Inferring Causal Relationships in Zebrafish-Robot Interactions through Transfer Entropy: A Small Lure to Catch a Big Fish

Maurizio Porfiri

Department of Mechanical and Aerospace Engineering, Tandon School of Engineering, New York University

Corresponding author (Email: mporfiri@nyu.edu)

Abstract – In the field of animal behavior, effective methods to apprehend causal relationships that underlie the interactions between animals are in dire need. How to identify a leader in a group of social animals or quantify the mutual response of predator and prey are exemplary questions that would benefit from an improved understanding of causality. Information theory offers a potent framework to objectively infer cause-and-effect relationships from raw experimental data, in the form of behavioral observations or individual trajectory tracks. In this targeted review, we summarize recent advances in the application of the information-theoretic concept of transfer entropy to animal interactions. First, we offer an introduction to the theory of transfer entropy, keeping a balance between fundamentals and practical implementation. Then, we focus on animal-robot experiments as a means for the validation of the use of transfer entropy to measure causal relationships. We explore a test battery of robotics-based protocols designed for studying zebrafish social behavior and fear response. Grounded in experimental evidence, we demonstrate the potential of transfer entropy to assist in the detection and quantification of causal relationships in animal interactions. The proposed robotics-based platforms offer versatile, controllable, and customizable stimuli to generate *a priori* known cause-and-effect relationships, which would not be feasible with live stimuli. We conclude the paper with an outlook on possible applications of transfer entropy to study group behavior and clarify the determinants of leadership in social animals.

Keywords – Ethorobotics, Fear response, Information theory, Robotics, Social behavior

The question of causality is pervasive to the study of animal behavior, yet, seldom do our quantitative methods enable precise measurements of the causal relationships that underlie animal interactions. Imagine a cheetah hunting a gazelle and the gazelle running for its life. As the cheetah spots a gazelle drifting from its herd, it will attack, running toward the selected prey. The gazelle will soon realize the imminent threat and will, in turn, start running. It will weave and zig-zag to lose its predator, but the cheetah will continue chasing, attempting to stay in pursuit and hold traction to the ground. Ultimately, the gazelle might be able to survive or the faster cheetah might succeed. The outcome of this predator-prey interaction, sketched in Figure 1, is less important than its revealing dynamics. Throughout the duration of this interaction, the cause and the effect alternate: sometimes it is the predator to influence the behavior of the prey, sometimes it is the prey that determines what the predator will do, and some other times the two will simultaneously adjust their behavior in response to what the other is doing. How can we tell what is truly happening through objective measures of their locomotion? In other words, if we were in possession

of time-resolved data of the behavior of both the cheetah and the gazelle, could we infer the cause-and-effect relationships underlying their interactions?
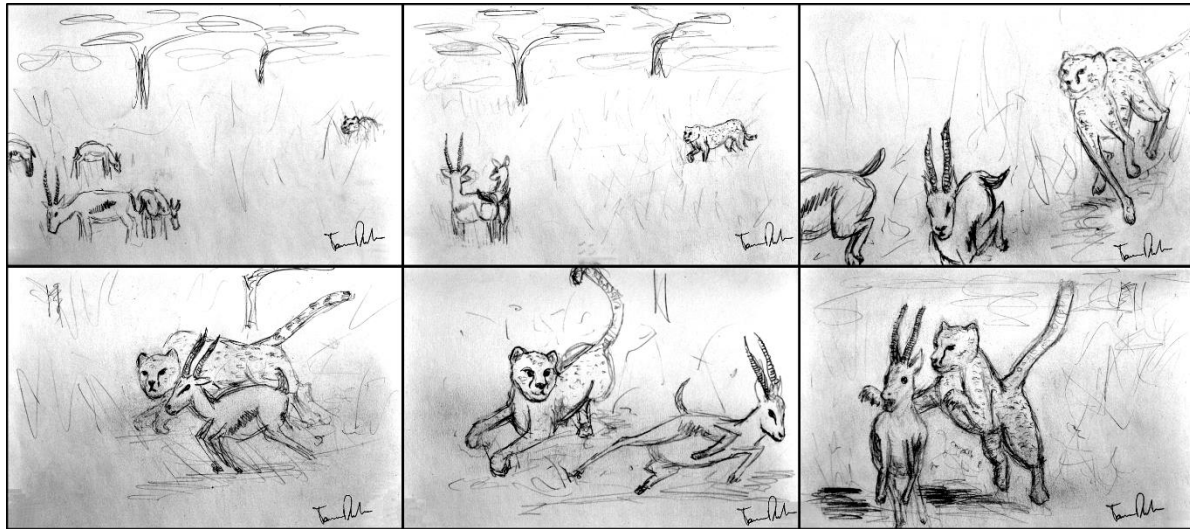

Figure 1. A cheetah hunting a gazelle and the gazelle running for its life. [Drawing by Tommaso Ruberto]

An answer to this question may be offered by the field of information theory, whose foundations date back to the seminal work of Shannon. In the words of Shannon (Shannon, 1948), information can be interpreted as "a measure of how much choice is involved in the selection of the event or of how uncertain we are of the outcome," such that more uncertainty means more information. A particularly empowering tool within this, now mature, field of investigation is the notion of transfer entropy, put forward nearly two decades ago by Schreiber (Schreiber, 2000). In his seminal paper (Schreiber, 2000), Schreiber introduced transfer entropy as a means to detect asymmetries in interaction of two coupled dynamical systems from their time series. Since then, we have seen applications of transfer entropy spawning in the most disparate fields of science and engineering, where the identification of cause-and-effect relationships is required. For instance, the concept of transfer entropy has been employed in neuroscience (Stetter, Battaglia, Soriano, & Geisel, 2012; Vicente, Wibral, Lindner, & Pipa, 2011), finance (Marschinski & Kantz, 2002; Sandoval, 2014), climatology (Hlinka et al., 2013; Runge, Heitzig, Petoukhov, & Kurths, 2012), sociology (Ver Steeg & Galstyan, 2012), and physiology (Faes et al., 2013). An excellent introduction to transfer entropy, along with an up-to-date overview of its applications can be found in (Bossomaier, Barnett, Harré, & Lizier, 2016).

In an information-theoretic sense, transfer entropy measures the reduction in the uncertainty of the prediction of the future of a system from its present due to additional knowledge about the present of another system. Following Wiener's principle of causality (Wiener, 1956), such a reduction can be associated with a cause-and-effect relationship between the two systems. If the two systems are decoupled, neither contains valuable information about the future of the other one, therefore transfer entropy should vanish. On the other hand, if one of the systems influences the other, a nonzero transfer entropy should be detected between the two time series, signifying that one system contains useful information to predict the future of the other.

In summer 2013, I had the fortune of being invited to a workshop supported by the United States Army Research Office (Bollt & Sun, 2014), which offered me a first exposure to transfer entropy. This is when I started my journey into the application of this promising technique to study causal relationships underlying animal interactions. Over the last four years, I have had the privilege of working with a large number of excellent collaborators under the support of the National Science Foundation to demonstrate the possibility of inferring causal relationships in animal behavior through the use of transfer entropy. With respect to animals, information should be associated with the degree of uncertainty in the behavior of an

animal, and transfer entropy should be consequently regarded as a measure of the exchange of information between individuals.

Most of our work has focused on zebrafish (*Danio rerio*), which is emerging as a powerful model organism to study functional and dysfunctional processes in humans (Orger & de Polavieja, 2017; Stewart, Braubach, Spitsbergen, Gerlai, & Kalueff, 2014). Native to the southeastern Himalayan region, this freshwater species is becoming ubiquitous in academic and industrial laboratories for an incredibly vast array of advantages that are bestowed by working with zebrafish (Dooley & Zon, 2000; Lieschke & Currie, 2007; Miklósi & Andrew, 2006). To name a few of them, zebrafish have impressive genetic and neural similarities with humans (Panula et al., 2010; Postlethwait et al., 2000), stocking and maintenance of zebrafish are easy (Brand, Granato, & Nüsslein-Volhard, 2002), and zebrafish come with a fully sequenced genome that has fueled the development of several sophisticated genetic modification techniques (Howe et al., 2013). An excellent example of the potential of zebrafish to aid preclinical research is their recent surge in the study of deficits relevant to autism spectrum disorders (Meshalkina et al., 2017).

Our research has turned to robotics for establishing a robust methodological framework that could help in manipulating the direction and strength of causal relationships, toward highly-controllable experimental datasets (Butail, Ladu, Spinello, & Porfiri, 2014). Within this framework, we designed and implemented a class of biologically-inspired robotic stimuli that could influence the behavior of live zebrafish without responding to them, or mirroring the behavior of live animals. In a sequence of ethorobotics experiments, we have leveraged this approach to: (i) investigate the role of body size (Bartolini et al., 2016) and mobility (Ruberto, Mwaffo, Singh, Neri, & Porfiri, 2016) on animal interactions; (ii) explore the effect of psychostimulant substances on zebrafish social behavior (Ladu, Mwaffo, Li, Macri, & Porfiri, 2015); and (iii) study fear response to a predatorial stimulus (Neri, Ruberto, Cord-Cruz, & Porfiri, 2017).

As a complementary strategy for the generation of surrogate datasets on which to test the potential of transfer entropy, we have also explored data-driven modeling (Butail, Mwaffo, & Porfiri, 2016; Mwaffo, Butail, & Porfiri, 2017a). Specifically, we have developed a class of mathematical models for zebrafish, in which each individual is described as a self-propelled particle moving in a two or three-dimensional space and interacting with other individuals and the boundary of the environment (Mwaffo, Anderson, Butail, & Porfiri, 2015; Mwaffo, Butail, & Porfiri, 2017b; A. Zienkiewicz, Barton, Porfiri, & di Bernardo, 2015a, b; A. K. Zienkiewicz, Ladu, Barton, Porfiri, & Di Bernardo, 2018). By differentially changing model parameters throughout the group, one can generate synthetic datasets where the strength and direction of the causal relationships are fully controlled. For example, by setting to zero the interaction parameters of a particular individual in the group, one would enforce the individual to be unresponsive to others, which would, however, still respond to this individual. Ultimately, this would create a causal relationship within the group.

Due to the focus of this special issue on animal-robot interaction, I will favor ethorobotics as a means to explain and validate the theoretical construct of transfer entropy. However, when possible, I will refer to the alternative approach of data-driven mathematical modeling, which I believe holds an equally important value in demonstrating the importance of transfer entropy to infer causal relationships in animal behavior. The rest of this paper is organized into three parts, dealing with theoretical aspects related to the use of transfer entropy, experiments on zebrafish-robot interactions, and open issues that warrant further investigation.

In the first part, I introduce the notion of transfer entropy in its original incarnation by Schreiber (Schreiber, 2000). I describe the main steps for building a computer code to calculate transfer entropy from raw time series. I point the reader to a few available toolboxes that can be used for rapid computation. Then, I highlight possible extensions to account for time delays and temporal correlations, which can be undertaken to improve the predictive power of transfer entropy. In the same vein, I also present a pre-processing technique for encoding salient aspects of animal behavior into meaningful symbols. Toward ultimately employing transfer entropy in the study of empirical datasets, I summarize potential statistical methods for hypothesis testing in the detection of causal relationships. I conclude the theoretical part by introducing a post-processing approach to combine transfer entropy with other estimators of causal

relationships. In the second part, I present our robotics-based experimental approach to manipulate interactions between live zebrafish and biologically-inspired robotic stimuli. I touch on both attractive stimuli inspired by zebrafish morphophysiology and fear-evoking stimuli based on zebrafish predators. I summarize experimental results from our group, demonstrating the feasibility of detecting causal relationships underlying animal interactions through transfer entropy. In the third and final part, I examine feasible extensions of the approach to the study of animal groups and identify a few gaps in knowledge that warrant future theoretical and experimental research.

Beyond previous work from my group on zebrafish, throughout this manuscript I attempt at summarizing research findings on other animal systems, which support the application of transfer entropy and recent theoretical developments to the study of other animal systems. While the number of these efforts is not very large, the breadth of their focus, from soldier crabs (Tomaru et al., 2016) to bats (Orange & Abaid, 2015), insects (Lord, Sun, Ouellette, & Bollt, 2016), and again fish (Hu, Nie, & Fu, 2015), makes a very compelling case for the promise of transfer entropy in animal behavior. The conclusions of these studies open the door for intriguing conclusions on the underpinnings of social behavior. For example, the work of (Tomaru et al., 2016) demonstrates that the influence of social interactions in soldier crabs increases with the size of the group and the work of (Hu et al., 2015) hints at the existence of a zone where prey fish are highly vigilant about the position of a predator.

Overall, I would like to catch a "big fish," in the sense that I aim to draw attention to a fundamental theoretical concept that has dramatic repercussions on our capacity to quantitatively study interactions in animals. Yet, I use a "small lure," such that the demonstration of the approach to the study of zebrafish-robot interactions is deliberately focused, toward a precise quantification of the advantages and limitations of the approach on a real case study.

Theory

A Primer on Transfer Entropy

In information theory (Cover & Thomas, 2012), the uncertainty of a discrete random variable X is quantified through the notion of entropy $H(X)$, defined as

$$H(X) = -\sum_{x \in \Omega} \Pr(X = x) \log \Pr(X = x), \tag{1}$$

where $\Pr(\cdot)$ is probability, Ω is the sample space that includes all possible realizations of X, and x is a generic realization. Entropy can be measured in "bits" or "nats," depending on the base of the logarithm, and is always nonnegative. As an example, consider the toss of a coin, such that the two possible realizations of the toss are "heads" and "tails." Let p be the probability that the coin toss will give a heads, then the entropy of the coin toss is simply $-(p \log p + (1-p) \log(1-p))$. This quantity varies from 0 to 1 bit, where 0 is attained for a completely unfair coin that will always return either heads or tails[1] and 1 bit corresponds to a fair coin. In an information-theoretic sense, no information is encoded in tossing a coin that has two heads or two tails, because we completely know the outcome of the process. On the other hand, the highest amount of information is encoded in a fair coin, where the odds of heads and tails are the same, and we cannot predict the outcome of the coin toss.

To the best of my knowledge, the first formal application of the concept of entropy to animal behavior was the work of Paulus, Geyer, Gold, and Mandell (1990), in an effort to assess change in rat locomotion induced by amphetamine and 3,4-methylenedioxymethamphetamine. Through the lens of entropy, one may seek to compute the degree of variability in the behavior of an animal, quantified from videos through behavioral observations or trajectory tracking. Specifically, one may opt for constructing a probability space from the ethogram of the animal, such that the random variable would correspond to the specific behavior exhibited by the individual. Alternatively, one may use tracked data, such that the random variable would be associated with the position or orientation of the individual. In both incarnations, large

---

[1] Note that $p \log p$ goes to zero as p approaches zero.

values of the entropy will highlight instances in which the individual will tend to spend comparable time across all the behaviors of the ethogram, change its position without a preference for a location versus another one, or randomly change their orientation without a preferred heading.

From a mathematical point of view, the right hand side of Equation (1) can be viewed as the expectation of -log Pr(X). Based on this, we can define the joint and conditional entropies of two random variables X and Y

$$H(X,Y) = - \sum_{x,y \in \Omega} \Pr(X = x, Y = y) \log \Pr(X = x, Y = y), \qquad (2a)$$

$$H(X|Y) = - \sum_{x,y \in \Omega} \Pr(X = x, Y = y) \log \Pr(X = x|Y = y), \qquad (2b)$$

where y is a generic realization of Y.[2]

We are now ready to introduce the notion of transfer entropy between two stationary stochastic processes $X = \{X_t\}_{t \in N}$ and $Y = \{Y_t\}_{t \in N}$ as follows:

$$\text{TE}^{Y \to X} = H(X_{t+1}|X_t) - H(X_{t+1}|X_t, Y_t) . \qquad (3)$$

In this sense, transfer entropy is a measure of the reduction in the uncertainty of the future of X from its present that is gained through knowledge of the present of Y. The notion of stationarity is used to compute a single value of transfer entropy for any choice of the time t; should the processes be non-stationary, the right hand side of Equation (3) would vary in time, thereby leading to an instantaneous measure of information transfer, rather than an aggregated quantity.

By replacing Equation (2b) into Equation (3), we arrive at the following common expression for transfer entropy:

$$\text{TE}^{Y \to X} = \sum_{x_+,x,y \in \Omega} \Pr(X_{t+1} = x_+, X_t = x, Y_t = y) \log \frac{\Pr(X_{t+1}=x_+|X_t=x,Y_t=y)}{\Pr(X_{t+1}=x_+|X_t=x)}, \qquad (4)$$

where we used $\Pr(X_{t+1} = x_+, X_t = x) = \sum_{y \in \Omega} \Pr(X_{t+1} = x_+, X_t = x, Y_t = y)$ to combine the summands.

In the context of animal interactions, X and Y will measure the behavior of two individuals and transfer entropy will be constructed based on raw time series extracted from behavioral observations or trajectory tracking. Going back to the example of the gazelle and the cheetah in the introduction, we may register the behavior of each animal at any time step (for example, running straight, zig-zagging, or jumping) or the orientation of the two and use transfer entropy to explore cause-and-effect relationships underlying the predator-prey interaction. The stationarity of the processes implies that the interaction repeats itself over and over during the experiment, without time effects that may modulate its strength or direction.

Computation of Transfer Entropy from Raw Data

To compute transfer entropy from two raw time series for X and Y, it may be convenient to express all the conditional probabilities in terms of $\Pr(X_{t+1} = x_+, X_t = x, Y_t = y)$ by using the definition of conditional probability[3]. We write

$$\text{TE}^{Y \to X} = \sum_{x_+,x,y \in \Omega} \Pr(X_{t+1} = x_+, X_t = x, Y_t = y) \log \frac{\Pr(X_{t+1}=x_+,X_t=x,Y_t=y)\Pr(X_t=x)}{\Pr(X_t=x,Y_t=y)\Pr(X_{t+1}=x_+,X_t=x)}, \qquad (5)$$

where the marginal distributions can all be retrieved from the joint probability mass function for the triplet (X_{t+1}, X_t, Y_t) by summing over all the missing variables, as we did for Equation (4). Thus, the only function that shall be estimated is $\Pr(X_{t+1} = x_+, X_t = x, Y_t = y)$, which is defined over $|\Omega|^3$ triplets, where $|\Omega|$ is the cardinality of the sample space. Such an estimation can be carried out through a simple plug-in estimator (Zhang & Grabchak, 2013), by counting the number of instances in which the selected triplet (x_+, x, y) will appear in the two time series as (X_{t+1}, X_t, Y_t) . For two time series of duration T, this will correspond to binning a total of T-1 triplets as the time step is varied from 1 to T-1.

---

[2] We assume that the probability space of X and Y is Ω, although one may consider a more general case in which each random variable, or process, has its own, different, probability space.

[3] For two random variables X and Y, conditional probability is defined as $\Pr(X = x|Y = y) = \frac{\Pr(X=x,Y=y)}{\Pr(Y=y)}$.

As a toy example, consider two binary ($|\Omega| = 2$) times series of duration $T = 5$: $X = \{0,1,1,0,1\}$ and $Y = \{1,0,1,1,1\}$. We have a total of $2^3 = 8$ triplets to count, namely: (0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), and (1,1,1). For each of these eight triplets, the corresponding probability is estimated by counting its frequency and scaling by T - 1. For instance, the triplet (0,0,0) never appears as a triplet since we do not find two consecutive zeros in X. On the other hand, there are two occurrences of (1,0,1), where we find twice that X switches from 0 to 1 in between two consecutive steps, t and $t+1$, with Y being equal to 1 at time t. Following this process, we estimate all the probabilities, as follows: Pr(0,0,0) = 0, Pr(0,0,1) = 0, Pr(0,1,0) = 0, Pr(0,1,1) = 1/4, Pr(1,0,0) = 0, Pr(1,0,1) = 2/4, Pr(1,1,0)= 1/4, and Pr(1,1,1) = 0. By marginalizing this probability mass function, we compute the remaining probability mass functions appearing in Equation (5). For instance[4],

$$\Pr(X_t = 1, Y_t = 0) = \Pr(X_{t+1} = 0, X_t = 1, Y_t = 0) + \Pr(X_{t+1} = 1, X_t = 1, Y_t = 0), \quad (6)$$

which is equal to 1/4 in our toy example. Finally, we estimate transfer entropy from Y to X as

$$\text{TE}^{Y \rightarrow X} = \frac{1}{4}\log\frac{\left(\frac{1}{4}\right)\left(\frac{2}{4}\right)}{\left(\frac{1}{4}\right)\left(\frac{1}{4}\right)} + \frac{2}{4}\log\frac{\left(\frac{2}{4}\right)\left(\frac{2}{4}\right)}{\left(\frac{2}{4}\right)\left(\frac{2}{4}\right)} + \frac{1}{4}\log\frac{\left(\frac{1}{4}\right)\left(\frac{2}{4}\right)}{\left(\frac{1}{4}\right)\left(\frac{1}{4}\right)} = \frac{1}{2} \text{ bits.} \quad (7)$$

Once a computer code has been written to calculate transfer entropy from times series, it is advisable to verify it against analytical results. For example, it contains closed form results for transfer entropy in a binary model of interaction that exemplifies firing of neurons in the brain or adoption of policy in a legal environment. By contrasting the predictions of the computer code against closed form results, one may verify its integrity, while appreciating first-hand the need for long time series to accurately estimate transfer entropy. Beyond this simple binary model, one can find closed form results against which to verify the developed computer code in (Porfiri, 2018; Porfiri & Ruiz Marín, 2018; Smirnov, 2013).

Working with experimental data on animal behavior, one often has the freedom to choose the size of the sample space $|\Omega|$, which corresponds to the number of behaviors constituting the ethogram when dealing with behavioral observations or the number of bins used to lump continuous data from trajectory tracking, such as position and orientation. The number of bins should be commensurate with the length of the time series, whereby excessively refined sample space demands long time series for accurate estimation of the probability mass function and transfer entropy. The bias of the estimation converges to zero linearly with the length of the time series and scales with the cardinality of sample space (Zhang & Grabchak, 2013).

As a rule of thumb, we should require several occurrences for each possible triplet, which implies $T \gg |\Omega|^3$. However, because we deal with animals, many of the triplets are biologically unfeasible, thereby alleviating this constraint. For example, if we work with position, then the animal will not be able to travel in a single time step between two positions that are far away. This restricts the need for estimating the probability mass functions only for triplets (X_{t+1}, X_t, Y_t) in which the first two numbers represent physically adjacent bins, such that $T \gg |\Omega|^2$.

The brute-force solution of increasing the density of the dataset during the design of an experiment, by maximizing the frequency of data collection, can also be pursued. However, some prudence is warranted. As further discussed in what follows dealing with experimental results, excessively high acquisition frequencies may enhance the effect of measurement noise (Paluš & Vejmelka, 2007), washing out possible interactions between individuals. Important theoretical insight into the selection of the binning size and the sampling frequency can be garnered from (Baptista et al., 2012), which clarifies the role of relevant dynamic measures (Lyapunov exponents, expansion rates, and capacity dimension) on the information transfer between coupled oscillators.

For continuous distributions, kernel density estimation may be used as an alternative approach to simple binning, but this may also require long time series and relies on the choice of a proper kernel-type function and a smoothing parameter (Silverman, 1986). An inadequate selection of the kernel function and/or smoothing parameter may hinder the accuracy of the estimations (Moon, Rajagopalan, & Lall, 1995). Alternative to binning and kernel density estimation, one might consider estimators based on k-nearest

---

[4] The marginal distributions can also be obtained by directly counting in the time series, although small discrepancies could be noted for a toy example like this, in which counting up to T or T-1 makes a difference.

neighbor statistics which are less intuitive than binning but more data efficient, such as the KSG estimator (Kraskov, Stögbauer, & Grassberger, 2004). A comprehensive discussion of estimation methods can be found in the thorough review by (Hlavackova-Schindler, Palus, Vejmelka, & Bhattacharya, 2007) and the implications of finite sample size on entropy estimations are detailed in (Roulston, 1999).

There are several toolboxes available to calculate transfer entropy from raw time series. For example, PROCESS NETWORK (Ruddell & Kumar, 2009a, 2009b; Ruddell, Oberg, Garcia, & Kumar, 2010) offers a comprehensive environment for computing transfer entropy and other useful information-theoretic measures. The software runs on MATLAB and takes as input raw time series in vector form. INFORM is another versatile and powerful toolbox written in C that can be used to run computations of transfer entropy and other information-theoretic quantities (Moore, 2017; Moore, Valentini, Walker, & Levin, 2017) in JULIA, PYTHON, and R; a less complete version is also available in MATHEMATICA. The input to INFORM is equivalent to PROCESS NETWORK. Other valuable tools include JIDT (Lizier, 2014), TRENTOOL (Lindner, Vicente, Priesemann, & Wibral, 2011), and MuTE (Montalto, Faes, & Marinazzo, 2014), which may be particularly valuable when dealing directly with continuous-valued datasets.

Refinements of the Concept of Transfer Entropy

There are several directions that can be pursued to generalize the notion of transfer entropy beyond Equation (5). First, one may consider the possibility of interactions between the two processes that involve more than a single time step, as originally posited by (Marschinski & Kantz, 2002; Schreiber, 2000). In other words, rather than looking only at the predictive power of the present on the future time step, one would take into account past time steps for both X and Y. Currently, INFORM allows for considering an arbitrarily long time history for X, to incorporate the influence of its past on its future. This approach helps in avoiding to misattribute effects that are actually explainable from the past of X to being from Y, thereby mitigating spurious inferences of causality. Further details on the technical importance of this approach along with criteria for the selection of appropriately long time histories can be found in the first Chapter of (Wibral, Vicente, & Lizier, 2014) Nevertheless, extending the length of the time histories escalates the computational cost of estimating transfer entropy, through a massive increase in the sample space.

An alternative strategy to treat delayed interactions would be downsampling the dataset. However, this is not always possible, due to the finite duration of the experimental acquisitions that would challenge the accuracy of transfer entropy estimation. Another option is to include a delay, sometimes called a time lag, in the definition of transfer entropy, such that we would look for the effect of past time steps on the future. PROCESS NETWORK allows for considering a common delay for both X and Y, such that one would replace the subscript t with t-τ in Equation (3), where τ is the time lag. This method has been recently questioned in (Wibral et al., 2013), who have suggested that only Y should be delayed in order to avoid incorrect predictions of cause-and-effect relationships. By monitoring transfer entropy as a function of the delay, one could isolate peaks in information transfer that shed light on the time scale of the interaction between the animals, which are not instantaneous.

Another line of expansion for transfer entropy consists of the use of a symbolic approach to encode the behavior of the animals (Porfiri & Ruiz Marín, 2017b). Symbolic dynamics is a classical area of research in dynamical systems theory, whose foundations date back to the pioneering work of Hadamard (Hadamard, 1898). With respect to animal behavior, a symbol may be viewed as a locomotory bout, identified from the time series of the motion of the individual. Thus, the motion of an animal is intuitively treated as a sequence of locomotory bouts, summarizing its behavior from the position or orientation.

By aggregating m consecutive time steps through ordinal patterns, we create time series taking values on a sample space of $m!$ symbols. The advantage of a symbolic approach lies in the reduced computational cost of estimating so-called permutation, or symbolic, transfer entropy (Staniek & Lehnertz, 2008), due to the small sample space. At the same time, symbolic transfer entropy naturally integrates memory effects in animal interactions, whereby each symbol encodes two or more consecutive time steps.

The approach seems also amenable for coping with multisensory cues in animal interactions. One might create symbols that combine measurements from multiple sources beyond behavioral observations or trajectory tracking, such as flow measurements for the case of swimming fish. While neither PROCESS NETWORK nor INFORM allow for computing permutation transfer entropy directly, one needs only to develop a separate code to preprocess the dataset and convert it into symbols to be input to either of these toolboxes.

How to Gain Biological Insight through Statistical Analysis

To identify the cause and the effect from experimental data, we compute net transfer entropy
$$\text{NetTE} = \text{TE}^{Y \rightarrow X} - \text{TE}^{X \rightarrow Y} \tag{8}$$
If net transfer entropy is different than zero, then we speak of an overall influence between the two processes, such that one process will have a greater influence on the other. In other words, a nonzero value of net transfer entropy is associated with a degree of asymmetry in the interaction. If net transfer entropy is positive, we say Y influences X more than how much X influences Y. Similarly, if transfer entropy is negative, then X influence Y, more than the other way around. To draw statistically meaningful conclusions from n independent trials of the same experimental condition, we can simply perform a two-sided t-test to test the null hypothesis that the mean of NetTE is equal to zero with a given level of confidence, typically 95%.

Bootstrapping can also be a viable option to test whether causal relationships underlie the empirical datasets, but care must be placed in preserving temporal correlations within each time series (Chávez, Martinerie, & Le Van Quyen, 2003; Verdes, 2005; Vicente et al., 2011). Excessive shuffling of the time series may produce incorrect predictions, where causality is artificially inferred. Alternatively, a surrogate dataset may be generated by pairing subjects that did not interact with each other but were part of the same experimental condition. Following this approach, one may generate up to $n(n-1)$ surrogate trials from which to create a distribution for net transfer entropy. The mean of the distribution will represent an average, spurious, interaction between time series. By testing that a repetition of an experimental condition lies in the tails of the distribution, one may infer a cause-and-effect relationship in that particular condition. More specifically, if that repetition has a value of net transfer entropy that is in any of the tails of the distribution, we may argue that a cause-and-effect relationship exists between the time series. The sign of net transfer entropy will tell the direction of influence.

If a control condition is available, in which there is not interaction between the two individuals, then one can contemplate the use of an ANOVA to compare the means of net transfer entropy between the experimental and control conditions, similar to the approach presented in (Vicente et al., 2011) for neuroimaging experiments. Although this scheme is certainly desirable, control conditions are difficult to obtain in experimental settings with live animals, whereby controlling for their interaction may disrupt the behavioral repertoire of the animals. Ultimately, this will challenge the comparison between control and experimental conditions. For example, the locomotory activity of a social animal may completely change when it is in isolation, potentially causing a stress response that could manifest in freezing or erratic movements. The sample space underlying the computation of transfer entropy will thus vary between control and experimental conditions, possibly masking potential causal relationships or amplifying secondary effects.

Beyond net transfer entropy, one may compute total transfer entropy
$$\text{TotTE} = \text{TE}^{Y \rightarrow X} + \text{TE}^{X \rightarrow Y} \tag{9}$$
This quantity should be related to the overall degree of mutual influence of the two individuals. From total transfer entropy, one may be able to discriminate the case in which a predator and a prey are mutually influencing each other from the case in which they are not interacting. In both cases, net transfer entropy should be zero, but in the former, we expect to detect a large value of total transfer entropy. When a control condition is available, then a t-test might be used to test whether total transfer entropy is different from zero, otherwise the use of a surrogate dataset from randomly pairing individuals in the experimental condition might be pursued.

Beyond net and total transfer entropies, it may also be useful to simply consider the values of transfer entropies $\text{TE}^{Y \to X}$ and $\text{TE}^{X \to Y}$. Using ANOVA, one may compare $\text{TE}^{Y \to X}$ and $\text{TE}^{X \to Y}$ with each other for each condition and with the corresponding quantities for other conditions. Following this process, one would recover equivalent claims on the relative influence between the processes to those bestowed by NetTE, and also test differences in $\text{TE}^{Y \to X}$ and $\text{TE}^{X \to Y}$ across conditions to gain insight on the explanation for findings on NetTE and TotTE.

If sufficient data are available, all these analyses could be performed in sequential segments of the datasets to explore time-effects in the interaction, which should not be *a priori* dismissed. This can be especially relevant if one expects a highly dynamic interaction between the individuals, like in leader-follower interactions, where leadership and followership roles can change in time and with the biological context (King, Johnson, & Van Vugt, 2009). Beyond the study of time-effects, the idea of segmenting may be useful in other scenarios, where interaction between the individuals is intermittent or sporadic. In this case, it may be appropriate to isolate instances of interactions, so that they will not be diluted in the analysis.

Finally, I would like to mention a recent research direction that we have pursued to combine transfer entropy with other methods for the detection of cause-and-effect relationships (Mwaffo et al., 2017a). Our idea is grounded on the notion of receiver operating characteristic (ROC) curve, which is typically used to ascertain the performance of a binary classifier (Fukunaga, 2013) and we have first introduced in the study of animal behavior in (Butail et al., 2016). The ROC curve plots the number of true positives against false positives for a range of thresholds on the classifier. A perfect classifier has 100% true positive rate for any false positive rate. Following the general approach in (Barreno, Cardenas, & Tygar, 2008), we combine transfer entropy with other classifiers in a maximum likelihood sense, such that resulting optimal classifier maximizes the true positive rate for a given false positive rate. The combined classifier should be designed based on experiments in which the cause-and-effect relationship is known *a priori*. Once the classifier is designed, its use in detecting causality follows the implementation of net transfer entropy detailed earlier.

As a closing remark to this subsection, I would like to emphasize that irrespective of the statistical design, it is advisable to verify the robustness of predictions by varying some of the parameters selected for the computation of transfer entropy. Simple pre-processing of the datasets can be used to explore the role of binning and sampling frequencies. Through PROCESS NETWORK and INFORM, one might gain further confidence in the validity of the biological findings, by systematically exploring the role of the delay in the prediction and the length of the time histories, respectively. I would question significant results that are obtained for a very limited set of combinations of these computational parameters, or at least warrant the researcher to delve into possible biological implications for the reasons underlying a strong dependence on simulation parameters.

An Important Point of Discussion: Is it Really Causality?

The idea of apprehending a cause-and-effect relationship from transfer entropy requires some more thoughts and discussion. As acknowledged by Wibral et al. (Wibral et al., 2013), transfer entropy is based on Wiener's principle of causality, which may not be fully revealing of true causal interactions in every domain of application.

Perhaps, the best terminology to be used for describing the type of relationships that could be inferred based on Wiener's principle is "predictive information transfer", which should be distinguished from "causal information flow" (Lizier & Prokopenko, 2010). Quoting the work of Lizier and Proponenko:

> Causal effect refers to the extent to which the source variable has a direct influence or drive on the next state of a destination variable, i.e. "if I change the state of the source, to what extent does that alter the state of the destination?" (p. 605)

The difference between the two concepts is subtle and, generally, predictive information transfer may not correspond to causal information flow.

In the context of animal behavior, predictive information transfer is associated with the possibility

to predict the behavior of an individual from a given stimulus (such as a conspecific or a predator), whereas causal information flow relates to the opportunity of modulating the behavior of the individual by altering the stimulus. This very manipulation requires that the experimenter has the ability to intervene on the stimulus, that is, manipulate its behavior and simultaneously measure the response of the individual, if any. Although this may not be feasible for a typical observational experiment in which a live stimulus is employed, our robotics-based platforms constitute a first promising step toward establishing authentic causal information flows. In this vein, our experimental results on zebrafish-robot interaction offer evidence for the possibility of detecting some *a priori* known cause-and-effect relationships from transfer entropy. Through the empirical findings summarized below, we suggest the feasibility of associating causal information flow with predictive information transfer in animal behavior.

Experiments

Where Animal Behavior and Robotics Meet

The recent surge of robots to produce a desired behavioral response in animals is grounded upon the fundamental concept of "social releasers" developed by Tinbergen (Tinbergen, 1948). Robots can be engineered to exhibit precise morphological characteristics and move in sophisticated locomotory patterns, which can be leveraged in the design of unprecedented experiments on emotions, perception, and cognition (Klein, Stein, & Taylor, 2012; Krause, Winfield, & Deneubourg, 2011).

The first demonstration of robotics in the investigation of the response of live animals likely dates back to the work by Michelsen et al. (Michelsen, Andersen, Storm, Kirchner, & Lindauer, 1992), which examined the role of sound and wagging in the dance of honeybees (*Apis mellifera*). Since then, we have witnessed applications across a number of species, from birds to mammals. To name a few of the classical studies in ethorobotics, we could cite experiments by Takanishi et al. (Takanishi, Aoki, Ito, Ohkawa, & Yamaguchi, 1998) on the behavioral response of rats (*Rattus norvegicus*) to custom-made rat-like robots, the work of Patricelli et al. (Patricelli, Uy, Walsh, & Borgia, 2002) on the courtship of male satin bowerbirds (*Ptilonorhynchus violaceus*), and research by Kubinyi et al. (Kubinyi et al., 2004) on the interactions between dogs (*Canis lupus familiaris*) and commercially available dog-like robots.

Beyond zebrafish on which our group has heavily focused (Butail, Abaid, Macrì, & Porfiri, 2015), robotics-based platforms have been used to study a large number of fish species, including bluefin killifish (*Lucania goodei*, Phamduy, Polverino, Fuller, & Porfiri, 2014), giant danios (*Devario aequipinnatus*, Aureli & Porfiri, 2010), golden shiners (*Notemigonus crysoleucas,* Abaid, Marras, Fitzgibbons, & Porfiri, 2013; Marras & Porfiri, 2012; Polverino, Phamduy, & Porfiri, 2013; Swain, Couzin, & Leonard, 2012), guppies (*Poecilia reticulata*, Landgraf et al., 2016), mackerels (*Scomber scombrus*, Kruusmaa, Rieucau, Montoya, Markna, & Handegard, 2016), mosquitofish (*Gambusia affinis*, Polverino & Porfiri, 2013), Siamese fighting fish (*Betta splendens*, Romano et al., 2017), three-spined sticklebacks (*Gasterosteus aculeatus*, Faria et al., 2010; Ward, Sumpter, Couzin, Hart, & Krause, 2008), and trunkfish (*Mormyrus rume,* Donati et al., 2016; Worm, Kirschbaum, & von der Emde, 2017). Collectively, these efforts have contributed to an improved understanding of fish behavior, illuminating several open areas of research, spanning from hydrodynamic advantage in schooling to the role of morphophysiological factors on aggression and courtship.

Our Robotics-based Approach to Validate the use of Transfer Entropy

In the context of causal relationships underlying zebrafish behavior, we have developed two different robotics-based platforms that offer the researcher the opportunity to create versatile, customizable, and highly controllable stimuli. The two platforms share the same design concept, which leverages 3D-printing and robotic manipulators to attain life-sized robotic models that can be maneuvered at a speed comparable to that of their live counterpart. The first platform was developed to study social behavior,

while the second targeted fear response. These platforms offer time resolved datasets on the interactions between controlled stimuli and live subjects, which we leverage to assess the feasibility of transfer entropy analysis to infer causal relationships underlying animal behavior. Here, we briefly summarize both.

The platform to study zebrafish social behavior was first presented in (Butail et al., 2014), as a modification of a setup that we developed for investigating courtship in bluefin killifish (Phamduy et al., 2014). This original platform used a robotic arm that consisted of three rigid links, each actuated by a separate servo-motor. The servo-motors were controlled using pre-programmed microcontroller, which allowed for oscillating and maneuvering a transparent support rod in two dimensions. 3D-printed replicas of zebrafish were connected to the bottom of the support rod to produce the visual stimulus of zebrafish swimming in a horizontal plane and beating their tails. Figure 2 shows a computer-aided-drawing of the platform, along with nomenclature for its components.
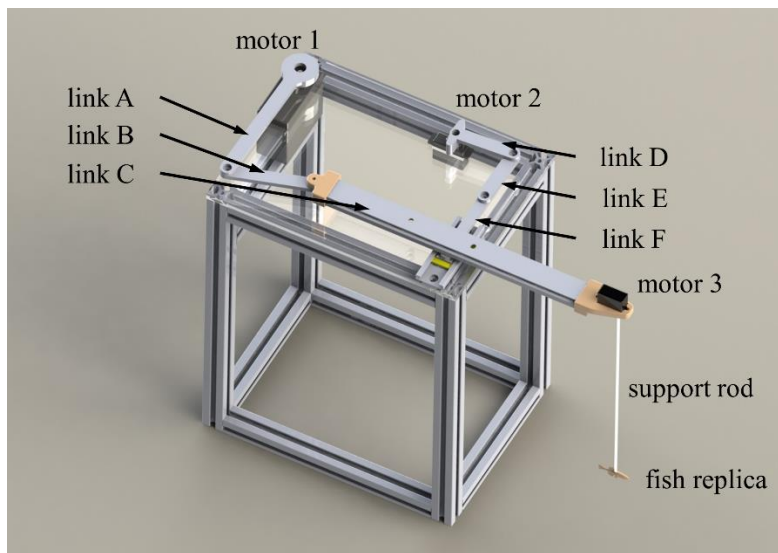
Figure 2. Computer aided design of the robotic platform. A 3D-printed replica was attached to the bottom end of the support rod and was passively translated on a horizontal plane, along desired trajectories while imposing an oscillatory motion. [Reproduced from (Phamduy et al., 2014)]

This platform was designed for conducting binary choice tests, in which a focal subject would be placed in the central compartment of a tripartitioned tank and the robotic stimulus in one of the lateral compartments. In (Butail et al., 2014), we presented experimental results in which the 3D-printed model consisted of a replica of a single zebrafish and its motion followed an eight-shaped trajectory. Figure 3 illustrates a top view of the experimental tank and the replica used in the trials. Two experimental conditions were included in (Butail et al., 2014). In condition "Replica," we installed two platforms to provide an equivalent visual background to the focal subject, but a replica of a zebrafish was mounted on only one side. In condition "Conspecific," a live stimulus zebrafish was placed in one of the side compartments and no robotic platforms were installed.

Videos were recorded from an overhead camera and the motion of the focal subject and the stimulus were tracked using the in-house developed software PEREGRINE, which was first presented by our group in (Butail, Bartolini, & Porfiri, 2013). The software was later refined to accommodate shape tracking and fusion of multiple views in (Bartolini, Butail, & Porfiri, 2015; Ladu, Bartolini, et al., 2015). We used the tank-wide coordinate, along which the motion of the replica and stimulus fish was wider, to measure information transfer between the focal subject and the stimulus (replica or conspecific). Our hypotheses at the beginning of the study were that: (i) in condition Replica, transfer entropy would reveal that the replica was causing the behavior of the focal subject, and (ii) in condition Conspecific, neither the focal subject nor the stimulus fish would cause the behavior of the other, that is, no asymmetry should be detected from the interaction.
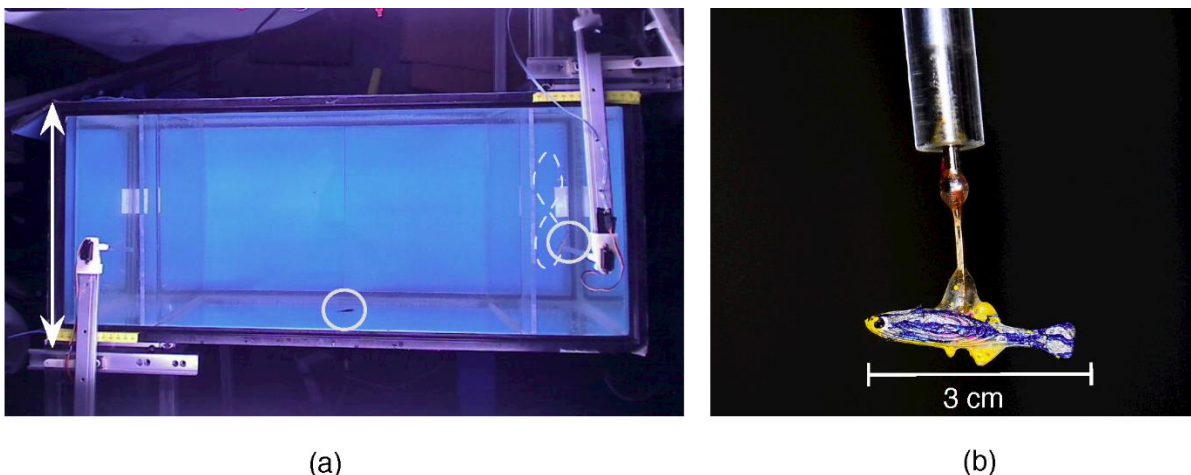
Figure 3. Experimental setup for ethorobotics experiments on zebrafish social behavior. (a) Top view of the tripartitioned tank with two robotic platforms on each side. The fish in the central compartment and the replica in the right compartment are circled. The dashed curve is the figure-eight that is travelled by the replica and the bidirectional arrow labels the vertical axis that is used for the transfer entropy analysis. (b) 3-D printed painted replica of the zebrafish that is attached to the support rod. [Reproduced from (Butail et al., 2014)]

The same platform was then used in (Bartolini et al., 2016; Ladu, Mwaffo, et al., 2015) to study the effect of caffeine administration and body size on social behavior of zebrafish. In both studies, we replaced the figure-eight with a simpler linear motion and opted for the use of a different 3D-printed model, consisting of a small shoal of zebrafish, rather than a single individual. This latter choice was motivated by the need to increase the saliency of the robotic stimulus, potentially strengthening its attractiveness to focal fish. Grounded in prior evidence indicating that shoaling tendency increases with anxiety (Maximino et al., 2010), in Ladu, Mwaffo, et al. (2015) we sought to test the hypothesis that increasing caffeine concentration would lead to a stronger interaction between the focal subject and the robotic stimulus. Resting upon the preference of social fish towards similarly-sized conspecifics (Hoare, Krause, Peuhkuri, & Godin, 2000), in (Bartolini et al., 2016) we attempted to test the hypothesis that the influence of the shoal of replicas on the focal subject would be maximized when using models of the same size as live zebrafish.

Since these first three studies, we have invested considerable effort in improving upon the design of the robotic platform to afford a more realistic motion for the robotic stimulus. The driving motives for our work were the needs to: (i) enable the robotic stimulus to be maneuvered along the water column, similar to live subjects (Macrì et al., 2017), and (ii) generate aperiodic, information-rich motions that could better proxy a stationary stochastic process. Toward these aims, we modified the design of the robotic platform to include an additional degree of freedom for diving, via a threaded rod that was actuated by a dedicated motor. Inspired by the work of Landgraf et al. (2016), we also improved on the appearance of the 3D-printed models by using glass, rather than painted, eyes. Figure 4 shows a live zebrafish close to the improved replica. Experiments in Ruberto et al. (2016) sought to verify that the inclusion of the diving motion of the robotic stimulus would increase its influence on live subjects. At the same time, we also attempted to clarify the relative role of the visual appearance of the robotic stimulus, through controlled experiments using a transparent model in lieu of the replica.

Figure 4. Zebrafish swimming close to an improved 3D-printed life-sized replica with glass eyes, suspended by a transparent rod.

Interestingly, our recent work (Ruberto, Polverino, & Porfiri, 2017) suggests that the appraisal of this improved robotic stimulus (Ruberto et al., 2016) by zebrafish is equivalent to their appraisal of a conspecific. Specifically, zebrafish were shown to display an equivalent degree of attraction for live zebrafish separated by one-way glass partitions and the robotic stimulus. Encouraged by this evidence, we have attempted to take a further, critical step in the development of the robotic platform. Specifically, we have integrated a closed-loop control system that allows the replica to respond in real time to the behavior of the live subjects. In our latest setup, we can track a live subject in three dimensions using two independent cameras and use the position of the live subject to maneuver the robotic stimulus in real-time. Figure 5 illustrates the working principle of the setup. Experimental results from this study have been recently been published (Kim, Ruberto, Phamduy, & Porfiri, 2018), and the symbolic analysis of a small subset of them has been also published (Porfiri & Ruiz Marín, 2017b).
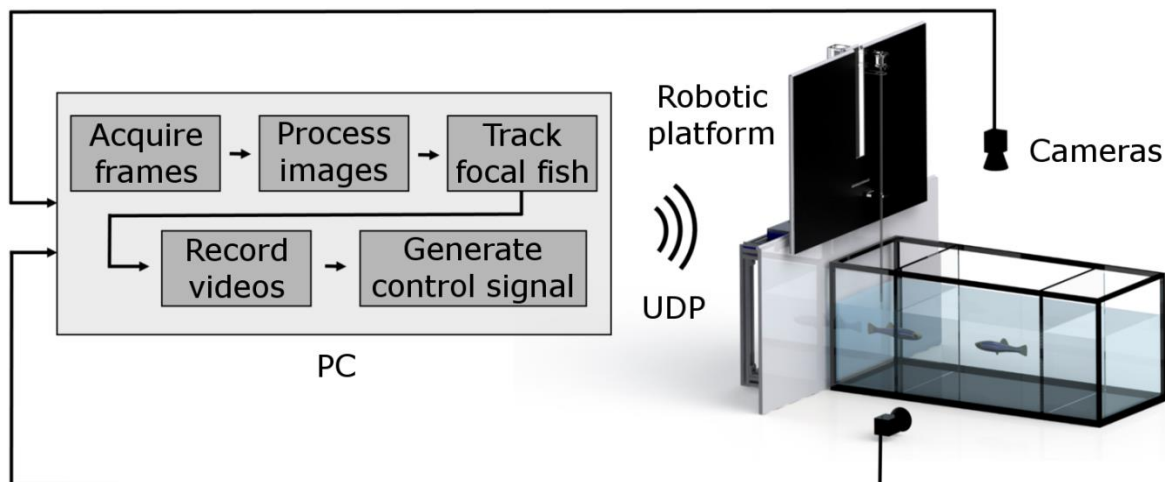
Figure 5. Schematics of the interactive robotic platform. The base station receives an input (images) from the top and front camera. The images are processed and the fish is tracked in real-time to generate a feedback control signal that is used to maneuver the replica. For clarity, the focal fish and the robotic stimulus are magnified. [Reproduced from (Kim et al., 2018)]

With respect to fear response, we have taken an analogous robotics-based approach through the design of a new platform that simulates anti-predatorial response of zebrafish in Neri et al. (2017). Our setup shares similarities with the work of Hu et al. (2015), where a circular arena was also used to investigate information transfer between a prey and a predator fish from their position. Based on prior work by our group (Ladu, Bartolini, et al., 2015), the robotic stimulus was designed after a red tiger oscar fish (*Astronotus ocellatus*), an allopatric predator of zebrafish. Figure 6 shows the 3D-printed model of the red tiger oscar fish, along with its live counterpart.
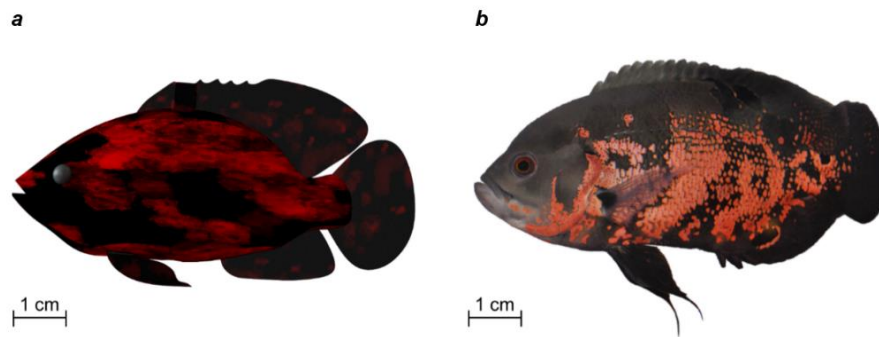
Figure 6. Stimuli used in the fear-evoking test. (a) Painted 3D-printed model of a red tiger oscar fish and (b) its live counterpart. [Reproduced from (Neri et al., 2017)].

The replica was maneuvered along preprogrammed, information-rich, circular trajectories using a robotic arm in the central portion of the arena. A cam system was used to induce oscillatory movements in the stimulus. Figure 7 illustrates the working principle of the setup. Zebrafish were free to swim in the outer annular region of the arena, separated by a transparent panel from the stimulus. In addition to studying individual swimming of red tiger oscar fish and zebrafish in the arena, we conducted three experimental conditions. Specifically, we performed a condition using a live predator ("Live") as the stimulus; a condition in which the robotic replica replaced the live predator ("Replica"); and a control condition with no replica attached to the robotic platform ("CTRL Replica"). Through the application of transfer entropy we sought to demonstrate an influence of the robotic stimulus on zebrafish in condition Replica and explore mutual influence between the live predator and zebrafish in condition Live.
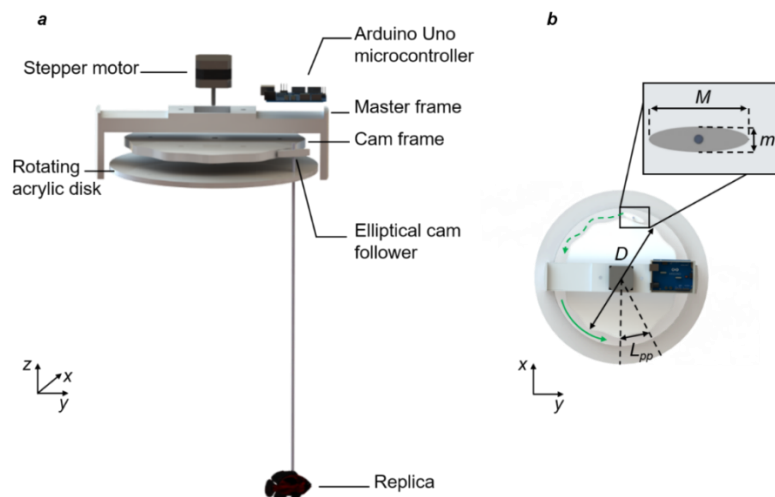

Figure 7. Computer-aided design of the robotic platform to study fear response of zebrafish. (a) Exploded schematics and (b) illustration of the working principle of the robotic platform from a top view with a close-up of the elliptical cam follower. The trajectory of the replica is displayed as a dashed green line, while the motion of the rotating acrylic disk is shown as a solid green line. [Reproduced from (Neri et al., 2017)].

Zebrafish Interactions with Zebrafish-Inspired Robots

We start our synoptic presentation of results with our first study on the use of transfer entropy to study zebrafish-robot interactions (Butail et al., 2014). Figure 8 displays the tank-wide position for two sample trials in conditions, Conspecific and Replica, which we use to compute transfer entropy. The

position is scaled with respect to motion along the width of the tank, such that ±1 correspond to positions close to the walls. It is very easy to differentiate the two conditions: in condition Conspecific, we observe a continuous adaptation of the motions of both the focal and stimulus fish, while in condition Replica, the stimulus moves consistently along the same path, irrespective of the motion of the focal fish.
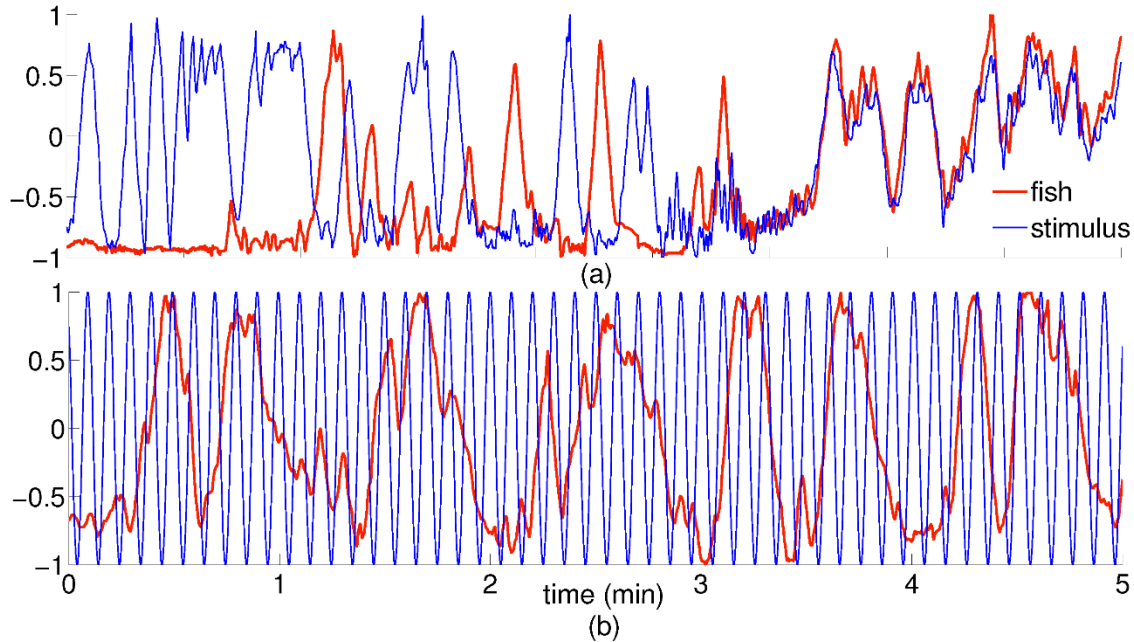

Figure 8. Sample tank-wide positions of focal fish and stimuli (conspecific and replica) in ethorobotics experiments on social behavior of zebrafish. (a) Normalized tank-wide positions for a fish (red) interacting with a conspecific (blue); for this sample, $TE^{fish \to stimulus} = 0.615$ bits and $TE^{stimulus \to fish} = 0.524$ bits, where "stimulus" is the conspecific. (b) Sample tank-wide positions for a fish (red) interacting with a replica (blue); for this sample, $TE^{fish \to stimulus} = 0.357$ bits and $TE^{stimulus \to fish} = 0.638$ bits, where "stimulus" is the replica. [Reproduced from (Butail et al., 2014)].

These sample trajectories bring to light interesting features that we seek to systematically dissect through the use of transfer entropy. For example, Figure 8a suggests that in condition Conspecific, there are long time segments in which the focal subject and the live stimulus stay close to each other. As one can observe from the negative and positive peaks, sometimes the focal subject will take the lead, and other times it is the stimulus that will anticipate the motion of the focal subject. In contrast, Figure 8b hints at a different form of interaction, in which the focal subject will respond to some of the locomotory bouts of the replica. Specifically, the focal subject tends to reverse its direction of motion when the replica reverses its motion, as observed by the peaks and dips in the two time traces.

We adopt the standard definition of transfer entropy in Equation (5) that uses a single time step in the prediction and does not account for delays. To compute transfer entropy between the tank-wide position of the focal fish and the stimulus, we bin the normalized positions using seven equal bins ($|\Omega| = 7$) and downsample the data such that the sampling time interval for the analysis is equal to 2 seconds, giving a total of 150 time steps ($T = 150$). These values were selected on the basis of theoretical arguments proposed by (Baptista et al., 2012), along with biological insight based on zebrafish behavior. With respect to the latter, seven bins correspond to a bin size of a little more than one fish body length and a time interval of 2 seconds is a rough estimate of how long the fish will take to swim between the far ends of two neighboring bins at a nominal speed of one body length per second.

In retrospect, one may argue that the time series was too short, but we verified the stability of our predictions by randomly selecting an experimental trial and exploring the sign of net transfer entropy as a function of the number of bins and the duration of the time step. Our results indicate that equivalent

predictions are obtained for most values of the sampling time interval greater than 0.5 seconds. Shorter sampling will reverse the sign of the net transfer entropy, possibly due to the correlation between the noise from the fish and the replica trajectories. Specifically, correlations between the motion of the robotic stimulus, which moves fast along the tank width, and the noise in tracking the fish trajectory may be detected at small temporal resolutions.

Numerical values of transfer entropies for the times series in Figure 8 are presented in the caption, exemplifying the expected difference between the two experimental conditions[5]. This evidence is echoed by Figure 9, which shows aggregated results on 20 naïve fish, 10 for condition Conspecific and nine[6] for condition Replica. In agreement with our hypotheses, we found that: (i) transfer entropy is successful in detecting a causal relationship underlying the response of the focal fish to the robotic stimulus, and (ii) the influence of the live stimulus on the focal subject does not differ from the influence of the focal subject on the live stimulus. Whereas the robotic stimulus was independently controlled, the live stimulus was always visually aware of the focal subject, and therefore likely to be influenced by it. Later, we further corroborated these results using our symbolic approach in (Porfiri & Ruiz Marín, 2017b).
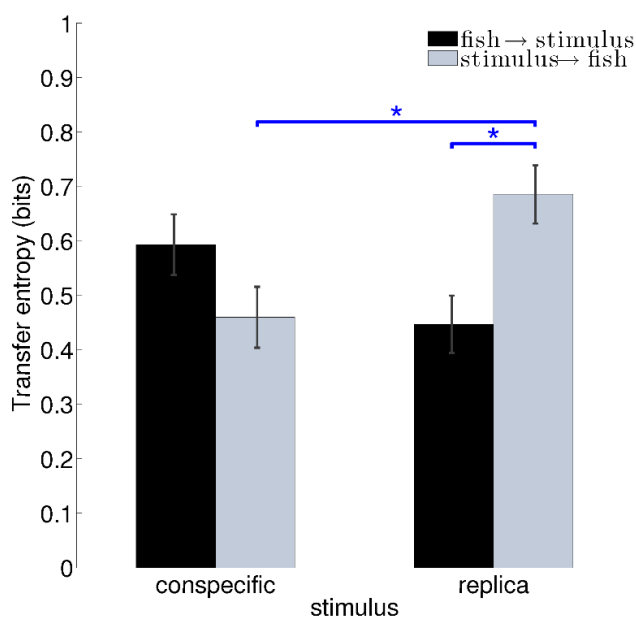

Figure 9. Transfer entropy computation for experiments on the interaction between two live zebrafish (Conspecific) and between a zebrafish and a life-sized replica maneuvered by a robotic arm along pre-programmed trajectories (Replica). Error bars represent ± standard errors and overhead bars with an asterisk indicate statistically significant differences. [Reproduced from (Butail et al., 2014)].

Between conditions, we found that the transfer entropy from the stimulus to the focal subject was significantly greater when the stimulus was a replica. This result may be regarded as indirect evidence in favor of the efficacy of robotics-based platform in eliciting reliable and repeatable behavioral responses (Cianca, Bartolini, Porfiri, & Macri, 2013; Ladu, Bartolini, et al., 2015; Spinello, Macrì, & Porfiri, 2013). In this sense, the marked asymmetric influence of the replica on the focal fish might suggest that consistent responses are favored by the use of robotic compared to live stimuli.

---

[5] Because the motion of the replica was programmed to be periodic, increasing the length of the time history in the computation of transfer entropy should reduce the value of $TE^{fish \rightarrow stimulus}$ and help in improving the accuracy of the inferences.

[6] In all of our experiments on zebrafish, we systematically used 10 fish per condition. Sometimes, one or two subjects were excluded from the analysis if anomalous behavior was observed, such as excessive freezing response.

Digging More into Zebrafish Response to Zebrafish-Inspired Robots

Figure 10 summarizes results on information transfer between the robotic stimulus, composed of a shoal of replicas, and zebrafish exposed to acute caffeine concentration (Ladu, Mwaffo, et al., 2015). Similar to Figure 9, the computation of transfer entropy was based on the tank-wide positions of focal fish and stimuli and the bin size was selected to correspond to approximately one zebrafish body length. Four experimental conditions were included in the study, in which caffeine concentration varied from moderate to high levels, with the expectation of eliciting a dose-dependent response in the behavior of focal subjects. Our findings suggest that at intermediate and high caffeine concentrations the shoal of 3D-printed replicas influenced the behavior of the focal subject. More precisely, the stimulus influenced the focal subject more than the focal subject influenced the pre-programmed stimulus in two conditions, although we did not register differences among conditions in terms of information flow between fish and replica. Combined with the analysis of other salient measures of animal behavior, these findings point at an anxiety-related behavioral response evoked by caffeine, which, in turn, produced a more robust shoaling tendency.
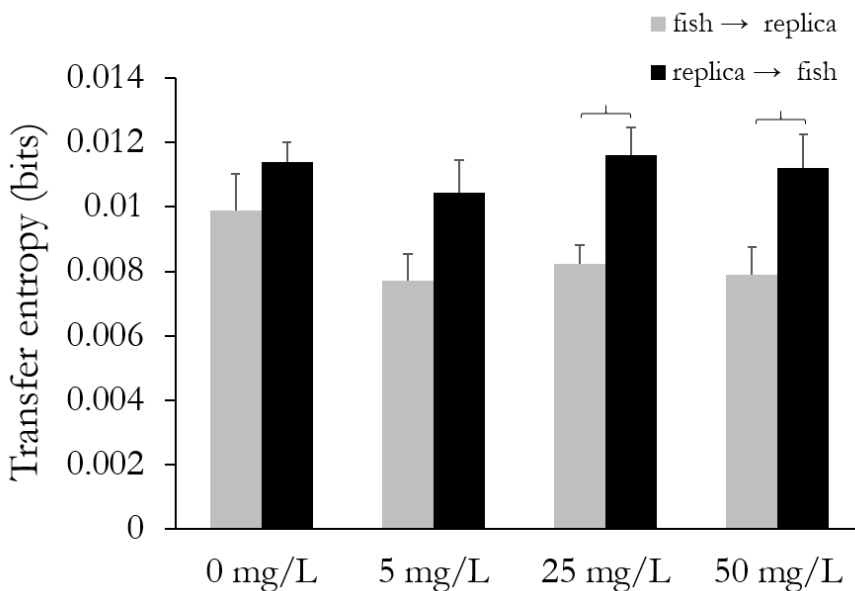

Figure 10. Transfer entropy for experiments on the interaction between live zebrafish treated with caffeine and a shoal of life-sized replicas maneuvered by a robotic arm along pre-programmed trajectories. Experiments refer to four conditions at caffeine concentrations: 0, 5, 25, and 50 mg/L. Error bars represent ± standard errors and curly braces indicate statistically significant differences. [Reproduced from (Ladu, Mwaffo, et al., 2015)].

Based on an equivalent robotics-based platform, Figure 11 demonstrates the influence of shoals of replicas of 2 and 3 cm in size on focal subjects (Bartolini et al., 2016). Transfer entropy computation mirrored the analysis in Figures 10 and 11, whereby we focused on the tank-wide position and used a bin size corresponding to one body length. Combined with evidence from other behavioral metrics, these results suggest the possibility of some form of a leader-follower relationship underlying the observed response, although the precise role of either the focal fish or the replicas is difficult to quantify. It is conceivable that zebrafish perceived shoals of life-sized replicas as conspecifics and, therefore, engaged in a highly dynamic interaction. The inference of leader-follower relationships was also investigated in our recent work (Butail et al., 2016), dealing with synthetic data generated from a data-driven model of zebrafish social behavior. An interesting feature of the analysis presented in Figure 11 is the treatment of the control condition, in which we randomly paired each control subject with the shoal of replicas from another condition. This approach was used to generate surrogate data on which to confirm the validity of transfer entropy in detecting causality.
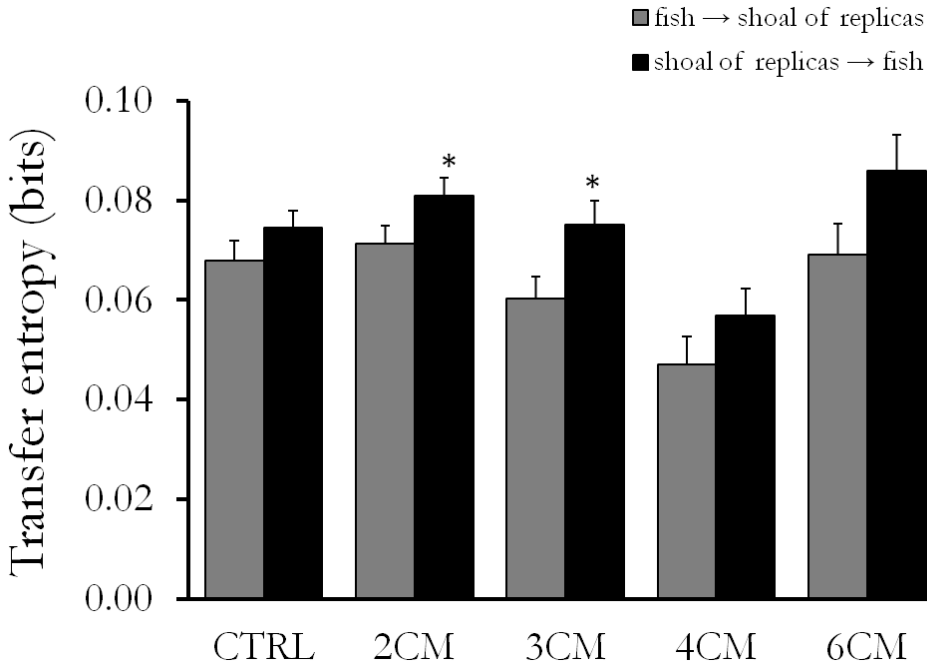
Figure 11. Transfer entropy for experiments on the interaction between live fish and a shoal of replicas of different body length maneuvered by a robotic arm along pre-programmed trajectories. Experiments refer to six conditions, including a control condition (CTRL) in which the lateral compartments were empty, and four experimental conditions with replicas of size: 2, 3, 4, and 6 cm. Error bars represent ± standard errors and a star indicates a significant difference between the two directions of information flow. [Reproduced from (Bartolini et al., 2016)].

Figure 12 illustrates information flow between zebrafish and the improved replica pictured in Figure 4. Results presented therein are again based on computing transfer entropy with respect to the tank-wide position of the focal fish and the stimulus, using a bin size of one approximately one body length. In agreement with our hypotheses, experimental results in (Ruberto et al., 2016) demonstrate that the three-dimensional motion of the replica influenced the behavior of focal subjects, and such a causal relationship was lost when the visual appearance or the motion along the water column were controlled. Another interesting claim that can be garnered from Figure 12 pertains to the increased flow of information from the stimulus to the fish in both the conditions in which the support rod was maneuvered along three-dimensional trajectories, with respect to the control condition, where we computed transfer entropy through surrogate data. This finding indicates that the transparent model was perceived by the subjects, likely due to: (i) visual signaling from the movement of the water generated by the platform, and (ii) acoustic signaling from the repeated collisions of the model on the partitions.
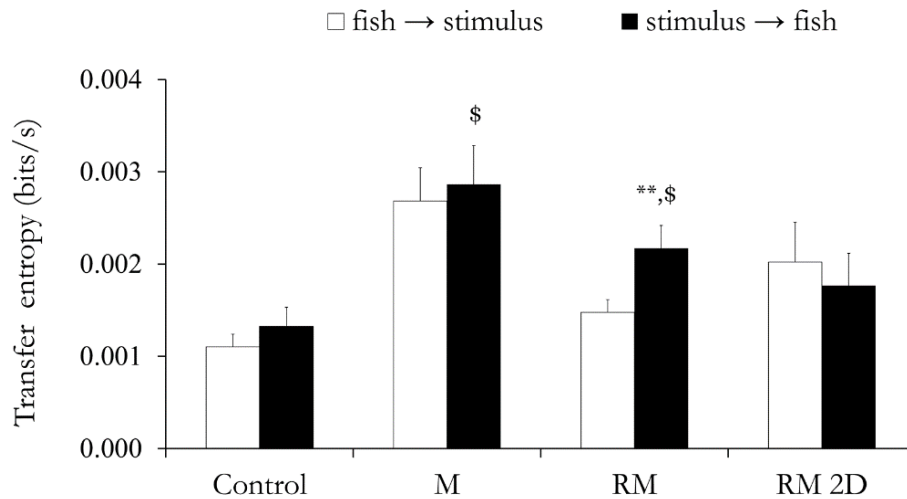
Figure 12. Transfer entropy for experiments on the interaction between live fish and an improved replica maneuvered by a robotic arm along three-dimensional pre-programmed trajectories. Experiments refer to four conditions, in which the replica was maneuvered along realistic three-dimensional trajectories (RM), the diving motion of the replica was abolished (RM2D), the replica was replaced by a transparent model moving along realistic three-dimensional trajectories (M), and neither a replica nor a transparent model were attached to the rod that was held fixed on the water surface. Error bars represent ± standard errors, asterisks indicate significant difference in transfer entropy from the stimulus to the fish and vice versa, and dollar symbol indicates a significant difference with respect to the Control condition. [Reproduced from (Ruberto et al., 2016)].

Beyond offering insight into the role of three-dimensional maneuvers and the visual appearance of the replica, the experiment in (Ruberto et al., 2016) constitutes a significant improvement toward a rigorous methodological validation of transfer entropy as a tool to infer causal relationships in animal behavior. Different from the afore-described experiments on zebrafish-robot interactions, here the trajectory of the replica was generated based on the motion of a live subject, scored in a previous observation. As a result, the replica was maneuvered along information-rich trajectories, in contrast with almost periodic motions which might hinder the rigorous application of transfer entropy.

I would like to close this subsection with discussion of our most recent paper on the interactions between zebrafish and zebrafish-inspired robots (Kim et al., 2018). Rather than attempting at summarizing the details of this experiment, which were rather intricate, I will try to summarize the key differences with respect to our previous work and the main insights we gained in the context of transfer entropy. First, in this experiment, we considered the possibility of closing the loop between animal behavior and robotics, such that the replica could be maneuvered in real-time based on the three-dimensional position of the focal fish. Second, we scored transfer entropy between the replica and the focal fish by using both the tank-wide coordinate, like we did in all our previous work, and the coordinate running through the water column. Interestingly, we found that cause-and-effect relationships depend on which coordinate is used for the analysis, such that one might misinterpret a specific interaction if behavior was scored from a single view. For example, if we observed only information transfer through the tank-wide coordinate, we might not be able to identify experimental instances in which the replica was controlled to follow in real-time the motion of the focal subject along the water column.

Zebrafish Interactions with Predator-Inspired Robots

Experiments on fear-response of zebrafish in (Neri et al., 2017) offer further compelling evidence in favor of the application of transfer entropy to study animal interactions. In this case, we measured transfer entropy from the angular coordinates of the stimulus and the focal fish by considering four sectors of 90°

each ($|\Omega| = 4$). We used a time interval between samples of 1/30s, corresponding to full resolution from the video. The consistency of our predictions was verified by systematically varying the bin size and the time interval with respect to these reference values. Specifically, we computed net transfer entropy by varying the number of bins from 3 to 10, while keeping the time interval to 1/30 s, and by downsampling the data until 1s.

Numerical values of transfer entropy are presented in Figure 13 for the four experimental conditions considered in this study. Our results indicate that transfer entropy from the stimulus to zebrafish did not vary as a function of the stimulus. On the other hand, transfer entropy from the zebrafish to the stimulus changed with the condition, such that focal fish transferred more information to the live predators than the robotic replica. This finding demonstrates that the live predator reacted to the focal subject, differently from the pre-programmed replica whose motion was independent of the focal subject. Combined with other behavioral metrics, this suggests that the visual feedback from the focal subject was responsible for an increased responsiveness of the predator.
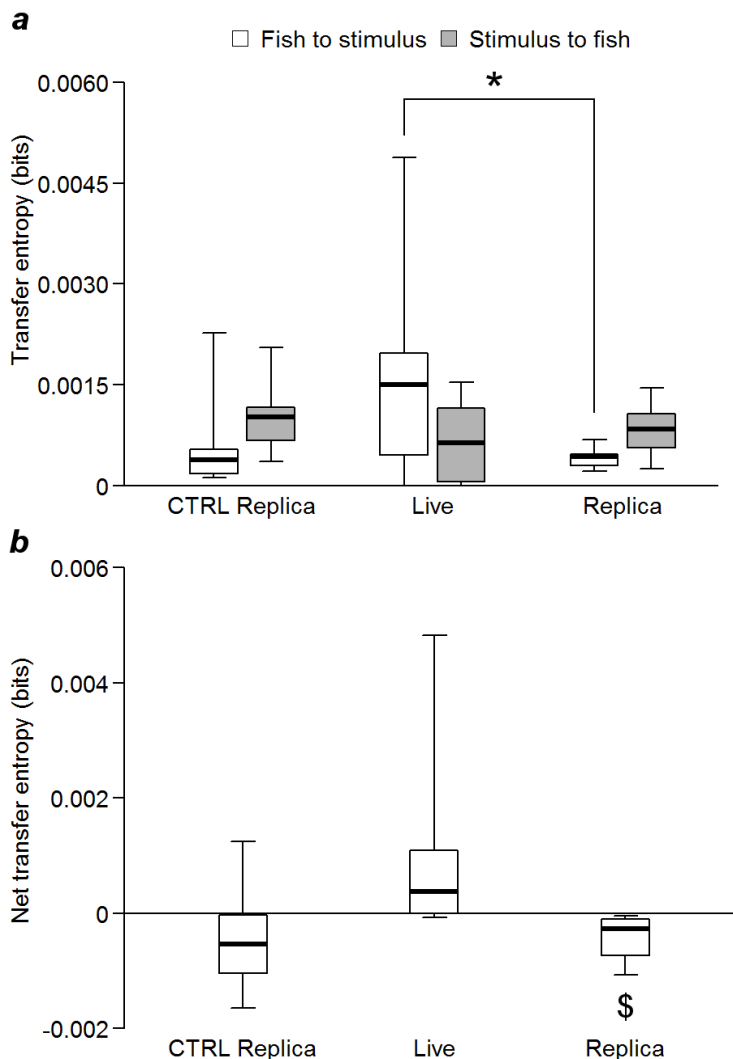

Figure 13. Transfer entropy for experiments on the interaction between live fish and fear-evoking stimuli. Experiments refer to three conditions, in which we used as stimulus a live red tiger oscar (Live), a replica of the predator (Replica), and a simple moving rod (CTRL Replica). The bottom and top of each box are the first and third quartiles, the band inside the box is the median, and whiskers identify the entire distribution of the data. (a) Transfer entropy from zebrafish to the stimulus (white bars) and from stimulus to zebrafish (grey bars) and (b) net transfer entropy. The dollar symbol indicates a significant difference in comparison with zero. [Reproduced from (Neri et al., 2017)].

In agreement with our predictions, the analysis of net transfer entropy registers a net information flow in condition Replica, with the stimulus causing the behavior of the focal subject, and failed to reveal a causal relationship between the stimulus and the zebrafish in condition CTRL Replica. Combined with behavioral metrics on avoidance response and locomotory activity, these results confirm our previous findings that a replica of the red tiger oscar fish constitutes a robust fear-evoking stimulus, which could potentially replace the use of live predators (Ladu, Bartolini, et al., 2015). The lack of net information flow between the focal subject and live predator indicates a mutual influence of the predator on the prey, and vice-versa. Just as the predator increases its responsiveness to the focal subject, it is possible that zebrafish becomes vigilant to the behavior of the stimulus.

Conclusions and Outlook on Future Work

In this targeted review, I presented a summary of the state-of-the-art in the application of transfer entropy to infer causal relationships in animal behavior. I tried to offer a balanced overview between theory and experiments, in an effort to empower readers with a set of methodological tools they could translate to other areas of animal behavior. From a theoretical point of view, this paper offers a synoptic introduction to the concept of transfer entropy and a practical guide for implementation on real data. The presentation of the experimental component revolves around the growing field of ethorobotics, which constitutes a powerful foundation for manipulating cause-and-effect relationships. Specifically, I focused on experiments in zebrafish-robot interactions as a means for validating the application of transfer entropy in the study of social behavior and fear.

Experiments on zebrafish-robot interactions are what I called the "small lure" in the title, and the "big fish" is the improved understanding of interactions in animals through the fundamental theoretical concept of transfer entropy. Transfer entropy promises the detection of causal relationships underlying behavioral interactions, and ethorobotics experiments on zebrafish provide an ideal testbed for validation of the approach. Recent efforts have started to expand on the application of transfer entropy to other animal systems, including bats and humans. Orange and Abaid (2015) have put forward a transfer entropy analysis of the interaction between two bats, supporting the intuition that information propagates from the foremost located individual to the subject in the rear. In our recent work on technology-mediated interactions between humans (Nakayama, Ruiz Marín, Camacho, & Porfiri, 2017), we have applied transfer entropy to study leader-follower relationships and found interesting evidence of behavioral plasticity. More insight on leader-follower relationship has been gained through controlled experiments on hydrodynamic interactions of pitching airfoils (Zhang, Rosen, Peterson, & Porfiri, in press), upon which we plan to formulate a novel methodology to quantify hydrodynamic interactions in schooling fish.

Perhaps, the most pressing open question that emerges from this targeted review is how to extend the approach to the study of animal groups. In other words, how can we scale from two to an arbitrary number of interacting individuals? The theoretical extension of the approach to animal groups is far from trivial, due to the dyadic nature of transfer entropy. Specifically, when computing transfer entropy for a chosen pair of animals in a group, the influence of the remaining animals may act as confound in determining whether a cause-and-effect relationship truly exists between the behaviors of the selected individuals. For example, if the two animals are isolated but they both interact with a third individual, transfer entropy might incorrectly pick up a causal relationship, when an interaction does not actually exist.

A possible way around the problem of spurious coupling between individuals has been presented by Bollt and Sun in a sequence of papers dealing with the notion of optimal causation entropy (Sun & Bollt, 2014; Sun, Taylor, & Bollt, 2015). The idea of causation entropy is based on a natural extension of transfer entropy to include relationships among many, more than two, systems. Put simply, a causal relationship between two systems is inferred by systematically assessing that such a relationship remains valid upon conditioning on any other system. Grounded on a mathematically robust framework, optimal causation entropy constitutes a promising methodology to infer causal relationships in animal behavior, as recently shown in (Lord et al., 2016) for insect swarms.

However, there are several reasons due to which one might still opt for transfer entropy, rather than other more sophisticated approaches. Among these reasons, I would like to mention the intuitiveness of transfer entropy from a conceptual viewpoint, the availability of a number of computer packages to perform computations, the large technical literature supporting the use of transfer entropy in complex systems, and the possibility of generating inferences with relatively small datasets. In this case, prudence is recommended to mitigate the occurrence of false inferences and enhance the accuracy of detecting true interactions.

A possible line of reasoning to gain confidence in the approach is to generate synthetic data sharing similar dynamic features to the experimental observations and test the validity of transfer entropy on those. Toward this aim, one may resort to data-driven models of collective behavior to generate datasets of fish behavior with known leader-follower interactions. By calibrating model parameters on experimental observations, one may construct realistic datasets on which to test the range of applicability of transfer entropy in terms of key factors, such as number of leaders, length of time series, and differences among individuals.

The application of transfer entropy on model data of collective behavior was first proposed in (Wang, Miller, Lizier, Prokopenko, & Rossi, 2012), and, recently, we have applied this approach to specific model data from zebrafish collective behavior (Butail et al., 2016; Mwaffo et al., 2017a). Beyond the originality of the endeavor, which constitutes the first application of transfer entropy in the study of collective behavior, Wang et al., (2012) should be mentioned for the technical approach to evaluate information transfer. Rather than the absolute position of each individual, the authors employ relative coordinates, thereby reducing the sample space for the computation. In (Tomaru et al., 2016), this line of argument has been used to study the effect of group size on the extent of social interaction among soldier crabs, by using relative orientation between individuals. From the analysis of experimental datasets, it is found that individuals in small groups tend to be modestly influenced by other group members. On the other hand, for larger groups, social interactions seem to play a significant role in the group collective response, whereby the influence of other group members on an individual tends to increase.

In a similar vein, our work (Butail et al., 2016; Mwaffo et al., 2017a) also uses angular coordinates. However, rather than considering the relative orientation, therein we study information transfer from the time series of the absolute angular velocities. By taking this approach, we simplify the problem of dealing with wrapping and unwrapping angles, which would induce potential artifacts in the computation of transfer entropy. Interestingly, by calibrating inferences on synthetic data, we can attempt at a first quantification of the effect of caffeine concentration on leadership in zebrafish groups (Mwaffo et al., 2017a) from data published in Neri, Ruberto, Mwaffo, Bartolini, and Porfiri (in press) thereby showing that caffeine-treated fish may be more likely to take on leadership roles in the group.

Curiously, in a parallel line of research, we have explored an equivalent strategy to infer causal relationships underlying the diffusion of alcohol-related policies in the United States (US) (Anderson et al., 2016; Grabow, Macinko, Silver, & Porfiri, 2016; Porfiri & Ruiz Marín, 2017a). Similar to social behavior of fish, some US states may take a leading role in the adoption of new policies, while others act as followers. Through transfer entropy, we attempted at clarifying the ideological, geographical, and political factors that modulate leader-follower relationships in this context.

I hope this targeted review will inspire scientists in animal behavior to explore the combined potentials of information theory and robotics toward the development of more refined methods of data analysis and highly controllable experimental schemes. At the same time, I hope that engineers will complete the reading of this paper with gained interest in the field of animal behavior and higher appreciation for the benefits that our research, as engineers, can bestow to animal behavior.

Focusing on the notion of transfer entropy, this targeted review exemplifies the value of new research at the intersection of animal behavior, applied mathematics, and robotics. Such an intersection constitutes a fertile area of study, whose growth relies on the collaboration of experts from different fields and the multidisciplinary training of junior scholars.

Acknowledgements

References

Abaid, N., Marras, S., Fitzgibbons, C., & Porfiri, M. (2013). Modulation of risk-taking behaviour in golden shiners (*Notemigonus crysoleucas*) using robotic fish. *Behavioural Processes, 100*, 9−12.

Anderson, R. P., Jimenez, G., Bae, J. Y., Silver, D., Macinko, J., & Porfiri, M. (2016). Understanding policy diffusion in the US: An information-theoretical approach to unveil connectivity structures in slowly evolving complex systems. *SIAM Journal on Applied Dynamical Systems, 15*, 1384−1409.

Aureli, M., & Porfiri, M. (2010). Coordination of self-propelled particles through external leadership. *EPL (Europhysics Letters), 92*, 40004.

Baptista, M. S., Rubinger, R. M., Viana, E. R., Sartorelli, J. C., Parlitz, U., & Grebogi, C. (2012). Mutual information rate and bounds for it. *PLoS ONE, 7*, e46745.

Barreno, M., Cardenas, A., & Tygar, J. D. (2008, December). *Optimal ROC curve for a combination of classifiers.* Paper presented at the Advances in Neural Information Processing Systems, Vancouver, British Columbia.

Bartolini, T., Butail, S., & Porfiri, M. (2015). Temperature influences sociality and activity of freshwater fish. *Environmental Biology of Fishes, 98*, 825−832.

Bartolini, T., Mwaffo, V., Showler, A., Macrì, S., Butail, S., & Porfiri, M. (2016). Zebrafish response to 3D printed shoals of conspecifics: The effect of body size. *Bioinspiration & Biomimetics, 11*, 026003.

Bollt, E. M., & Sun, J. (2014). Editorial comment on the special issue of "Information in dynamical systems and complex systems." *Entropy, 16*, 5068-5077.

Bossomaier, T., Barnett, L., Harré, M., & Lizier, J. T. (2016). *An introduction to transfer entropy: Information flow in complex systems.* Berlin: Springer.

Brand, M., Granato, M., & Nüsslein-Volhard, C. (2002). Keeping and raising zebrafish. *Zebrafish, 261*, 7−37.

Butail, S., Abaid, N., Macrì, S., & Porfiri, M. (2015). Fish–robot interactions: Robot fish in animal behavioral studies. In R. Du, Z. Li, K. Youcef-Toumi, & P. Valdivia y Alvarado (Eds.), *Robot Fish* (pp. 221−240). Berlin: Springer.

Butail, S., Bartolini, T., & Porfiri, M. (2013). Collective response of zebrafish shoals to a free-swimming robotic fish. *PLoS ONE, 8*, e76123.

Butail, S., Ladu, F., Spinello, D., & Porfiri, M. (2014). Information flow in animal-robot interactions. *Entropy, 16*, 1315−1330.

Butail, S., Mwaffo, V., & Porfiri, M. (2016). Model-free information-theoretic approach to infer leadership in pairs of zebrafish. *Physical Review E, 93*, 042411.

Chávez, M., Martinerie, J., & Le Van Quyen, M. (2003). Statistical assessment of nonlinear causality: Application to epileptic EEG signals. *Journal of Neuroscience Methods, 124*, 113−128.

Cianca, V., Bartolini, T., Porfiri, M., & Macri, S. (2013). A robotics-based behavioral paradigm to measure anxiety-related responses in zebrafish. *PLoS ONE, 8*, e69661.

Cover, T. M., & Thomas, J. A. (2012). *Elements of information theory*. Hoboken, NJ: John Wiley & Sons.

Donati, E., Worm, M., Mintchev, S., van der Wiel, M., Benelli, G., …Stefanini, C. (2016). Investigation of collective behaviour and electrocommunication in the weakly electric fish, *Mormyrus rume*, through a biomimetic robotic dummy fish. *Bioinspiration & Biomimetics, 11*, 066009.

Dooley, K., & Zon, L. I. (2000). Zebrafish: A model system for the study of human disease. *Current Opinion in Genetics & Development, 10*, 252−256.

Faes, L., Porta, A., Rossato, G., Adami, A., Tonon, D., …Nollo, G. (2013). Investigating the mechanisms of cardiovascular and cerebrovascular regulation in orthostatic syncope through an information decomposition strategy. *Autonomic Neuroscience, 178*, 76−82.

Faria, J. J., Dyer, J. R. G., Clément, R. O., Couzin, I. D., Holt, N., ...Krause, J. (2010). A novel method for investigating the collective behaviour of fish: Introducing 'Robofish.' *Behavioral Ecology and Sociobiology, 64*, 1211−1218.

Fukunaga, K. (2013). *Introduction to statistical pattern recognition*. San Diego: Academic Press.

Grabow, C., Macinko, J., Silver, D., & Porfiri, M. (2016). Detecting causality in policy diffusion processes. *Chaos: An nterdisciplinary Journal of Nonlinear Science, 26*, 083113.

Hadamard, J. (1898). Les surfaces à courbures opposées et leurs lignes géodésique. *Journal de Mathématiques Pures et Appliquées, 4*, 27−73.

Hlavackova-Schindler, K., Palus, M., Vejmelka, M., & Bhattacharya, J. (2007). Causality detection based on information-theoretic approaches in time series analysis. *Physics Reports-Review Section of Physics Letters, 441*, 1−46.

Hlinka, J., Hartman, D., Vejmelka, M., Runge, J., Marwan, N., …Paluš, M. (2013). Reliability of inference of directed climate networks using conditional mutual information. *Entropy, 15*, 2023−2045.

Hoare, D. J., Krause, J., Peuhkuri, N., & Godin, J. G. J. (2000). Body size and shoaling in fish. *Journal of Fish Biology, 57*, 1351−1366.

Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., ...Matthews, L. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature, 496*, 498−503.

Hu, F., Nie, L. J., & Fu, S. J. (2015). Information dynamics in the interaction between a prey and a predator fish. *Entropy, 17*, 7230−7241.

Kim, C., Ruberto, T., Phamduy, P., & Porfiri, M. (2018). Closed-loop control of zebrafish behaviour in three dimensions using a robotic stimulus. *Scientific Reports, 8*, 657.

King, A. J., Johnson, D. D. P., & Van Vugt, M. (2009). The origins and evolution of leadership. *Current Biology, 19*, R911−R916.

Klein, B. A., Stein, J., & Taylor, R. C. (2012). Robots in the service of animal behavior. *Communicative & Integrative Biology, 5*, 466−472.

Kraskov, A., Stögbauer, H., & Grassberger, P. (2004). Estimating mutual information. *Physical Review E, 69*, 066138.

Krause, J., Winfield, A. F. T., & Deneubourg, J. L. (2011). Interactive robots in experimental biology. *Trends in Ecology & Evolution, 26*, 369−375.

Kruusmaa, M., Rieucau, G., Montoya, J. C. C., Markna, R., & Handegard, N. O. (2016). Collective responses of a large mackerel school depend on the size and speed of a robotic fish but not on tail motion. *Bioinspiration & Biomimetics, 11*, 056020.

Kubinyi, E., Miklósi, Á., Kaplan, F., Gácsi, M., Topál, J., & Csányi, V. (2004). Social behaviour of dogs encountering AIBO, an animal-like robot in a neutral and in a feeding situation. *Behavioural Processes, 65*, 231−239.

Ladu, F., Bartolini, T., Panitz, S. G., Chiarotti, F., Butail, S., …Porfiri, M. (2015). Live predators, robots, and computer-animated images elicit differential avoidance responses in zebrafish. *Zebrafish, 12*, 205−214.

Ladu, F., Mwaffo, V., Li, J., Macri, S., & Porfiri, M. (2015). Acute caffeine administration affects zebrafish response to a robotic stimulus. *Behavioural Brain Research, 289*, 48−54.

Landgraf, T., Bierbach, D., Nguyen, H., Muggelberg, N., Romanczuk, P., & Krause, J. (2016). RoboFish: Increased acceptance of interactive robotic fish with realistic eyes and natural motion patterns by live Trinidadian guppies. *Bioinspiration & Biomimetics, 11*, 015001.

Lieschke, G. J., & Currie, P. D. (2007). Animal models of human disease: Zebrafish swim into view. *Nature Reviews Genetics, 8*, 353−367.

Lindner, M., Vicente, R., Priesemann, V., & Wibral, M. (2011). TRENTOOL: A Matlab open source toolbox to analyse information flow in time series data with transfer entropy. *BMC Neuroscience, 12*, 119.

Lizier, J. T. (2014). JIDT: An information-theoretic toolkit for studying the dynamics of complex systems. *Frontiers in Robotics and AI, 1*, 11.

Lizier, J. T., & Prokopenko, M. (2010). Differentiating information transfer and causal effect. *The European Physical Journal B-Condensed Matter and Complex Systems, 73*, 605−615.

Lord, W. M., Sun, J., Ouellette, N. T., & Bollt, E. M. (2016). Inference of causal information flow in collective animal behavior. *IEEE Transactions on Molecular, Biological and Multi-Scale Communications, 2*, 107−116.

Macrì, S., Neri, D., Ruberto, T., Mwaffo, V., Butail, S., & Porfiri, M. (2017). Three-dimensional scoring of zebrafish behavior unveils biological phenomena hidden by two-dimensional analyses. *Scientific Reports, 7*, 1962.

Marras, S., & Porfiri, M. (2012). Fish and robots swimming together: Attraction towards the robot demands biomimetic locomotion. *Journal of the Royal Society Interface, 9*, 1856−1868.

Marschinski, R., & Kantz, H. (2002). Analysing the information flow between financial time series. *The European Physical Journal B-Condensed Matter and Complex Systems, 30*, 275−281.

Maximino, C., de Brito, T. M., da Silva Batista, A. W., Herculano, A. M., Morato, S., & Gouveia, A. (2010). Measuring anxiety in zebrafish: A critical review. *Behavioural Brain Research, 214*, 157−171.

Meshalkina, D. A., Kizlyuk, M., Kisel, E., Collier, A. D., Echevarria, D. J., ...Kyzar, E. J. (2017). Zebrafish models of autism spectrum disorder. *Experimental Neurology*.

Michelsen, A., Andersen, B. B., Storm, J., Kirchner, W. H., & Lindauer, M. (1992). How honeybees perceive communication dances, studied by means of a mechanical model. *Behavioral Ecology and Sociobiology, 30*, 143−150.

Miklósi, Á., & Andrew, R. J. (2006). The zebrafish as a model for behavioral studies. *Zebrafish, 3*, 227−234.

Montalto, A., Faes, L., & Marinazzo, D. (2014). MuTE: A MATLAB toolbox to compare established and novel estimators of the multivariate transfer entropy. *PLoS ONE, 9*, e109462.

Moon, Y.-I., Rajagopalan, B., & Lall, U. (1995). Estimation of mutual information using kernel density estimators. *Physical Review E, 52*, 2318.

Moore, D. G. (2017). Retrieved from https://github.com/dglmoore/Inform

Moore, D. G., Valentini, G., Walker, S. I., & Levin, M. (2017). *Inform: A toolkit for information-theoretic analysis of complex systems.* Paper presented at the IEEE Symposium Series on Computational Intelligence, Symposium on Artificial Life. Honolulu, HI, November 27 - December 1.

Mwaffo, V., Anderson, R. P., Butail, S., & Porfiri, M. (2015). A jump persistent turning walker to model zebrafish locomotion. *Journal of The Royal Society Interface, 12*, 20140884.

Mwaffo, V., Butail, S., & Porfiri, M. (2017a). Analysis of pairwise interactions in a maximum likelihood sense to identify leaders in a group. *Frontiers in Robotics & AI, 4*, 35.

Mwaffo, V., Butail, S., & Porfiri, M. (2017b). In-silico experiments of zebrafish behaviour: Modeling swimming in three dimensions. *Scientific Reports, 7*, 39877.

Nakayama, S., Ruiz Marín, M., Camacho, M., & Porfiri, M. (2017). Plasticity in leader–follower roles in human teams. *Scientific Reports, 7*, 14562.

Neri, D., Ruberto, T., Cord-Cruz, G., & Porfiri, M. (2017). Information theory and robotics meet to study predator-prey interactions. *Chaos: An Interdisciplinary Journal of Nonlinear Science, 27*, 073111.

Neri, D., Ruberto, T., Mwaffo, V., Bartolini, T., & Porfiri, M. (in press). Social environment modulates anxiogenic effects of caffeine in zebrafish, *Behavioural Pharmacology*.

Orange, N., & Abaid, N. (2015). A transfer entropy analysis of leader-follower interactions in flying bats. *The European Physical Journal Special Topics, 224*, 3279−3293.

Orger, M. B., & de Polavieja, G. G. (2017). Zebrafish behavior: Opportunities and challenges. *Annual Review of Neuroscience*, 40, 125-147.

Paluš, M., & Vejmelka, M. (2007). Directionality of coupling from bivariate time series: How to avoid false causalities and missed connections. *Physical Review E, 75*, 056211.

Panula, P., Chen, Y. C., Priyadarshini, M., Kudo, H., Semenova, S., …Sallinen, V. (2010). The comparative neuroanatomy and neurochemistry of zebrafish CNS systems of relevance to human neuropsychiatric diseases. *Neurobiology of Disease, 40*, 46−57.

Patricelli, G. L., Uy, J. A. C., Walsh, G., & Borgia, G. (2002). Sexual selection: Male displays adjusted to female's response. *Nature, 415*, 279−280.

Paulus, M. P., Geyer, M. A., Gold, L. H., & Mandell, A. J. (1990). Application of entropy measures derived from the ergodic theory of dynamical systems to rat locomotor behavior. *Proceedings of the National Academy of Sciences, 87*, 723−727.

Phamduy, P., Polverino, G., Fuller, R. C., & Porfiri, M. (2014). Fish and robot dancing together: Bluefin killifish females respond differently to the courtship of a robot with varying color morphs. *Bioinspiration & Biomimetics, 9*, 036021.

Polverino, G., Phamduy, P., & Porfiri, M. (2013). Fish and robots swimming together in a water tunnel: Robot color and tail-beat frequency influence fish behavior. *PLoS ONE, 8*, e77589.

Polverino, G., & Porfiri, M. (2013). Mosquitofish (*Gambusia affinis*) responds differentially to a robotic fish of varying swimming depth and aspect ratio. *Behavioural Brain Research, 250*, 133−138.

Porfiri, M. (in press). Information flow in a Boolean network model of collective behavior. *IEEE Transactions on Control of Network Systems*.

Porfiri, M., & Ruiz Marín, M. (2017a). Information flow in a model of policy diffusion: An analytical study. *IEEE Transactions on Network Science and Engineering*, 5, 42-54.

Porfiri, M., & Ruiz Marín, M. (2017b). Symbolic dynamics of animal interaction. *Journal of Theoretical Biology, 435*, 145−156.

Porfiri, M., & Ruiz Marín, M. (2018). Information flow in a model of policy diffusion: An analytical study. *IEEE Transactions on Network Science and Engineering, 5*, 42−54.

Postlethwait, J. H., Woods, I. G., Ngo-Hazelett, P., Yan, Y. L., Kelly, P. D.,...Talbot, W. S. (2000). Zebrafish comparative genomics and the origins of vertebrate chromosomes. *Genome Research, 10*, 1890−1902.

Romano, D., Benelli, G., Donati, E., Remorini, D., Canale, A., & Stefanini, C. (2017). Multiple cues produced by a robotic fish modulate aggressive behaviour in Siamese fighting fishes. *Scientific Reports, 7*, 4667.

Roulston, M. S. (1999). Estimating the errors on measured entropy and mutual information. *Physica D: Nonlinear Phenomena, 125*, 285−294.

Ruberto, T., Mwaffo, V., Singh, S., Neri, D., & Porfiri, M. (2016). Zebrafish response to a robotic replica in three dimensions. *Royal Society Open Science, 3*, 160505.

Ruberto, T., Polverino, G., & Porfiri, M. (2017). How different is a 3D-printed replica from a conspecific in the eyes of a zebrafish? *Journal of the Experimental Analysis of Behavior, 107*, 279-293.

Ruddell, B. L., & Kumar, P. (2009a). Ecohydrologic process networks: 1. Identification. *Water Resources Research, 45,* W03419.

Ruddell, B. L., & Kumar, P. (2009b). Ecohydrologic process networks: 2. Analysis and characterization. *Water Resources Research, 45,* W03420.

Ruddell, B. L., Oberg, N., Garcia, M., & Kumar, P. (2010). Using information-theoretic statistics in MATLAB to understand how ecosystems affect regional climates. Retrieved from MathWorks.com

Runge, J., Heitzig, J., Petoukhov, V., & Kurths, J. (2012). Escaping the curse of dimensionality in estimating multivariate transfer entropy. *Physical Review Letters, 108*, 258701.

Sandoval, L. (2014). Structure of a global network of financial companies based on transfer entropy. *Entropy, 16*, 4443−4482.

Schreiber, T. (2000). Measuring information transfer. *Physical Review Letters, 85*, 461.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27*, 623-656.

Silverman, B. W. (1986). *Density estimation for statistics and data analysis* (Vol. 26). Boca Raton, FL: CRC Press.

Smirnov, D. A. (2013). Spurious causalities with transfer entropy. *Physical Review E, 87*, 042917.

Spinello, C., Macrì, S., & Porfiri, M. (2013). Acute ethanol administration affects zebrafish preference for a biologically inspired robot. *Alcohol, 47*, 391−398.

Staniek, M., & Lehnertz, K. (2008). Symbolic transfer entropy. *Physical Review Letters, 100*, 158101.

Stetter, O., Battaglia, D., Soriano, J., & Geisel, T. (2012). Model-free reconstruction of excitatory neuronal connectivity from calcium imaging signals. *PLoS Computational Biology, 8*, e1002653.

Stewart, A. M., Braubach, O., Spitsbergen, J., Gerlai, R., & Kalueff, A. (2014). Zebrafish models for translational neuroscience research: From tank to bedside. *Trends in Neurosciences, 37*, 264−278.

Sun, J., & Bollt, E. M. (2014). Causation entropy identifies indirect influences, dominance of neighbors and anticipatory couplings. *Physica D: Nonlinear Phenomena, 267*, 49−57.

Sun, J., Taylor, D., & Bollt, E. M. (2015). Causal network inference by optimal causation entropy. *SIAM Journal on Applied Dynamical Systems, 14*(1), 73-106.

Swain, D. T., Couzin, I. D., & Leonard, N. E. (2012). Real-time feedback-controlled robotic fish for behavioral experiments with fish schools. *Proceedings of the IEEE, 100*, 150−163.

Takanishi, A., Aoki, T., Ito, M., Ohkawa, Y., & Yamaguchi, J. (1998, October). *Interaction between creature and robot: Development of an experiment system for rat and rat robot interaction.* Paper presented at the Intelligent Robots and Systems, 1998, Proceedings, Victoria, British Columbia.

Tinbergen, N. (1948). Social releasers and the experimental method required for their study. *The Wilson Bulletin*, 60, 6−51.

Tomaru, T., Murakami, H., Niizato, T., Nishiyama, Y., Sonoda, K., …Gunji, Y.-P. (2016). Information transfer in a swarm of soldier crabs. *Artificial Life and Robotics, 21*, 177−180.

Ver Steeg, G., & Galstyan, A. (2012). *Information transfer in social media.* Proceedings of the 21st international conference on World Wide Web (pp. 509-518). New York: ACM Digital Library.

Verdes, P. F. (2005). Assessing causality from multivariate time series. *Physical Review E, 72*, 026222.

Vicente, R., Wibral, M., Lindner, M., & Pipa, G. (2011). Transfer entropy--a model-free measure of effective connectivity for the neurosciences. *Journal of Computational Neuroscience, 30*, 45−67.

Wang, X. R., Miller, J. M., Lizier, J. T., Prokopenko, M., & Rossi, L. F. (2012). Quantifying and tracing information cascades in swarms. *PLoS ONE, 7*, e40084.

Ward, A. J. W., Sumpter, D. J. T., Couzin, I. D., Hart, P. J. B., & Krause, J. (2008). Quorum decision-making facilitates information transfer in fish shoals. *Proceedings of the National Academy of Sciences, 105*, 6948−6953.

Wibral, M., Pampu, N., Priesemann, V., Siebenhühner, F., Seiwert, H., ...Vicente, R. (2013). Measuring information-transfer delays. *PLoS ONE, 8*, e55809.

Wibral, M., Vicente, R., & Lizier, J. T. (2014). *Directed information measures in neuroscience*. Berlin: Springer-Verlag.

Wiener, N. (1956). The theory of prediction. *Modern Mathematics for Engineers, 1*, 125−139.

Worm, M., Kirschbaum, F., & von der Emde, G. (2017). Social interactions between live and artificial weakly electric fish: Electrocommunication and locomotor behavior of *Mormyrus rume proboscirostris* towards a mobile dummy fish. *PLoS ONE, 12*, e0184622.

Zhang, Z., & Grabchak, M. (2013). Bias adjustment for a nonparametric entropy estimator. *Entropy, 15*, 1999−2011.

Zhang, P., Rosen, M., Peterson, S. D., & Porfiri, M. (in press). An information-theoretic approach to study fluid-structure interactions, *Journal of Fluid Mechanics*.

Zienkiewicz, A., Barton, D. A. W., Porfiri, M., & di Bernardo, M. (2015a). Data-driven stochastic modelling of zebrafish locomotion. *Journal of Mathematical Biology, 71*, 1081−1105.

Zienkiewicz, A., Barton, D. A. W., Porfiri, M., & Di Bernardo, M. (2015b). Leadership emergence in a data-driven model of zebrafish shoals with speed modulation. *The European Physical Journal Special Topics, 224*, 3343−3360.

Zienkiewicz, A. K., Ladu, F., Barton, D. A. W., Porfiri, M., & Di Bernardo, M. (2018). Data-driven modelling of social forces and collective behaviour in zebrafish. *Journal of Theoretical Biology, 443*, 39−51.