G. Debrégeas, *Thermal modulation of Zebrafish exploratory statistics reveals constraints on individual behavioral variability*, *BMC Biology* **19** (Sept., 2021) 208.
- [21] D. Tadres and M. Louis, *PiVR: An affordable and versatile closed-loop platform to study unrestrained sensorimotor behavior*, *PLOS Biology* **18** (July, 2020) e3000712.
- [22] F. Svara, D. Förster, F. Kubo, M. Januszewski, M. Dal Maschio, P. J. Schubert, J. Kornfeld, A. A. Wanner, E. Laurell, W. Denk, and H. Baier, *Automated synapse-level reconstruction of neural circuits in the larval zebrafish brain*, *Nature Methods* **19** (Nov., 2022) 1357–1366.

- [23] R. Wu, O. Deussen, I. D. Couzin, and L. Li, *Non-invasive eye tracking and retinal view reconstruction in free swimming schooling fish*, *Communications Biology* **7** (Dec., 2024) 1636.
- [24] R. Portugues and F. Engert, *The neural basis of visual behaviors in the larval zebrafish*, *Current Opinion in Neurobiology* **19** (Dec., 2009) 644–647.
- [25] J. Xie, P. R. Jusuf, B. V. Bui, and P. T. Goodbourn, *Experience-dependent development of visual sensitivity in larval zebrafish*, *Scientific Reports* **9** (Dec., 2019) 18931.
- [26] D. A. Guggiana-Nilo and F. Engert, *Properties of the Visible Light Phototaxis and UV Avoidance Behaviors in the Larval Zebrafish*, *Frontiers in Behavioral Neuroscience* **10** (Aug., 2016).
- [27] A. Avdesh, M. Chen, M. T. Martin-Iverson, A. Mondal, D. Ong, S. Rainey-Smith, K. Taddei, M. Lardelli, D. M. Groth, G. Verdile, and R. N. Martins, *Regular Care and Maintenance of a Zebrafish (*Danio rerio*) Laboratory: An Introduction*, *Journal of Visualized Experiments* (Nov., 2012) 4196.
- [28] N. Villamizar, L. M. Vera, N. S. Foulkes, and F. J. Sánchez-Vázquez, *Effect of Lighting Conditions on Zebrafish Growth and Development*, *Zebrafish* **11** (Apr., 2014) 173–181.
- [29] F. Romero-Ferrero, M. G. Bergomi, R. C. Hinz, F. J. H. Heras, and G. G. De Polavieja, *idtracker.ai: tracking all individuals in small or large collectives of unmarked animals*, *Nature Methods* **16** (Feb., 2019) 179–182.