

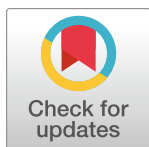
RESEARCH ARTICLE

Advancing statistical treatment of photolocomotor behavioral response study data

Natalie Mastin¹, Luke Durell¹, Bryan W. Brooks^{2,3}, Amanda S. Hering^{1*}

1 Department of Statistical Science, Baylor University, Waco, TX, United States of America, **2** Department of Environmental Science, Baylor University, Waco, TX, United States of America, **3** Institute of Biomedical Studies, Baylor University, Waco, TX, United States of America

* mandy_hering@baylor.edu



OPEN ACCESS

Citation: Mastin N, Durell L, Brooks BW, Hering AS (2024) Advancing statistical treatment of photolocomotor behavioral response study data. PLoS ONE 19(5): e0300636. <https://doi.org/10.1371/journal.pone.0300636>

Editor: Robyn L Tanguay, Oregon State University, UNITED STATES

Received: November 14, 2023

Accepted: March 2, 2024

Published: May 21, 2024

Copyright: © 2024 Mastin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Mastin, Natalie; Durell, Luke; Brooks, Bryan W.; Hering, Amanda S., 2023, "Replication Data for: Advancing Statistical Treatment of Photolocomotor Behavioral Response Study Data", <https://doi.org/10.7910/DVN/2XEL5H>, Harvard Dataverse, V1.

Funding: ASH: National Science Foundation Harnessing the Data Revolution: Data Science Corps project #1924146, <https://www.nsf.gov> BWB: National Science Foundation award #CHE-1339637 and National Institute of Environmental Health Sciences of the National Institutes of Health

Abstract

Fish photolocomotor behavioral response (PBR) studies have become increasingly prevalent in pharmacological and toxicological research to assess the environmental impact of various chemicals. There is a need for a standard, reliable statistical method to analyze PBR data. The most common method currently used, univariate analysis of variance (ANOVA), does not account for temporal dependence in observations and leads to incomplete or unreliable conclusions. Repeated measures ANOVA, another commonly used method, has drawbacks in its interpretability for PBR study data. Because each observation is collected continuously over time, we instead consider each observation to be a function and apply functional ANOVA (FANOVA) to PBR data. Using the functional approach not only accounts for temporal dependency but also retains the full structure of the data and allows for straightforward interpretation in any subregion of the domain. Unlike the traditional univariate and repeated measures ANOVA, the FANOVA that we propose is nonparametric, requiring minimal assumptions. We demonstrate the disadvantages of univariate and repeated measures ANOVA using simulated data and show how they are overcome by applying FANOVA. We then apply one-way FANOVA to zebrafish data from a PBR study and discuss how those results can be reproduced for future PBR studies.

1 Introduction

In photolocomotor behavioral response (PBR) studies, researchers utilize known brightness preference/avoidance in fish in order to compare the behavior of treatment groups in alternating light/dark conditions over time. This is also often called the "light/dark transition test." PBR studies have become increasingly prevalent over the past two decades in pharmacological and toxicological research due to their useful indication of the behavioral effect of chemical compounds and potential underlying mode of action. In *Google Scholar*, when the keywords "fish AND locomotor AND behavioral response" are searched, over 90,000 works are returned. When the same keywords are searched in both *PubMed* and *Scopus*, over 15,000

award #1P01ES028942, <https://www.nih.gov>

These sponsors did not play any role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

works are represented. When the keywords “fish AND light/dark AND transition” are searched across all three search engines, over 39,000 works are listed. The widespread presence of these studies reflects the fundamental importance of behavioral effects when examining biological activities of chemicals. In fact, the zebrafish model is increasingly used in basic biomedical and environmental research and translational applications because of societal animal welfare concerns and due to governmental mandates that are moving away from use of rodent models. Behavioral changes reflect physiological changes and can impact population-level development, thus acting as an effective tool for assessing the environmental impact of chemical contaminants [1] and behavioral phenotypes during drug discovery [2].

1.1 Standard statistical methods used in PBR studies

The standard model used to analyze the behavioral effects of chemical concentrations in the PBR literature is univariate analysis of variance (ANOVA). Responses, such as swim speed, are averaged for each observation either over the entire domain or a section of the domain, and the averages are then used in subsequent analysis. The most common strategy is to calculate the average of the response variable across the entire domain, regardless of light/dark period [3–5]. It is also common for univariate ANOVA to be repeated twice, once for the measurements in the light periods and once for those in the dark periods [6–8]. Alternatively, it may be repeated several times for each individual light/dark period itself [9–11]. When assumptions for univariate one-way ANOVA are not met, Fisher’s Exact test [12] or the nonparametric Kruskal-Wallis test [4, 13, 14] are often used. Other approaches include repeated measures ANOVA [15, 16]; multiple *t*-tests for each level of the factor against the control group [17]; *t*-tests for every time point [18]; and a Kolmogorov-Smirnov test comparing the area under the curve (AUC) of differential entropy between treatment and control groups [14, 19].

We surveyed over fifty papers in the PBR literature, and while these are not necessarily representative of the entire population of PBR studies, even in this small sample, many different statistical methods to analyze PBR data are represented. These papers were selected based on the expertise of our colleagues who use them for reference in their work. Fig 1 shows ten different categories of methods being used to analyze PBR data in these papers with the majority being a type of one-way ANOVA in which measurements are averaged for each fish. This inconsistency makes comparing study results challenging, if not impossible. Furthermore, [20] mention that there are no available publications comparing the different statistical methods for behavioral studies of zebrafish and recognizes that handling time series data requires models with greater complexity. Over 80% of the papers that we surveyed did not account for the temporal dependency of PBR data, losing useful information about the observations.

Because the data in PBR studies are recorded over time, each individual fish’s observations are dependent. By averaging across the domain and applying univariate ANOVA, the variability within each individual is not accounted for, thereby underestimating the error in the model. This could result in misleading conclusions, as we demonstrate in this paper. Because repeated measures ANOVA models time as a factor, it accounts for individual variability across the domain and is a more favorable way to handle PBR data. However, it is difficult to interpret the interaction between treatment groups and time. When the interaction is significant, post hoc tests must be performed at every time point to determine where in time the differences in treatment levels occur. Most studies that use repeated measures ANOVA test the average differences of the treatment levels in each light/dark period rather than the differences at each time point [21, 22]. Because the raw data are not consistently made publicly available for each PBR study, it is difficult to assess the effects of the statistical methods used on the conclusions drawn. In spite of the fact that many statistical methods are being used in PBR

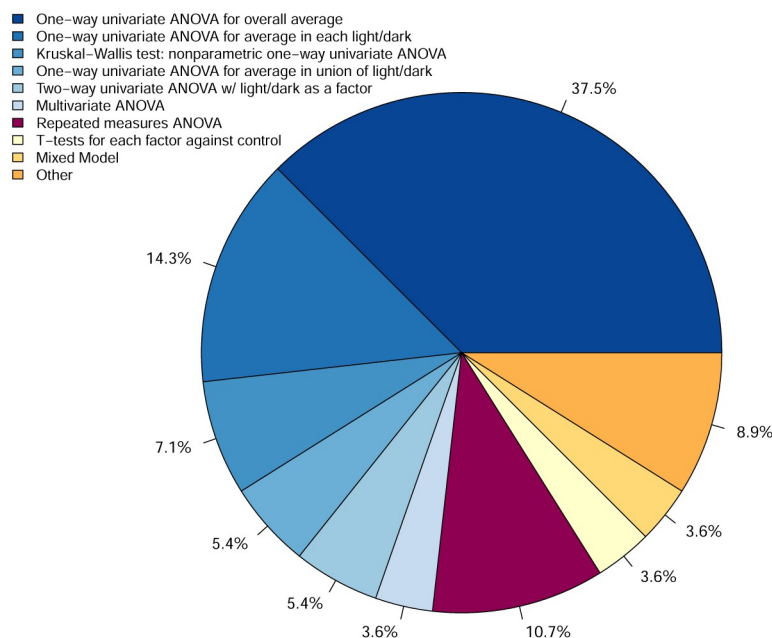


Fig 1. Pie chart representing survey of methods used to analyze PBR study data. Methods represented in shades of blue are variations of traditional ANOVA.

<https://doi.org/10.1371/journal.pone.0300636.g001>

literature, none of the standard methods involve treating the observations as continuous functions, as we propose in this paper. This may occur because once a methodology has been established, it gains momentum, and many researchers continue to use similar methods. Furthermore, functional data analysis (FDA) itself is still a developing branch of statistics, so we hope to encourage researchers to adopt this new paradigm.

1.2 The functional approach

FDA is a branch of statistical science that deals with data observed over a continuous domain, such as time, and can be resampled as smooth functions [23, 24]. FDA assumes that the data are smooth, naturally accounting for the dependence across the domain. Data in PBR studies are functional in their structure, making FDA the proper approach to use in these studies. Because there are an infinite number of ways that functions can differ and yet still share the same mean, information is lost by averaging across the domain and applying univariate ANOVA. For example, Fig 2 shows eight different functions that share the same mean over their domain, and their shapes clearly differ. If univariate ANOVA were to be performed on the means of these functions, no difference would be found. Similarly, the AUC used in [14, 19] is a summary value and suffers from the same problem in that equal values of AUC can arise from two functions with very different patterns. However, functional ANOVA (FANOVA) would detect a difference in the functions. Results obtained using the functional approach, unlike the univariate approach, are both valid and straightforward to interpret.

In this paper, we present a framework for implementing one-way FANOVA that has minimal assumptions and demonstrate how it can be implemented in PBR study datasets. In Section 2, we review the methodology for univariate ANOVA and repeated measures ANOVA and present functional ANOVA. For each method, we define its models, hypotheses, tests, and assumptions. The FANOVA that we recommend uses a nonparametric permutation test for

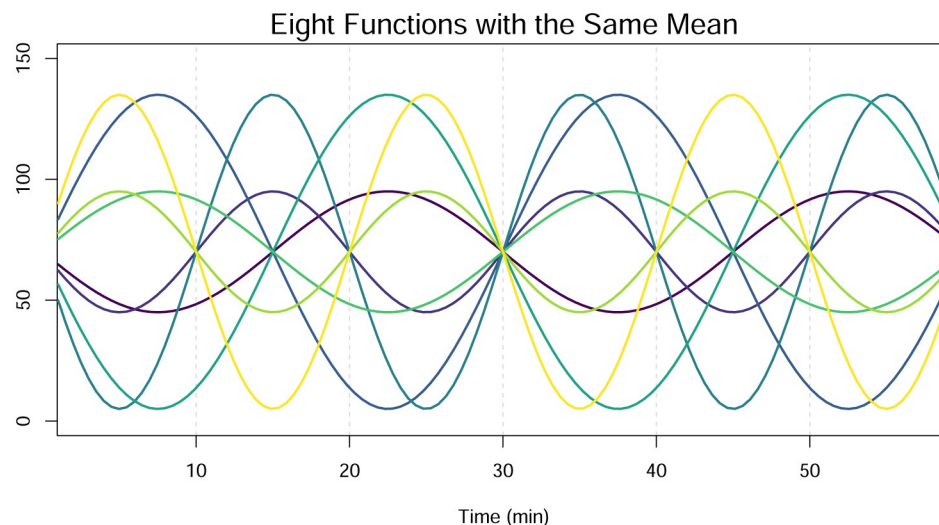


Fig 2. Eight functions that share an overall mean of 70 across their domain.

<https://doi.org/10.1371/journal.pone.0300636.g002>

an F -based test statistic and bootstrapped pointwise confidence intervals for the estimated functional effects and contrasts. In Section 3, we illustrate the most extreme possible consequences of applying univariate or repeated measures ANOVA to PBR study data in a simulated example, and we compare these results with one-way FANOVA. Section 4 details a real example using PBR study data and shows how the functional representation of the data provides a seamless analysis of the startle response when brightness conditions change by constructing acceleration functions. Section 5 presents concluding remarks.

2 Methodology

In this section, we describe the methodology for traditional univariate ANOVA, repeated measures ANOVA, and FANOVA. To streamline the presentation, we present the balanced one-way case with equal sample sizes in each treatment group.

2.1 Univariate ANOVA

In this section, we review the univariate ANOVA model, hypotheses, F -statistic, and assumptions. Traditional analysis of variance became widely known after being presented in the classic text by Fisher [25].

2.1.1 Univariate ANOVA model. The response, y_{ij} , is the observed value of subject j at level i of factor A . The one-way univariate ANOVA model is

$$y_{ij} = \mu + \alpha_i + \epsilon_{ij}, \quad (1)$$

where μ is the overall mean response across the domain; α_i is the treatment effect of factor A with levels $i = 1, \dots, a$; and ϵ_{ij} is the subject effect for $j = 1, \dots, n_i$ replicates. In the balanced case, $n_i = n$.

2.1.2 Univariate ANOVA hypotheses. The following set of hypotheses is used to test the equality of the treatment effect of factor A :

$$\begin{aligned} H_0 &: \alpha_1 = \dots = \alpha_a = 0, \\ \text{vs. } H_1 &: \alpha_i \neq 0, \text{ for some } i. \end{aligned} \quad (2)$$

If H_0 is rejected, then we conclude that there is evidence to support the presence of at least one treatment effect, and post hoc tests are performed on the following hypotheses to determine which pairs of treatment effects differ:

$$\begin{aligned} H_{0,\alpha_i-\alpha_{i'}} &: \alpha_i - \alpha_{i'} = 0, \text{ for } i \neq i', \\ \text{vs. } H_{1,\alpha_i-\alpha_{i'}} &: \alpha_i - \alpha_{i'} \neq 0, \text{ for } i \neq i'. \end{aligned} \quad (3)$$

2.1.3 Univariate F -Test. The parameters in the model are estimated as follows:

$$\begin{aligned} \hat{\mu} &= \bar{y}_{..}, \text{ and} \\ \hat{\alpha}_i &= \bar{y}_{i.} - \bar{y}_{..}, \end{aligned}$$

where

$$\begin{aligned} \bar{y}_{..} &= \frac{1}{an} \sum_{i=1}^a \sum_{j=1}^n y_{ij}, \text{ and} \\ \bar{y}_{i.} &= \frac{1}{n} \sum_{j=1}^n y_{ij}. \end{aligned}$$

Then, the sums of squares due to the hypothesis (SSA) and error (SSE) are calculated as

$$SSA = n \sum_{i=1}^a [\bar{y}_{i.} - \bar{y}_{..}]^2 = n \sum_{i=1}^a \hat{\alpha}_i^2, \text{ and} \quad (4)$$

$$SSE = \sum_{i=1}^a \sum_{j=1}^n [\bar{y}_{ij} - \bar{y}_{i.}]^2 \quad (5)$$

The univariate F -statistic is

$$F^{obs} = \frac{SSA/q_A}{SSE/q_E}, \quad (6)$$

and the degrees of freedom (df) associated with SSA and SSE are $q_A = a - 1$ and $q_E = a(n - 1)$, respectively.

To obtain the critical value for the rejection region, F^{crit} , the $(1 - \tau)^{\text{th}}$ percentile of the F_{q_A, q_E} distribution for significance level τ is calculated. If $F^{obs} > F^{crit}$, then reject H_0 , and post hoc analyses are performed. Otherwise, we fail to reject H_0 .

2.1.4 Univariate ANOVA assumptions. The following assumptions are required to perform univariate ANOVA:

1. The response variable for each factor level is normally distributed.
2. The distributions of the response variable for each level have the same variance.
3. Each observation is independent of the others.

Averaging the PBR values of an individual across the domain and then applying univariate ANOVA to them ignores the fact that values of each individual are dependent as well as the variability of an individual over the course of the experiment. Therefore, the results may be misleading and should consider biological relevance about the differences in the individuals as a whole. Violation of the independence assumption is considered to have the highest impact of all of these assumptions [26].

2.2 Repeated measures ANOVA

In this section, a review of repeated measures ANOVA model, hypotheses, and F -statistics is given. This exposition follows largely from [27].

2.2.1 Repeated measures ANOVA model. The response, y_{ijk} , represents the response at time k from the j th subject in level i of factor A . The repeated measures ANOVA model is

$$y_{ijk} = \mu + \alpha_i + \beta_k + (\alpha\beta)_{ik} + \pi_{j(i)} + \epsilon_{ijk}, \quad (7)$$

where μ is the overall mean response across the domain, and α_i is the fixed effect of factor A with levels $i = 1, \dots, a$. Additionally, β_k is the fixed effect of the time point with levels $k = 1, \dots, t$ with $\sum_{k=1}^t \beta_k = 0$, and finally, $(\alpha\beta)_{ik}$ is the fixed effect of the interaction between the i th level and the k th time point, with constraints

$$\sum_{i=1}^a (\alpha\beta)_{ik} = \sum_{k=1}^t (\alpha\beta)_{ik} = 0.$$

The parameters $\pi_{j(i)}$ are random effects for the j th subject at the i th level with $j = 1, \dots, n$. The $\pi_{j(i)}$ are assumed to be independent and normally distributed with mean zero and variance σ_π^2 . Finally, the ϵ_{ijk} are independent random error terms with $\epsilon_{ijk} \sim N(0, \sigma_\epsilon^2)$.

2.2.2 Repeated measures ANOVA hypotheses. The following series of hypotheses is used to test the equality of the treatment effect of factor A for all of the time points. The first hypothesis tests the significance of the interaction between the factor and time,

$$\begin{aligned} H_{0AT} : & \quad (\alpha\beta)_{11} = \dots = (\alpha\beta)_{at} = 0, \\ \text{vs. } H_{1AT} : & \quad (\alpha\beta)_{ik} \neq 0, \text{ for some } ik. \end{aligned} \quad (8)$$

If H_{0AT} is rejected, then there is evidence to support the presence of at least one interaction effect between factor A and time, and the cell means, $\eta_{ik} = \mu + \alpha_i + \beta_k + (\alpha\beta)_{ik}$, are then tested for equality based on the following post hoc hypothesis:

$$\begin{aligned} H_{0, \eta_{ik} - \eta_{(ik)'}} : & \quad \eta_{ik} - \eta_{(ik)'} = 0, \text{ for } ik \neq (ik)', \\ \text{vs. } H_{1, \eta_{ik} - \eta_{(ik)'}} : & \quad \eta_{ik} - \eta_{(ik)'} \neq 0, \text{ for } ik \neq (ik)'. \end{aligned} \quad (9)$$

Otherwise, failure to reject H_{0AT} indicates that there is no significant evidence that the interaction terms differ from zero, leading to tests of the main effects based on the following two hypotheses:

$$\begin{aligned} H_{0A} : & \quad \alpha_1 = \dots = \alpha_a = 0, \\ \text{vs. } H_{1A} : & \quad \alpha_i \neq 0, \text{ for some } i, \end{aligned} \quad (10)$$

and

$$\begin{aligned} H_{0T} : & \quad \beta_1 = \dots = \beta_t = 0, \\ \text{vs. } H_{1T} : & \quad \beta_k \neq 0, \text{ for some } k. \end{aligned} \quad (11)$$

If H_{0A} or H_{0T} are rejected, then the following hypotheses are tested:

$$\begin{aligned} H_{0, \alpha_i - \alpha_{i'}} : & \quad \alpha_i - \alpha_{i'} = 0, \text{ for } i \neq i', \\ \text{vs. } H_{1, \alpha_i - \alpha_{i'}} : & \quad \alpha_i - \alpha_{i'} \neq 0, \text{ for } i \neq i', \end{aligned} \quad (12)$$

and

$$\begin{aligned} H_{0,\beta_j-\beta_{k'}} &: \beta_k - \beta_{k'} = 0, \text{ for } k \neq k', \\ \text{vs. } H_{1,\beta_j-\beta_{k'}} &: \beta_k - \beta_{k'} \neq 0, \text{ for } k \neq k'. \end{aligned} \quad (13)$$

When there are a large number of repeated measurements, performing post hoc tests for each individual time point or interaction between factor A and time becomes impractical and unwieldy.

2.2.3 Repeated measures F -Test. First, the effects are estimated as,

$$\begin{aligned} \hat{\mu} &= \bar{y}_{...}, \\ \hat{\alpha}_i &= \bar{y}_{i..} - \bar{y}_{...}, \\ \hat{\beta}_k &= \bar{y}_{..k} - \bar{y}_{...}, \\ (\hat{\alpha}\hat{\beta})_{ik} &= \bar{y}_{i.k} - \bar{y}_{i..} - \bar{y}_{..k} + \bar{y}_{...}, \text{ and} \\ \hat{\pi}_{j(i)} &= \bar{y}_{ij.} - \bar{y}_{i..}, \end{aligned}$$

where

$$\begin{aligned} \bar{y}_{...} &= \frac{1}{ant} \sum_{i=1}^a \sum_{j=1}^n \sum_{k=1}^t y_{ijk}, \\ \bar{y}_{i..} &= \frac{1}{nt} \sum_{j=1}^n \sum_{k=1}^t y_{ijk}, \\ \bar{y}_{..k} &= \frac{1}{an} \sum_{i=1}^a \sum_{j=1}^n y_{ijk}, \\ \bar{y}_{i.k} &= \frac{1}{n} \sum_{j=1}^n y_{ijk}, \text{ and} \\ \bar{y}_{ij.} &= \frac{1}{t} \sum_{k=1}^t y_{ijk}. \end{aligned}$$

Table 1 displays the sum of squares and degrees of freedom for each source of variation. The sums of squares are calculated as

$$SSA = t \sum_{i=1}^a n [\bar{y}_{i..} - \bar{y}_{...}]^2 = t \sum_{i=1}^a n \hat{\alpha}_i^2, \quad (14)$$

$$SSS(A) = t \sum_{i=1}^a \sum_{j=1}^n [\bar{y}_{ij.} - \bar{y}_{i..}]^2 = t \sum_{i=1}^a \sum_{j=1}^n \hat{\pi}_{j(i)}^2, \quad (15)$$

Table 1. Sums of squares and degrees of freedom for repeated measures ANOVA.

Source	DF	Sum Sq
Factor A	$a - 1$	SSA
Subjects(Factor A)	$a(n - 1)$	SSS(A)
Time	$t - 1$	SST
Factor A×Time	$(a - 1)(t - 1)$	SSAT
Residual	$a(n - 1)(t - 1)$	SSE

<https://doi.org/10.1371/journal.pone.0300636.t001>

$$SST = an \sum_{k=1}^t [\bar{y}_{..k} - \bar{y}_{...}]^2 = an \sum_{k=1}^t \hat{\beta}_k^2, \quad (16)$$

$$SSAT = \sum_{i=1}^a \sum_{j=1}^n \sum_{k=1}^t [\bar{y}_{ijk} - \bar{y}_{i..} - \bar{y}_{.jk} + \bar{y}_{...}]^2 = \sum_{i=1}^a \sum_{j=1}^n \sum_{k=1}^t (\hat{\alpha}\hat{\beta})_{ik}^2, \text{ and} \quad (17)$$

$$SSE = \sum_{i=1}^a \sum_{j=1}^n \sum_{k=1}^t [y_{ijk} - \bar{y}_{ijk} - \bar{y}_{ij.} + \bar{y}_{i..}]^2. \quad (18)$$

Eqs 14, 16 and 17 are the sums of squares needed to test the hypotheses in Eqs 8, 10 and 11, respectively, while Eq 18 is the SSE. The F -statistic for testing for differences among groups is given by

$$F_A = \frac{MSA}{MSSA} = \frac{SSA/(a-1)}{SSSA/a(n-1)} \quad (19)$$

with $a-1$ and $a(n-1)$ numerator and denominator df, respectively. This test requires the assumption that the within-group covariance matrices are equal. The F -statistic for testing differences among time points is given by

$$F_T = \frac{MST}{MSE} = \frac{SST/(t-1)}{SSE/[a(n-1)(t-1)]} \quad (20)$$

with $t-1$ and $a(n-1)(t-1)$ numerator and denominator df, respectively. Similarly, the F -statistic for testing the significance of the factor $A \times$ time interaction is given by

$$F_{AT} = \frac{MSAT}{MSE} = \frac{SSAT/[(a-1)(t-1)]}{SSE/[a(n-1)(t-1)]} \quad (21)$$

with $(a-1)(t-1)$ numerator and $a(n-1)(t-1)$ denominator df. Both of the prior tests require the assumption that the within-group covariance matrices are equal and that the variances of the differences between variables are equal.

2.2.4 Repeated measures ANOVA assumptions. The following assumptions are required to perform repeated measures ANOVA:

1. The response variable for each factor level is normally distributed.
2. All *ant* observations are independent.
3. The within-group covariance matrices are equal.
4. The variances of the pairwise differences between all combinations of groups are equal; i.e., $Var(y_{ijk} - y_{i'jk})$ is constant for all $i \neq i'$, and $Var(y_{ijk} - y_{ijk'})$ is constant for all $k \neq k'$. This is usually referred to as the sphericity assumption.

Because time is treated as a factor in repeated measures ANOVA, it is a valid way to analyze PBR study data, but there are several disadvantages to using this approach. It is difficult to interpret the interaction between groups and time. The assumptions of normality, within-group homogeneity of variance, and sphericity are often violated and can bias results. For example, [28] showed that if the sphericity assumption is violated, then the true type I error rate is positively biased, causing excessive false rejection of the null hypotheses for the

hypotheses in Eqs 8 and 11. Additionally, [29] notes that even repeated measures ANOVA does not account for temporal autocorrelation between measurements, which is likely present in PBR data, leading to lower power. This limitation is greatest when adjacent time points have very similar responses, which is the case for most of the adjacent time points in PBR study data. Finally, due to the number of repeated measurements over time in PBR data, post hoc analyses require many tests. Functional ANOVA, however, does not have these disadvantages. The functional ANOVA that we present makes minimal assumptions and provides simple interpretations and post hoc analyses.

2.3 FANOVA

In this section, we present the one-way FANOVA model, hypotheses, and F -based statistics. One-way FANOVA is described in [24, 30]. We also present a nonparametric approach to testing these hypotheses that relies on very few assumptions.

2.3.1 FANOVA model. The functional response, $y_{ij}(t)$, is defined over a domain $\mathcal{T}$, with i corresponding to the levels of factor A and j corresponding to the subject. In this case, $t \in \mathcal{T}$, where $\mathcal{T}$ represents an interval of time, and t represents a specific point in time. The one-way FANOVA model is

$$y_{ij}(t) = \mu(t) + \alpha_i(t) + \epsilon_{ij}(t), \quad t \in \mathcal{T}, \quad (22)$$

where $\mu(t)$ is the overall functional mean; $\alpha_i(t)$ is the functional effect of factor A with levels $i = 1, \dots, a$; and $\epsilon_{ij}(t)$ is the functional subject effect for $j = 1, \dots, n$ replicates.

2.3.2 FANOVA hypotheses. The following hypotheses test the equality of the functional effects.

$$\begin{aligned} H_0 &: \alpha_1(t) = \dots = \alpha_a(t) = 0, \text{ for all } t \in \mathcal{T}, \\ \text{vs. } H_1 &: \alpha_i(t) \neq 0, \text{ for some } t \in \mathcal{T} \text{ and level } i. \end{aligned} \quad (23)$$

If H_0 is rejected, then we can conclude that there is evidence of at least one functional effect, and post hoc tests are performed for the following hypotheses to determine which pairs of functional effects differ:

$$\begin{aligned} H_{0, \alpha_i - \alpha_{i'}} &: \alpha_i(t) - \alpha_{i'}(t) = 0, \text{ for all } t \in \mathcal{T} \text{ with } i \neq i', \\ \text{vs. } H_{1, \alpha_i - \alpha_{i'}} &: \alpha_i(t) - \alpha_{i'}(t) \neq 0, \text{ for some } t \in \mathcal{T} \text{ with } i \neq i'. \end{aligned} \quad (24)$$

2.3.3 Smoothing the functions. Before applying FANOVA to the data, the first step is to create functions of the discretized values by taking a weighted sum of basis functions. The `fda` package in R [31] has many helpful functions for smoothing, and we present a summary of this process [24]. Basis functions are sets of mathematically independent functions, and by taking linear combinations of them, they can be used to approximate any function. The degree to which the data are smoothed rather than interpolated is determined by the number of basis functions. As the number of basis functions increases, the function can be more closely approximated, but this introduces a possibility of overfitting to noise. Interpolation is achieved when the number of basis functions is equal to the number of discrete values recorded in the domain, but this is only desirable when the measurements are errorless. Once the functions have been fit, they can be evaluated across a fine grid of values over the domain.

The first choice to be made in smoothing is which basis set to use. Many basis systems are described in [24]. If possible, the set of basis functions chosen should have features that match those of the functions being approximated. For example, if the data have periodic features, a

Fourier basis can achieve a good approximation using fewer basis functions. Fourier bases are most useful when the data have no strong local features, and there is no need to reflect discontinuities in their derivatives. For data with no clear structure, cubic b-splines are the most common choice of approximation system, and they can be engineered to have discontinuity in the functions.

Less important than the choice of basis set is how many basis functions should be used in the approximation. However, the modern way to smooth functions is to use the maximum number of basis functions, or to “saturate” the model, allowing the number of basis functions to match the number of measured values for each observation. This would produce functions that are interpolated and thereby very rough, so a penalty is imposed that controls the trade-off between fit to the data and smoothness. The smoothness of the function is controlled by a smoothing parameter that penalizes its roughness using the square of the second derivative of the estimated functions. The smoothing parameter is chosen by minimizing generalized cross validation of the squared error between the smooth and discretized values. By minimizing the squared error, unusually high values are allowed to be influential, but their impact on the overall analysis is dampened.

The frequency of measurements necessary to construct a continuous function may vary based on the variability of the data. The data should be collected frequently enough to capture important changes in the process at the rate at which they occur. For example, if rapid changes are expected to occur at a temporal scale of less than a minute, data should be collected more frequently than a minute. In order to create a smooth function that represents the true data, features of interest must be captured through the discretized measurements.

2.3.4 Functional F -Statistic. Given the smoothed functions, we then estimate the functional effects as

$$\begin{aligned}\hat{\mu}(t) &= \bar{y}_{..}(t), \text{ and} \\ \hat{\alpha}_i(t) &= \bar{y}_i(t) - \bar{y}_{..}(t),\end{aligned}$$

where

$$\begin{aligned}\bar{y}_{..}(t) &= \frac{1}{an} \sum_{i=1}^a \sum_{j=1}^n y_{ij}(t), \text{ and} \\ \bar{y}_i(t) &= \frac{1}{n} \sum_{j=1}^n y_{ij}(t).\end{aligned}$$

Then, the functional sums of squares due to the hypothesis ($SSA(t)$) and error ($SSE(t)$) are calculated as

$$SSA(t) = n \sum_{i=1}^a [\bar{y}_i(t) - \bar{y}_{..}(t)]^2 = n \sum_{i=1}^a \hat{\alpha}_i^2(t), \text{ and} \quad (25)$$

$$SSE(t) = \sum_{i=1}^a \sum_{j=1}^n [\bar{y}_{ij}(t) - \bar{y}_i(t)]^2. \quad (26)$$

The general functional F -statistic is

$$F^{obs}(t) = \frac{SSA(t)/q_A}{SSE(t)/q_E}, \quad t \in \mathcal{T}, \quad (27)$$

where the degrees of freedom associated with $SSA(t)$ and $SSE(t)$ are $q_A = a - 1$ and $q_E = a(n$

– 1), respectively. These quantities are the functional versions of the scalar counterparts in one-way univariate ANOVA.

2.3.5 F_{FT} statistic and permutation test. The functional F_{FT} test statistic [32] summarizes the variation across levels over the domain through integration, providing a scalar F -value and corresponding p -value. This overall statistic is

$$F_{FT}^{obs} = \frac{\int_T SSA(t)dt/q_A}{\int_T SSE(t)dt/q_E}. \quad (28)$$

To approximate the distribution of F_{FT} under the null hypothesis in Eq 23, the functions are permuted (reassigned to a random level of Factor A) a large number of times, and F_{FT} is recalculated for each permutation of the functions. The values of F_{FT} based on the permuted samples are denoted F_{FT}^r for $r = 1, \dots, R$. Then, for a given significance level, τ , the $100 \times (1 - \tau)^{th}$ percentile of the F_{FT}^r values is calculated. If the observed statistic, F_{FT}^{obs} , exceeds the critical value, F_{FT}^{crit} , then we reject H_0 . Otherwise, we fail to reject H_0 . A scalar p -value can be estimated by calculating the proportion of F_{FT}^r permutations that exceed F_{FT}^{obs} .

Additional F -based statistics have been proposed to summarize the general $F(t)$ statistic, such as the L^2 -norm [33], globalizing- F [34], and F -max statistics [35]. Any of these F -based statistics could be used for this method, but as shown in [36], the F_{FT} test is very close to the nominal size and is slightly below $\tau = 0.05$, resulting in fewer rejections of a true null hypothesis. Furthermore, it has high power under a variety of alternative hypotheses. Because of the fundamental differences in the hypotheses between univariate ANOVA, repeated measures ANOVA, and FANOVA, the power of the three methods cannot be compared to one another; however, the power of the F_{FT} test compared to the other F -based FANOVA tests is explored further in [36]. The only assumption required for the FANOVA method described here is that the individuals are independent of each other.

There are some challenges with implementing FANOVA. Smoothing the curves reduces variability in the data, which underestimates the true error in the FANOVA model, but this underestimation can be reduced by fitting the smooth curves to minimize the error between them and the discretized values, as described in Section 2.3.3. Furthermore, smoothing reduces the impact of outliers in the data, but even with data-driven choices in place, such as using cross-validation to choose the smoothing parameter or choice of basis sets, some subjectivity in smoothing is unavoidable. Thus, a researcher should investigate the effects of smoothing on their conclusions. Secondly, knowledge of a programming language is currently required to implement the nonparametric FANOVA that we present. However, the code that we developed for this analysis is publicly available on the Harvard Dataverse repository [37], and many modifications can be made with a low level of effort to adapt the code to different PBR studies.

3 Illustration

To illustrate how traditional ANOVA can mask important features in PBR data, we develop a one-way functional ANOVA example using artificial data for three levels of a factor over a period of sixty minutes. The factor could correspond to a chemical toxicant with three concentration levels, and the response could be the swim speed.

3.1 Simulating the data

In this illustration, we simulate 20 observations under each level by adding random variation to the amplitude, period, and phase of sine curves. For $j = 1, \dots, 20$ and $t \in [1, 60]$, let $f_{i,j}(t)$ represent the functional observation of the i^{th} fish observed under level $i = 1, 2, 3$ evaluated at

minute t . All of the observations can be represented as

$$f_{ij}(t) = 70 + A_{ij} \sin\left(\frac{\pi t}{R_{ij}} + \pi P_{ij}\right) + W_i, \quad (29)$$

where A_{ij} represents variation in the amplitude; R_{ij} represents variation in the period; P_{ij} represents variation in the phase; and $W_i \sim \text{iid } N(0, \sigma^2 = 36)$ represents random individual variation. The distributions for the amplitude, period and phase of each level are shown in Table 2.

For each level, the true average swim pattern can be represented as

$$\bar{f}_i(t) = 70 + \bar{A}_i \sin\left(\frac{\pi t}{\bar{R}_i} + \pi \bar{P}_i\right), \quad (30)$$

where $\bar{A}_i$, $\bar{R}_i$, and $\bar{P}_i$ are the averages of the distributions in Table 2 and are given in Table 3. Note that the averages of the mean functions taken across the domain are equal, or

$$\int_T \bar{f}_1(t) dt = \int_T \bar{f}_2(t) dt = \int_T \bar{f}_3(t) dt = 70.$$

The left column of Fig 3 shows twenty functions simulated from Eq 29 for each level. All sets of functions are smoothed prior to analysis with a Fourier basis set with ten basis functions. The Fourier basis system is chosen because the simulated data are periodic, and the number of basis functions is selected to balance interpolation and smoothness of the data. The results of the Fourier smoothing are shown in the right column of Fig 3. For each function, the mean across the domain is calculated, and Fig 4 shows a boxplot of these values. The boxplots have very similar means for each of the three levels even though the function shapes at each level are clearly different.

3.2 Univariate and repeated measures ANOVA

If we are interested in determining if there is a difference in the swim speeds of the fish observed under the three levels, then it would be problematic to employ univariate ANOVA because reducing each function to its average does not account for the variability in swim speed over time. In this particular case, the true mean for each of the functions in Fig 3 is 70,

Table 2. Distributions for random variation in amplitude, period, and phase for each level of the factor.

Function	Amplitude	Period	Phase
$f_{1,j}(t)$	$A_{1,j} \sim \text{iid Unif}(37, 43)$	$R_{1,j} \sim \text{iid } N(15, 0.025)$	$P_{1,j} \sim \text{iid } N(0, 0.0225)$
$f_{2,j}(t)$	$A_{2,j} \sim \text{iid Unif}(17, 23)$	$R_{2,j} \sim \text{iid } N(15, 0.025)$	$P_{2,j} \sim \text{iid } N(1, 0.0225)$
$f_{3,j}(t)$	$A_{3,j} \sim \text{iid Unif}(57, 63)$	$R_{3,j} \sim \text{iid } N(10, 0.025)$	$P_{3,j} \sim \text{iid } N(1, 0.0225)$

<https://doi.org/10.1371/journal.pone.0300636.t002>

Table 3. Average amplitude, period, and phase for the mean swim pattern functions.

Mean Function	Amplitude	Period	Phase
$\bar{f}_1(t)$	$\bar{A}_1 = 40$	$\bar{R}_1 = 15$	$\bar{P}_1 = 0$
$\bar{f}_2(t)$	$\bar{A}_2 = 20$	$\bar{R}_2 = 15$	$\bar{P}_2 = 1$
$\bar{f}_3(t)$	$\bar{A}_3 = 60$	$\bar{R}_3 = 10$	$\bar{P}_3 = 1$

<https://doi.org/10.1371/journal.pone.0300636.t003>

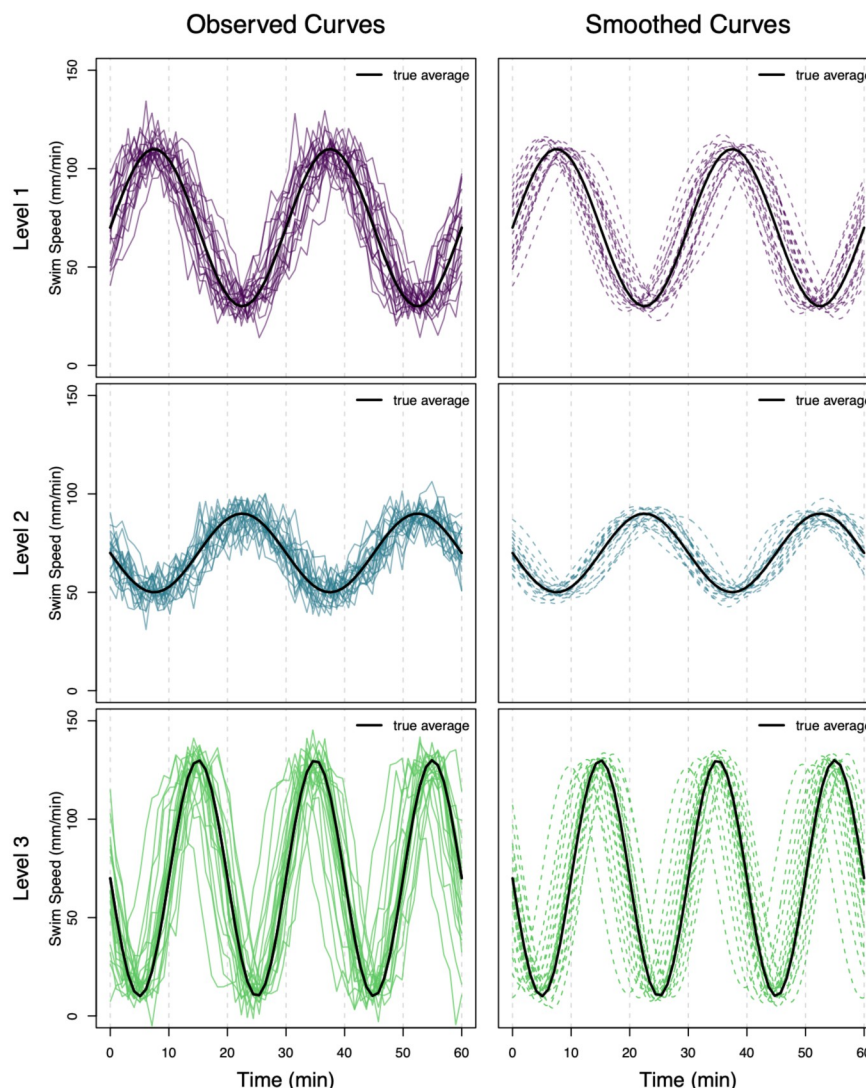


Fig 3. The left side shows swim speed over time of twenty fish simulated under level 1 of the factor (top), level 2 of the factor (middle), and level 3 of the factor (bottom), with the true average function overlaid in black. These are treated as the observed data. The right side shows the corresponding smoothed functions of swim speed.

<https://doi.org/10.1371/journal.pone.0300636.g003>

so the null hypothesis in the univariate ANOVA is true. Therefore, the correct decision is to fail to reject the null hypothesis. Indeed, when we apply a univariate ANOVA to the means of the functions in Fig 4, a large p -value is produced, as shown in Table 4, indicating that there is no significant difference in the mean swim speeds of the fish under the three conditions. However, when viewed over the domain, the fish observed under the three levels have very different swim patterns.

Next, we perform a repeated measures ANOVA, and the results are shown in Table 5. We find that the interaction between the factor and time is significant, which is an improvement upon the univariate ANOVA that did not detect any difference between the three levels. Due to the violation the normality and sphericity assumptions, however, these results are inconclusive. Usually when assumptions of a test are not met, a nonparametric method is used, but no

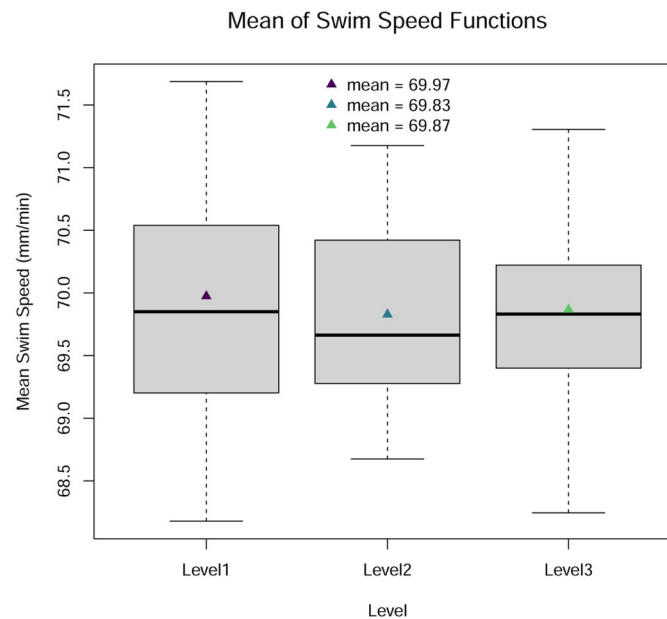


Fig 4. Boxplots of the averages of the functions simulated under three levels of a factor.

<https://doi.org/10.1371/journal.pone.0300636.g004>

such test can be performed in this context. Ignoring the violated assumptions, the significant interaction between the factor and time implies that a post hoc analysis should be performed. In this case, there would be $3 \times 60 = 180$ post hoc comparisons, and this is a prohibitive number of comparisons to make, so we do not present the post hoc analyses for the repeated measures ANOVA.

3.3 FANOVA application

Finally, we apply functional ANOVA to this simulated data. Using 10,000 permutations and a significance level of $\tau = 0.05$, the permuted critical value is $F_{FT}^{crit} = 2.945$, which is much smaller than the observed value of $F_{FT}^{obs} = 77.468$. The p -value for F_{FT}^{obs} is zero, meaning that 100% of the permuted test statistic values were less than F_{FT}^{obs} . Fig 5 shows the $F^{obs}(t)$ statistic from the one-way functional ANOVA results on the smoothed data. The $F^{obs}(t)$ statistic is particularly high in regions where the functional means differ the most. FANOVA not only captures the significant difference in the mean swim pattern functions, but it also identifies the approximate locations in the domain where the differences occur.

If the overall function mean and differences from this overall mean for each level are of interest, then the estimated functional effects can provide insight, as shown in Fig 6. The top panel of Fig 6 shows the overall functional mean effect, $\hat{\mu}(t)$, and the bottom panel shows the functional effect of each level of factor A, $\hat{\alpha}_i(t)$. To form 95% pointwise bootstrapped

Table 4. Univariate ANOVA results.

	Df	Sum Sq	Mean Sq	F-value	Pr(>F)
Factor	2	0.23	0.11	0.17	0.8475
Residual	57	38.85	0.68		

<https://doi.org/10.1371/journal.pone.0300636.t004>

Table 5. Repeated measures ANOVA results.

	Df	Sum Sq	Mean Sq	F-value	Pr(>F)
Factor	2	13.57	6.79	0.17	0.8475
Fish(Factor)	57	2331.24	40.90		
Time	59	574781.29	9742.06	38.79	<0.0001
Factor×Time	118	1980732.70	16785.87	66.84	<0.0001
Residual	3363	844573.03	251.14		

<https://doi.org/10.1371/journal.pone.0300636.t005>

confidence intervals for each of the estimated functional effects, we sample with replacement 1000 times from the sixty estimated functional subject effects,

$$\hat{v}_{ij}(t) = y_{ij}(t) - \hat{\mu}(t) - \hat{\alpha}_i(t), \quad i = 1, \dots, 3, \quad j = 1, \dots, 20,$$

resulting in $\hat{v}_{ij}^r(t)$ for $r = 1, \dots, 1000$. These are used to construct bootstrapped observations as

$$y_{ij}^r(t) = \hat{v}_{ij}^r(t) + \hat{\mu}(t) + \hat{\alpha}_i(t), \quad i = 1, \dots, 3, \quad j = 1, \dots, 20. \quad (31)$$

The functional effects, $\hat{\mu}^r(t)$ and $\hat{\alpha}_i^r(t)$, are reestimated based on the bootstrap samples in Eq 31. The difference between the bootstrapped and observed functional effects are calculated, and the 2.5% and 97.5% pointwise percentiles of the 1000 difference functions at each t are taken to be the upper and lower bounds of the confidence interval. These confidence intervals can only be interpreted at a point, but they provide a helpful visual reference for the variability of the estimated functional effects. The computational time to construct the permutation test and bootstrap intervals is minimal, requiring less than a minute on a personal laptop computer.

3.3.1 Post hoc analysis. Once the null hypothesis is rejected, we perform post hoc analyses to determine which pairs of levels differ significantly from one another. The tests employ a

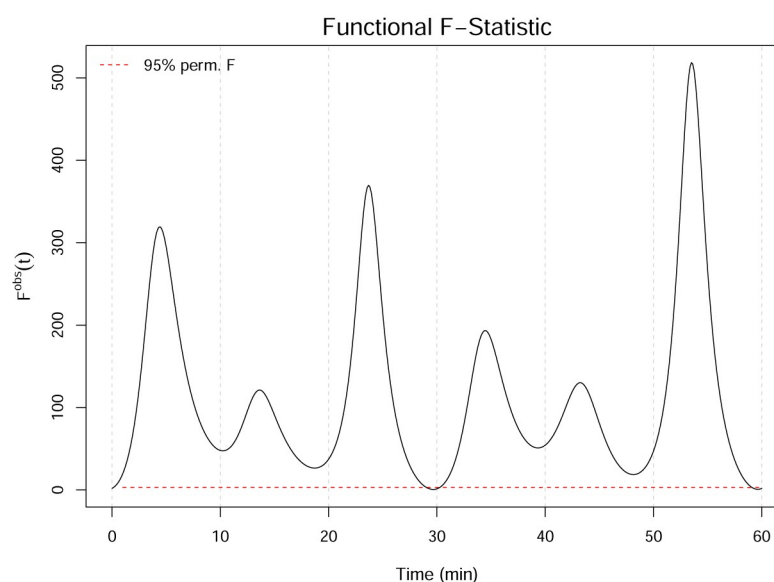


Fig 5. One-way FANOVA $F^{obs}(t)$ statistic based on the smoothed dataset. The red dashed line indicates the 95% pointwise percentile calculated from the permutations.

<https://doi.org/10.1371/journal.pone.0300636.g005>

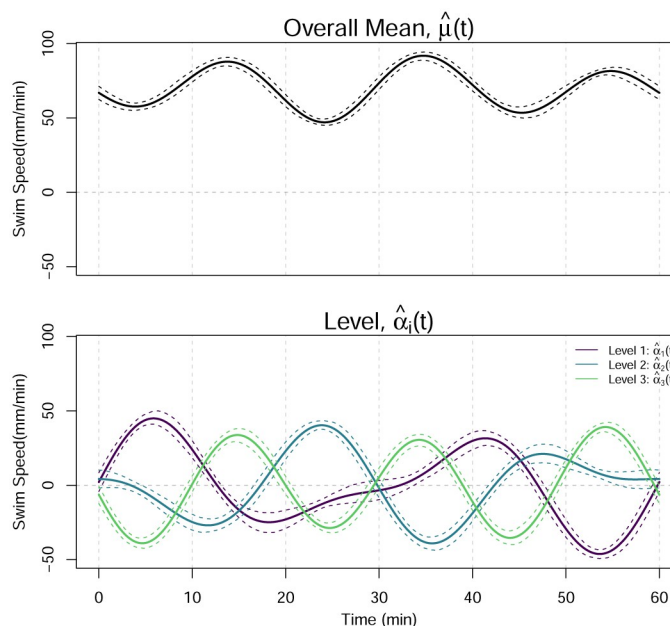


Fig 6. Estimated overall mean effect (top) and factor effect (bottom). Dashed lines are 95% pointwise confidence intervals formed using 1000 bootstrap replicates.

<https://doi.org/10.1371/journal.pone.0300636.g006>

Bonferroni correction so that the significance level for each test is the nominal significance level ($\tau = 0.05$) divided by the number of tests. With three levels, there are three pairwise comparisons formed by calculating $\hat{\alpha}_i(t) - \hat{\alpha}_{i'}(t)$ with $i \neq i'$, so the adjusted level of significance is $0.05/3 = 0.017$. The 95% pointwise bootstrapped confidence intervals are constructed as described in Section 3.3 with 1000 replicates and are shown in Fig 7. There is a significant difference between the mean swim speeds of each pair of levels.

4 Real data example

In this example, we demonstrate one-way FANOVA on a dataset from a PBR study that compares the effects of an environmental toxicant, 1-heptanol, of different concentrations on the swim patterns of zebrafish over alternating lightness conditions [8]. Experimental results examined here were obtained from studies that were approved by the Institutional Animal Care and Use Committee at Baylor University. Swim speed is recorded for a total of fifty minutes, divided into a ten-minute dark acclimation period followed by four alternating light and dark periods, each lasting ten minutes. A measurement is recorded every minute.

The zebrafish are divided into five groups of twenty four and exposed to concentrations of 1-heptanol at either 0%, 5%, 10%, 20%, and 40% of their median lethal concentration (LC_{50}). Fig 8 shows the observed swim speed of the 120 zebrafish, grouped by the concentration level. It is clear that the zebrafish are more active in the dark periods than the light periods. As the concentration of 1-heptanol increases, the zebrafish also appear to swim less overall.

For each individual fish, the mean swim speed across the light and dark periods is taken (excluding the acclimation period) and is plotted in Fig 9 by concentration. Each concentration level has a clearly different mean. This implies that a one-way univariate ANOVA is likely to be significant, but provides no information about where the differences occur in the domain.

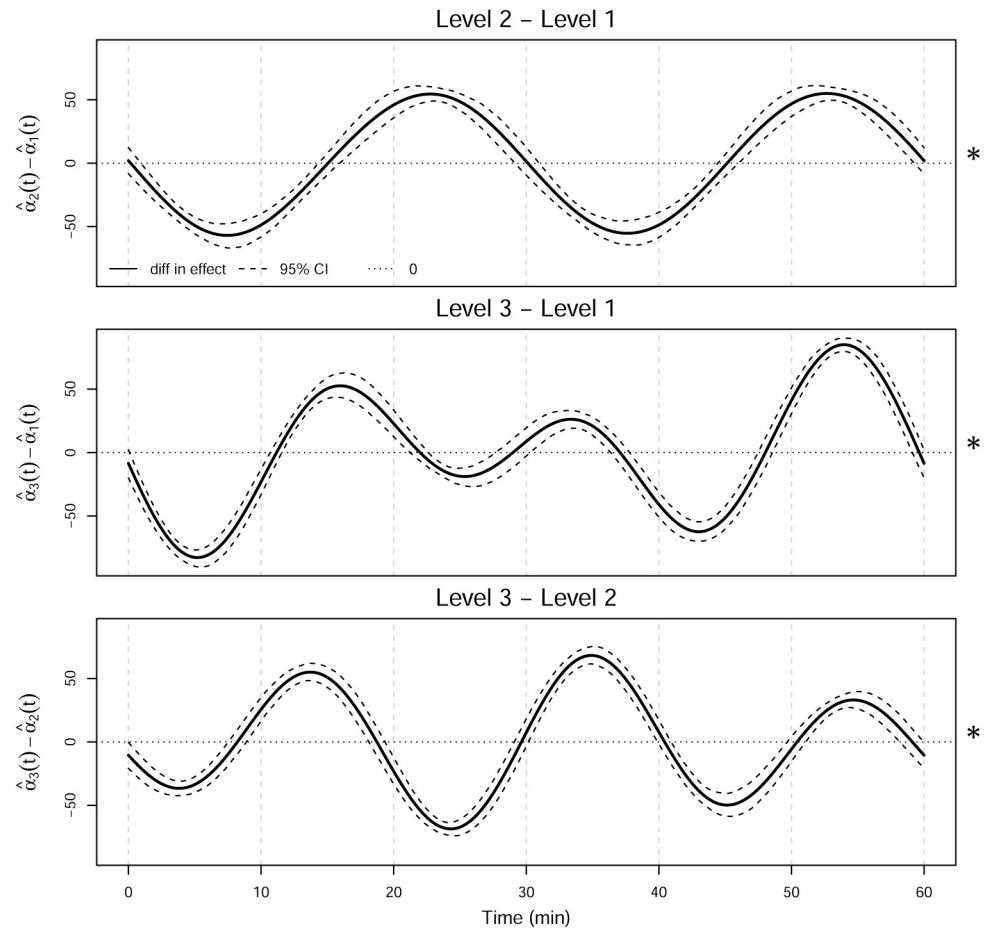


Fig 7. Functional post hoc contrasts for the levels. The dashed line is the pointwise 98.33% confidence interval using 1000 bootstrap replicates. The asterisk on the right indicates significance of the F_{FT} test for the factor effect.

<https://doi.org/10.1371/journal.pone.0300636.g007>

4.1 Univariate and repeated measures ANOVA

The result of a univariate ANOVA applied to the means of each individual over the domain is shown in Table 6. As expected, the univariate ANOVA results in a low p -value, indicating a significant difference in the mean swim speed due to the concentration levels. Because the null hypothesis is rejected, the proper procedure is to perform a post hoc analysis to determine which pairs of concentration levels differ significantly, using the hypotheses in Eq 3. However, the post hoc analysis does not provide any additional information about where in the domain the differences occur.

Similarly, we apply repeated measures ANOVA to the data. Because all of the assumptions for repeated measures ANOVA are violated, any results obtained from the analysis are inconclusive, but Table 7 provides the results for illustration. All of the terms in the model, including the interaction, are significant. Because the assumptions are violated, pairwise comparisons for the interaction between concentration and time would be invalid. Furthermore, with forty repeated measurements, over 750 pairwise comparisons would be needed, and this is a prohibitive number to evaluate. Therefore, we do not include any post hoc analyses for the repeated measures ANOVA.

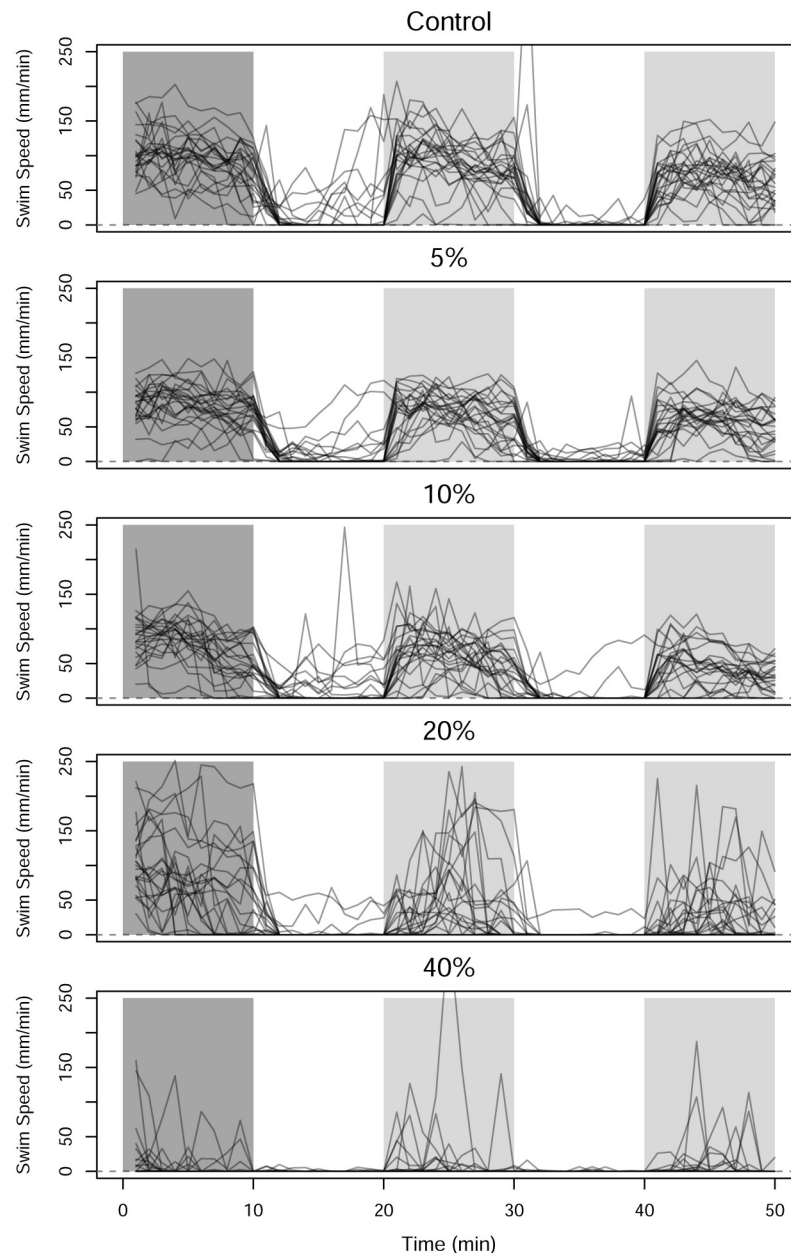


Fig 8. Pointwise observations of swim speed by concentration of 1-heptanol. The color of the background denotes the acclimation period (dark gray) and the alternating light (white) and dark (light gray) ten-minute periods.

<https://doi.org/10.1371/journal.pone.0300636.g008>

4.2 FANOVA application

Before applying FANOVA to the data, each function is smoothed using a cubic b-spline basis with sixty basis functions. The cubic b-spline basis set is selected because the data do not have a recognizable pattern. When the light turns on/off, the zebrafish display a rapid and immediate change in their swim pattern, so we build discontinuities into the smoothed functions at these times by allowing additional flexibility at these points. The basis functions are saturated, penalization is applied to the square of the second derivative, and the smoothing parameter is

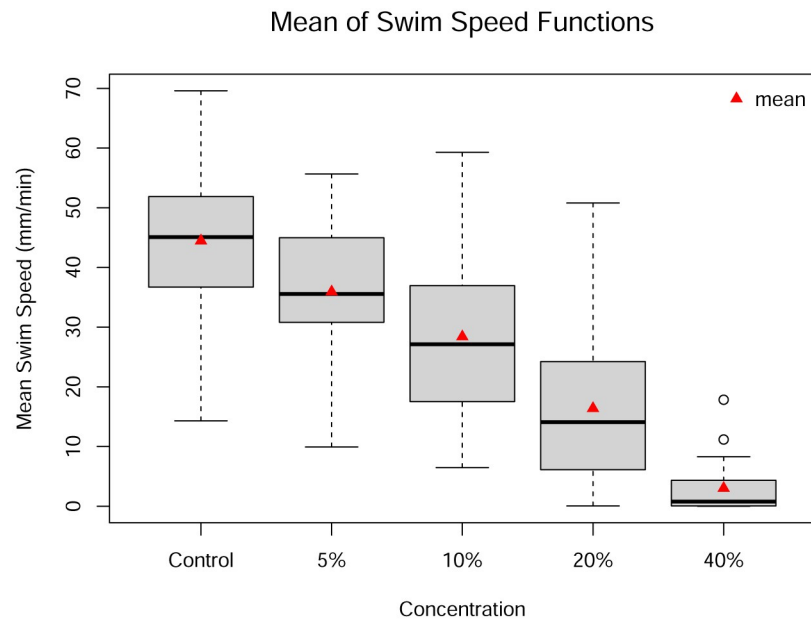


Fig 9. Boxplots of the averages of the functions for each concentration. Circles outside of the fences of the boxplots represent potential outliers.

<https://doi.org/10.1371/journal.pone.0300636.g009>

obtained through generalized cross validation to minimize the squared error between the smooth curves and discretized values. The smoothing results in the reduction of some large values that occur in the data, but the smooth curves also reduce the noise of the data. Another option would be to remove the functions with unusual values, and several methods for identification of functional outliers are available but are not explored here [38, 39]. Once the smoothing parameter is chosen, the smooth functions are then evaluated across a fine, evenly-spaced grid, and negative values are replaced with zeroes so that the swim speed functions remain non-negative across the domain. The results of the b-spline smoothing are shown in the center panel of Fig 10. To visualize the average swim speed for each concentration level, the average

Table 6. Univariate ANOVA results.

	Df	Sum Sq	Mean Sq	F-value	Pr(>F)
Concentration	4	25577.23	6394.31	48.51	<0.0001
Residual	115	15157.14	131.80		

<https://doi.org/10.1371/journal.pone.0300636.t006>

Table 7. Repeated measures ANOVA results.

	Df	Sum Sq	Mean Sq	F-value	Pr(>F)
Concentration	4	1023089.35	255772.34	48.51	<0.0001
Fish(Concentration)	115	606285.45	5272.05		
Time	39	2052231.27	52621.31	82.38	<0.0001
Concentration×Time	156	714188.17	4578.13	7.17	<0.0001
Residual	4485	2865001.94	638.80		

<https://doi.org/10.1371/journal.pone.0300636.t007>

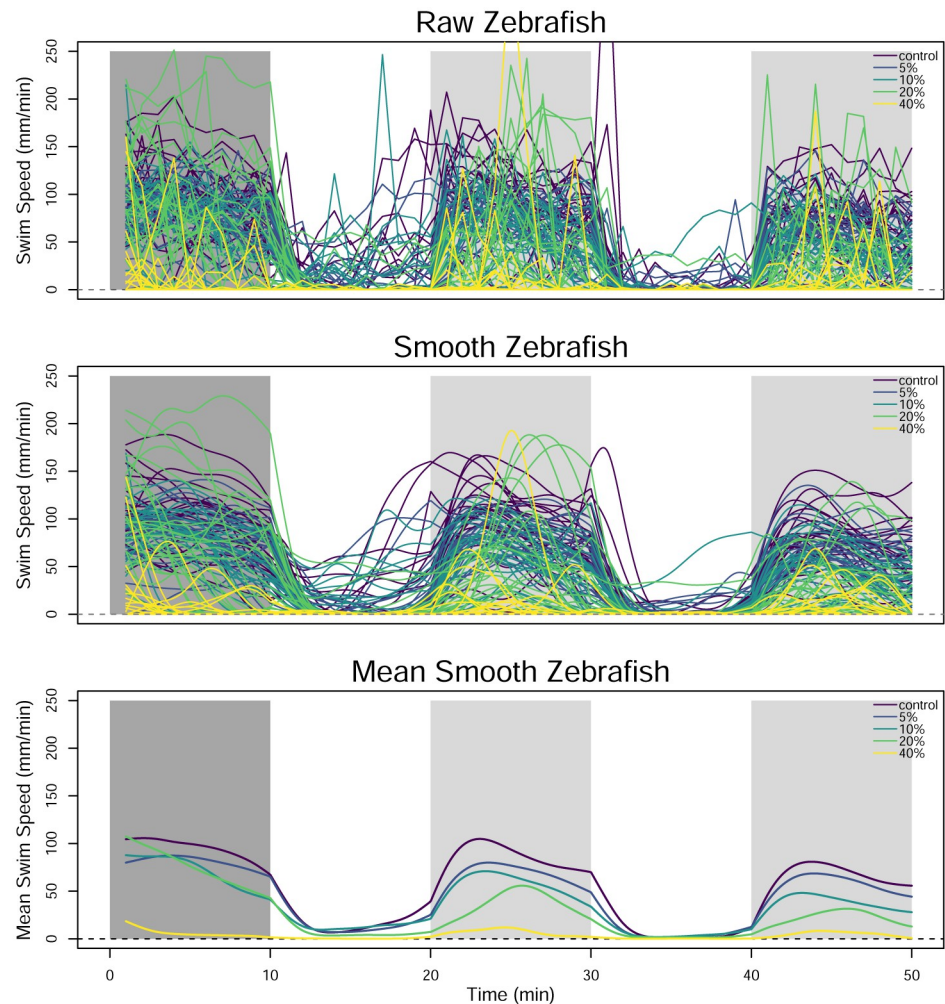


Fig 10. Raw observations of swim speed (top), smoothed observations of swim speed (center), and the functional averages of the smoothed swim speed curves for each concentration level (bottom).

<https://doi.org/10.1371/journal.pone.0300636.g010>

of the smooth swim speed curves for each concentration level is plotted in the bottom panel of Fig 10. In the two dark periods in particular, the average swim speed decreases as the concentration of 1-heptanol increases. In the light periods, there is only a small difference between the concentration levels.

One-way FANOVA is applied to the smoothed curves plotted in Fig 10, and the parameters in the model of Eq 22 are estimated. Fig 11 displays the estimated functional effects and 95% bootstrapped pointwise confidence intervals formed using the process described in Section 3.3. The overall mean effect illustrates the overall lightness aversion of the zebrafish, which is also clear in Figs 8 and 10. The concentration effects illustrate that higher concentrations result in lower swim speeds, particularly in the dark regions. There is less difference in the concentration effect during the light periods. Note that because we do not include the acclimation period in our analysis of the data, it is not included in this figure.

The functional F -statistic for the hypotheses given in Eq 23 is computed according to Eq 27. Fig 12 shows the $F^{obs}(t)$ statistic from the one-way FANOVA results on the smoothed data. The $F^{obs}(t)$ statistic is particularly high in the dark regions, where the functions differ the most.

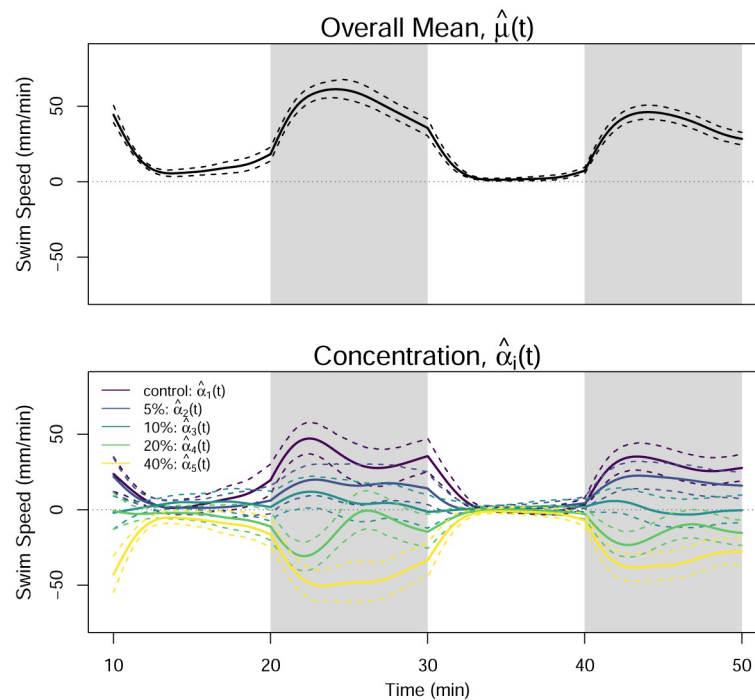


Fig 11. Estimated overall mean effect (top) and concentration effect (bottom). Dashed lines are 95% pointwise confidence intervals formed using 1000 bootstrap replicates.

<https://doi.org/10.1371/journal.pone.0300636.g011>

We then perform the permutation test for the F_{FT} statistic described in Section 2.3.5 using 10,000 permutations. Fig 13 shows a histogram of the 10,000 permuted F_{FT} statistics, illustrating its sampling distribution under the null hypothesis. Using a significance level of $\tau = 0.05$, the permuted critical value of F_{FT} is $F_{FT}^{crit} = 2.14$, which is much smaller than the observed

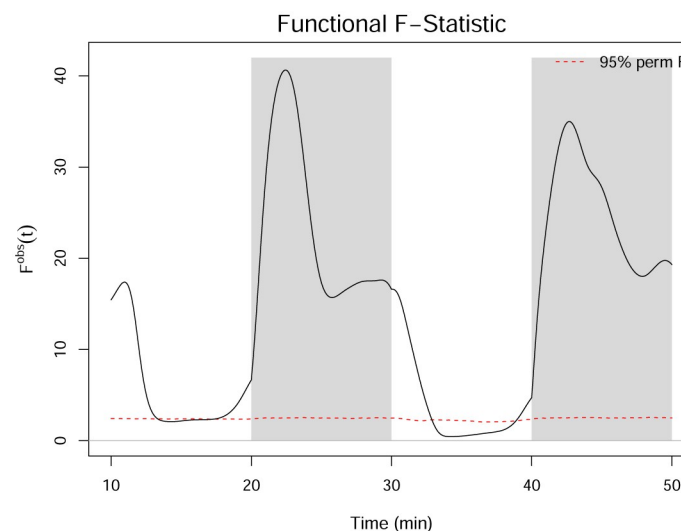


Fig 12. One-way FANOVA $F^{obs}(t)$ statistic applied to the smoothed dataset. The red dashed line indicates the 95% pointwise percentile calculated from the permutations.

<https://doi.org/10.1371/journal.pone.0300636.g012>

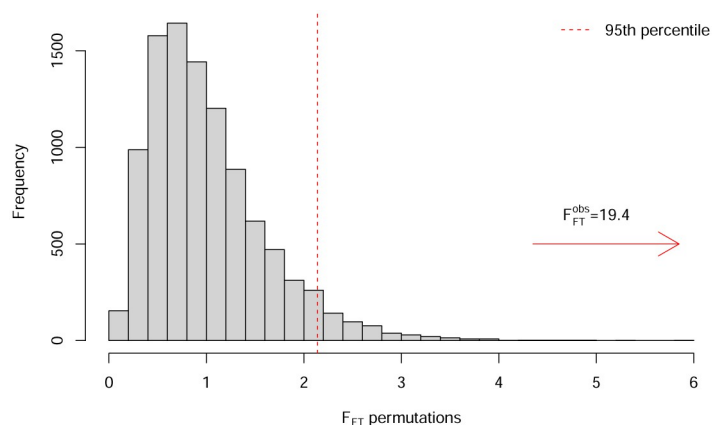


Fig 13. Histogram of 10,000 F_{FT} permutations. The red dashed line indicates the 95th percentile of the permutations.

<https://doi.org/10.1371/journal.pone.0300636.g013>

$F_{FT}^{obs} = 19.39$. The p -value for the F_{FT} test is zero, indicating that 100% of the permuted statistics are below the critical value. In addition to concluding that significant differences in the mean functions exist across concentrations, FANOVA also identifies approximately where in the domain the differences occur.

4.2.1 Post hoc analysis. Because we reject H_0 , a post hoc analysis of the concentration levels to test the hypotheses given in Eq 24 are performed next with results in Table 8. We only compare each concentration level with the control group because the average effect of an increase in the concentration level on swim speed is negative, which can be seen in Fig 11. The tests employ a Bonferroni correction so that the significance level for each test is the nominal significance level ($\tau = 0.05$) divided by the number of tests. The functional contrasts shown in

Table 8. The F_{FT} test results for the main effects and post hoc tests for the entire domain and each of the light/dark regions. For each region name, the letter refers to the brightness condition (light/dark), and the number corresponds to whether the region is the first or second light/dark period. In the “Decision” column, FTR means “fail to reject.”

Region	H_0	F_{FT}^{crit}	F_{FT}^{obs}	p -value	Sig. Level	Decision
L1-D2: $t \in [10, 50)$	$\alpha_1(t) = \dots = \alpha_5(t)$	2.14	19.39	0.0000	τ	Reject
	$\alpha_1(t) = \alpha_2(t)$	0.36	0.26	0.0393	$\tau/4$	FTR
	$\alpha_1(t) = \alpha_3(t)$	0.37	0.86	0.0001	$\tau/4$	Reject
	$\alpha_1(t) = \alpha_4(t)$	0.37	2.35	0.0000	$\tau/4$	Reject
	$\alpha_1(t) = \alpha_5(t)$	0.35	5.11	0.0000	$\tau/4$	Reject
L1: $t \in [10, 20)$	$\alpha_1(t) = \alpha_3(t)$	0.46	0.22	0.0695	$\tau/12$	FTR
	$\alpha_1(t) = \alpha_4(t)$	0.49	0.54	0.0028	$\tau/12$	Reject
	$\alpha_1(t) = \alpha_5(t)$	0.50	1.60	0.0000	$\tau/12$	Reject
D1: $t \in [20, 30)$	$\alpha_1(t) = \alpha_3(t)$	0.59	0.83	0.0003	$\tau/12$	Reject
	$\alpha_1(t) = \alpha_4(t)$	0.59	2.60	0.0000	$\tau/12$	Reject
	$\alpha_1(t) = \alpha_5(t)$	0.55	6.09	0.0000	$\tau/12$	Reject
L2: $t \in [30, 40)$	$\alpha_1(t) = \alpha_3(t)$	0.50	0.87	0.0003	$\tau/12$	Reject
	$\alpha_1(t) = \alpha_4(t)$	0.55	1.59	0.0000	$\tau/12$	Reject
	$\alpha_1(t) = \alpha_5(t)$	0.54	2.89	0.0000	$\tau/12$	Reject
D2: $t \in [40, 50)$	$\alpha_1(t) = \alpha_3(t)$	0.63	1.32	0.0000	$\tau/12$	Reject
	$\alpha_1(t) = \alpha_4(t)$	0.63	3.22	0.0000	$\tau/12$	Reject
	$\alpha_1(t) = \alpha_5(t)$	0.60	6.17	0.0000	$\tau/12$	Reject

<https://doi.org/10.1371/journal.pone.0300636.t008>

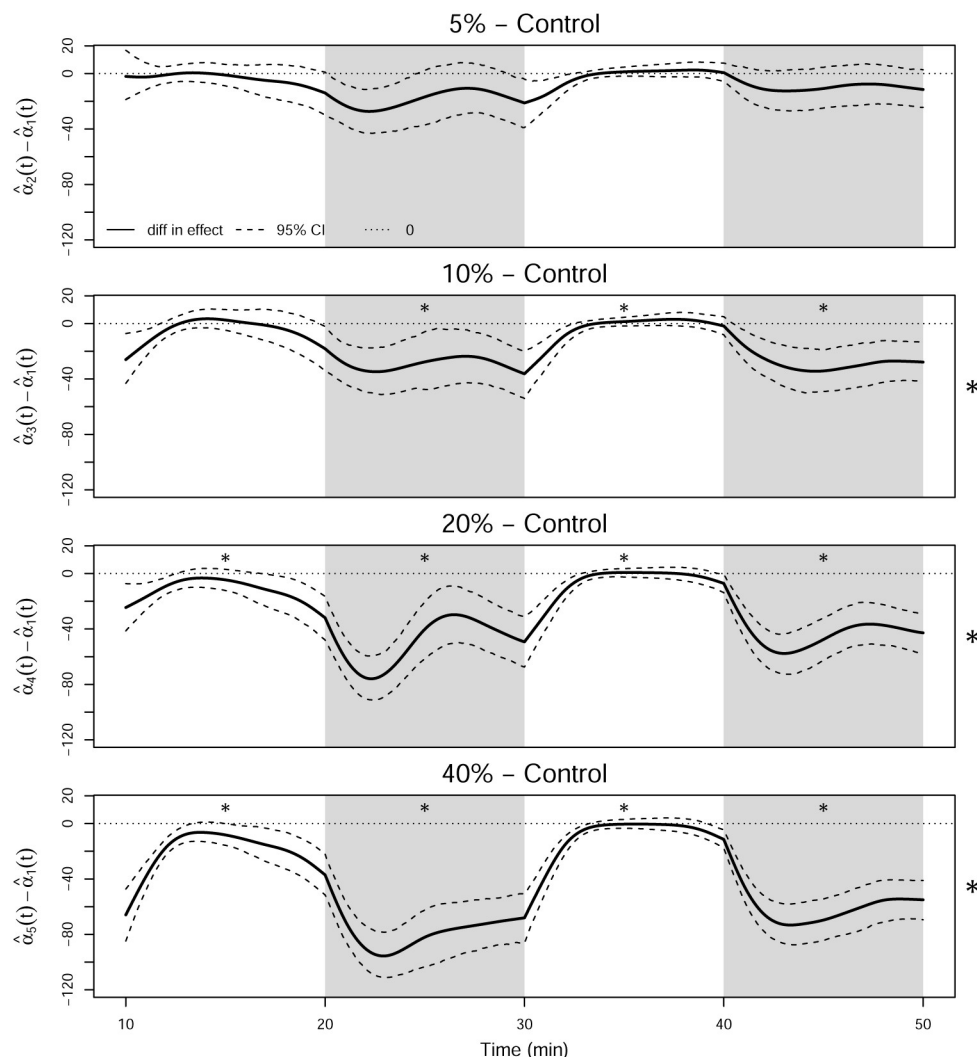


Fig 14. Functional post hoc contrasts for the concentration levels. The dashed line is the pointwise 95% confidence interval using 1000 bootstrap replicates. The asterisks on the right of the plots indicate significance of the F_{FT} test for the concentration effect across the entire domain, and the asterisks inside the plots indicate significance within the corresponding brightness period.

<https://doi.org/10.1371/journal.pone.0300636.g014>

Fig 14 are constructed by calculating $\hat{\alpha}_i(t) - \hat{\alpha}_1(t)$ with $i \neq 1$ and forming 95% pointwise bootstrapped confidence intervals with 1000 replicates. We determine that there is a significant difference between the mean swim speeds of the control group and the groups with concentrations of 1-heptanol at 10%, 20%, and 40% of their LC_{50} .

With the functional approach, it is also possible to explore subregions of the post-hoc contrasts to determine under which brightness conditions the differences occur. Note that because the 5% concentration level was not significantly different from the control group, no further comparisons are made. Thus, there are twelve comparisons, so the significance level for each test is $\tau/12$. Table 8 shows that the 1-heptanol 20% ($\alpha_4(t)$) and 40% ($\alpha_5(t)$) LC_{50} groups are significantly different from the control group in all four subregions, while the 1-heptanol 10% ($\alpha_3(t)$) LC_{50} is significant in every region except for the first light region. While significance in

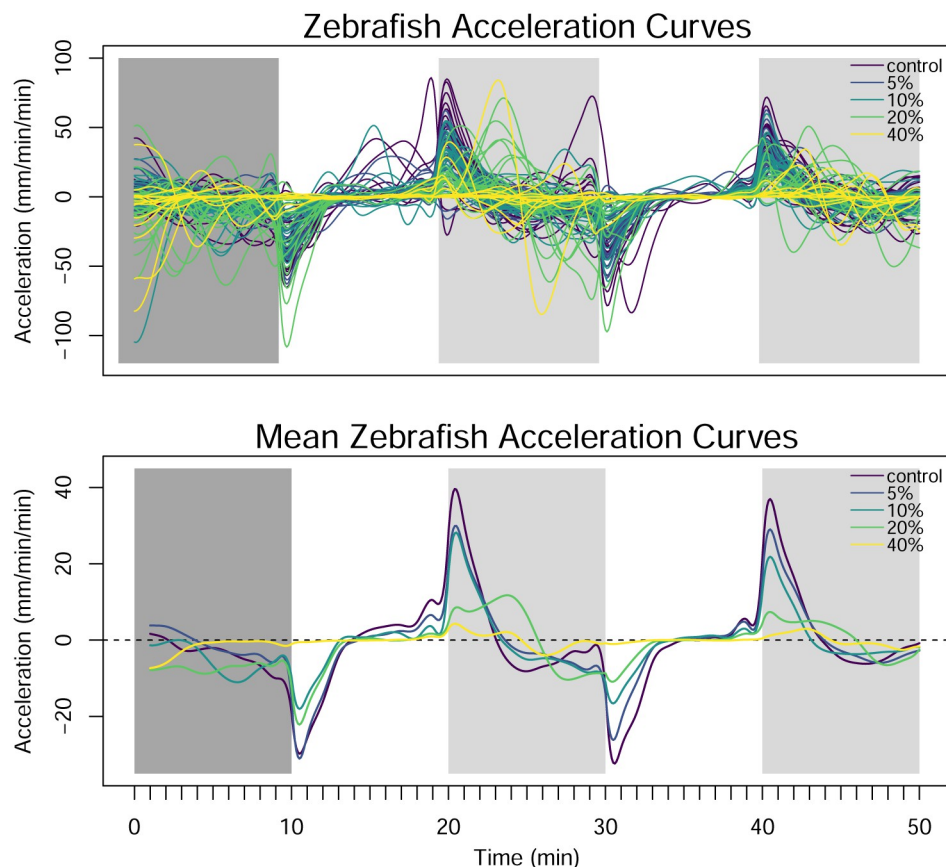


Fig 15. The first derivative of the smoothed swim speed curves (top), and the functional averages of the first derivative of the smoothed swim speed curves for each concentration level (bottom).

<https://doi.org/10.1371/journal.pone.0300636.g015>

the dark regions is expected, it is surprising to find significance in the light regions. This is likely due to the time it takes for the fish to slow down at the beginning of the light periods.

4.3 Startle response analysis

Some researchers are particularly interested in the change in swim speed in response to a change in conditions, called the startle response. There have been many approaches to analyzing the startle response, most of which involve performing univariate or repeated measures ANOVA for either the maximum swim speed post-stimulus [40, 41]; the average duration of the response [41]; or the average swim speed for short periods directly preceding and/or following a change in conditions [5, 18, 42, 43]. Sometimes, *t*-tests or nonparametric alternatives are performed for every time point against a baseline speed or control group [44, 45]. Another approach involves grouping swim behavior into speed thresholds that indicate bursting and/or freezing speeds [8, 15, 17]. Several researchers use univariate statistical methods to capture more information about the startle response, such as analyzing both the average swim speed and the net change in swim speed from the end of one period to the beginning of the following period [9, 46] or analyzing responsiveness, response latency, and maximum angular velocity during the startle [47]. All of this information is captured using the functional approach. Because the startle response occurs over some period of time and may last varying amounts of

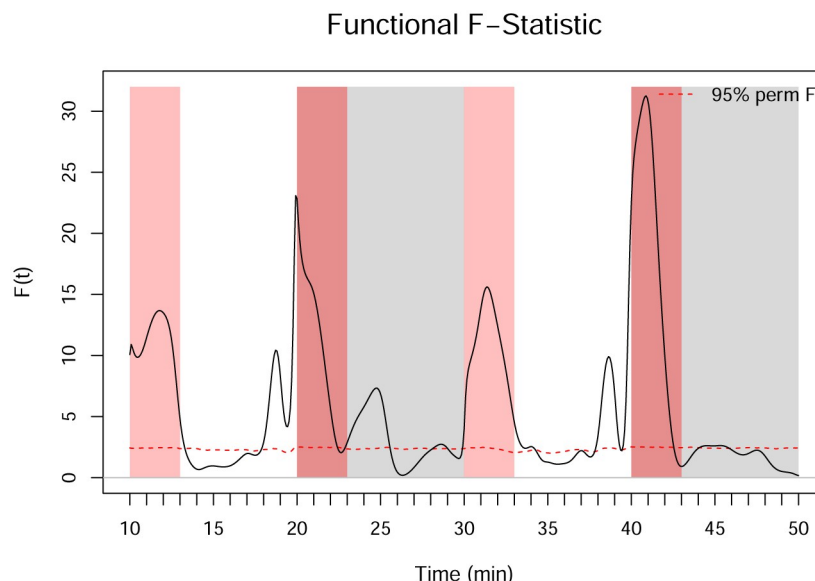


Fig 16. Functional $F^{obs}(t)$ statistic of the acceleration curves, where the red dashed line indicates the 95% pointwise percentile calculated from the permutations. Regions of interest for the startle response are represented in light pink shading for dark to light transitions and dark pink shading for light to dark transitions.

<https://doi.org/10.1371/journal.pone.0300636.g016>

time for each individual, the functional representation of the observations provides a natural way to visualize and quantify it.

A novel way to analyze the startle response is to compute the acceleration function of the zebrafish. Given smoothed curves of swim speed, the acceleration can be found by taking their first derivative, as shown in the top panel of Fig 15. When the light is turned on, there is a downward spike in the acceleration, and conversely, there is an upward spike when time the light is turned off. The zebrafish appear to accelerate or decelerate for approximately three minutes after a change in the brightness conditions. To assess if the acceleration differs across the concentration levels, the average of the acceleration curves for each concentration level is plotted in the bottom panel of Fig 15. As the concentration of 1-heptanol increases, the strength of the average response of the zebrafish to the change in brightness decreases.

One-way FANOVA can also be applied to the smooth acceleration curves to test for a significant difference in the acceleration of the groups for the three minutes after a change in brightness. The functional F -statistic for the entire domain is shown in Fig 16. Because the acclimation period is not relevant for this analysis, it is not included in the remaining figures in this section. There is a significant difference in the average acceleration across the concentration groups for approximately two to three minutes after the light turns on and a very strong spike in the F -statistic for two to three minutes after the light turns off. Thus, we perform post hoc analysis to determine which of the groups are significantly different from the control group during those periods. Complete results are given in Table 9. The contrasts shown in Fig 17 indicate that the group with the highest concentration of 1-heptanol has a significantly smaller startle response than the control group at the beginning of each period, which is shown in Table 9. The 20% LC_{50} concentration group has a significantly smaller startle response than the control group at the beginning of each period except for the first light period, and the 10% LC_{50} concentration group has a significantly smaller startle response than the control group at the beginning of the last two periods.

Table 9. The F_{FT} test results for the main effects and post hoc tests for the acceleration curves across the entire domain and the first three minutes of each of the light/dark regions. In the “Decision” column, FTR means “fail to reject.”

Region	H_0	F_{FT}^{crit}	F_{FT}^{obs}	p -value	Sig. Level	Decision
L1-D2: $t \in [10, 50]$	$\alpha_1(t) = \dots = \alpha_5(t)$	1.50	7.94	0.0000	τ	Reject
	$\alpha_1(t) = \alpha_2(t)$	0.23	0.14	0.1086	$\tau/4$	FTR
	$\alpha_1(t) = \alpha_3(t)$	0.23	0.39	0.0007	$\tau/4$	Reject
	$\alpha_1(t) = \alpha_4(t)$	0.22	1.16	0.0000	$\tau/4$	Reject
	$\alpha_1(t) = \alpha_5(t)$	0.23	1.76	0.0000	$\tau/4$	Reject
L1: $t \in [10, 13]$	$\alpha_1(t) = \alpha_3(t)$	0.67	0.59	0.0068	$\tau/12$	FTR
	$\alpha_1(t) = \alpha_4(t)$	0.66	0.31	0.0452	$\tau/12$	FTR
	$\alpha_1(t) = \alpha_5(t)$	0.67	2.80	0.0000	$\tau/12$	Reject
D1: $t \in [20, 23]$	$\alpha_1(t) = \alpha_3(t)$	0.60	0.28	0.0499	$\tau/12$	FTR
	$\alpha_1(t) = \alpha_4(t)$	0.60	2.03	0.0000	$\tau/12$	Reject
	$\alpha_1(t) = \alpha_5(t)$	0.62	2.88	0.0000	$\tau/12$	Reject
L2: $t \in [30, 33]$	$\alpha_1(t) = \alpha_3(t)$	0.57	0.87	0.0008	$\tau/12$	Reject
	$\alpha_1(t) = \alpha_4(t)$	0.58	1.58	0.0000	$\tau/12$	Reject
	$\alpha_1(t) = \alpha_5(t)$	0.60	3.08	0.0000	$\tau/12$	Reject
D2: $t \in [40, 43]$	$\alpha_1(t) = \alpha_3(t)$	0.58	0.95	0.0006	$\tau/12$	Reject
	$\alpha_1(t) = \alpha_4(t)$	0.59	3.28	0.0000	$\tau/12$	Reject
	$\alpha_1(t) = \alpha_5(t)$	0.57	4.74	0.0000	$\tau/12$	Reject

<https://doi.org/10.1371/journal.pone.0300636.t009>

The startle response is an important behavioral response of interest in PBR literature, and the functional approach provides researchers with a systematic approach to analyzing the startle response. By using FANOVA, we are able to analyze both the duration and magnitude of the startle response and determine differences in treatment groups over regions of the domain of varying lengths. This could offer a new approach for diagnostic evaluation of tipping points of chemical contaminants following exposure.

5 Conclusion

Zebrafish PBR studies are of great value to toxicological and pharmacological research. However, the typical statistical methods applied in PBR studies prevents researchers from fully exploiting all of the information contained in the data. The most common approach, univariate ANOVA for each individual's average across the entire domain, reduces each individual's values to a single value and can produce untrustworthy results due to violation of the independence of observations within each individual. Even repeated measures ANOVA, the method used in PBR literature that best accounts for the temporal dependence in observations, has limitations in its interpretability, practical use, and assumptions that may not be met. To overcome these problems, we suggest the use of functional ANOVA, a method that naturally retains the functional structure of PBR data and produces valid, fast, and interpretable results. The FANOVA presented in this paper only requires that the observations are independent and can be used to analyze the response over the entire domain, detecting which regions of the domain have the largest differences. Because of the smooth functional representation of the data, it becomes possible to take the derivative of the speed functions to find the acceleration curves, so researchers can analyze the startle response in addition to the original measurement. Furthermore, FANOVA can be applied to any subregion of the domain, so researchers can apply FANOVA to the regions where the startle responses occur.

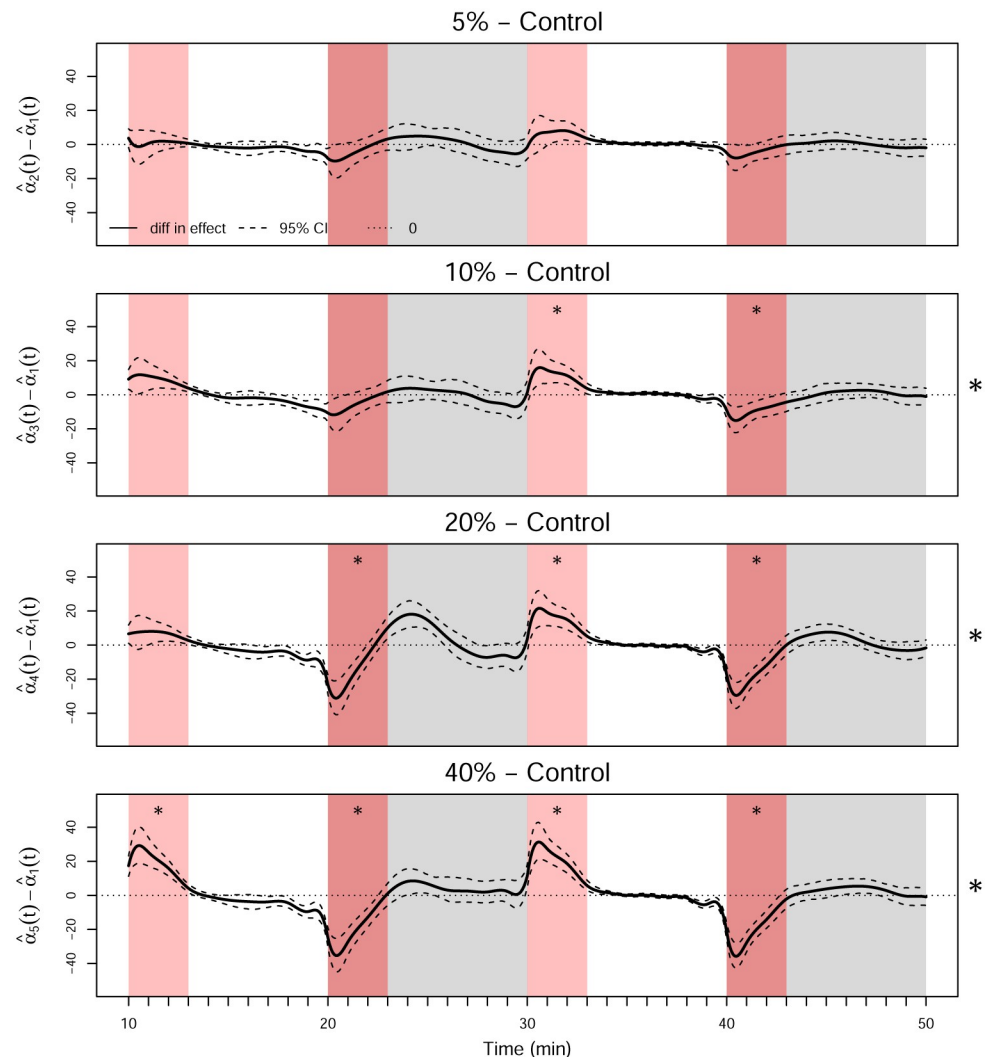


Fig 17. Functional post hoc contrasts for the acceleration curves for the concentration levels. The dashed line is the pointwise 98.75% confidence interval using 1000 bootstrap replicates. Regions of interest for the startle response are represented in light pink shading for dark to light transitions and dark pink shading for light to dark transitions. The asterisks indicate significance of the F_{FT} test for the concentration effect in the corresponding region.

<https://doi.org/10.1371/journal.pone.0300636.g017>

In this paper, we present a framework for implementing one-way FANOVA that has minimal assumptions and demonstrate how to apply it. We compare the results to those obtained from univariate and repeated measures ANOVA, illustrating the ways FANOVA overcomes their limitations. In order to facilitate the use of FANOVA in future PBR research, the R code and data used to perform the analyses in this paper and to create every figure and table are provided on the Harvard Dataverse data repository [37].

It is possible to extend this work beyond light/dark transition tests to other responses that are considered functional. Many studies that examine zebrafish also collect data on spontaneous movement/tail flexes [21, 48–50], bioaccumulation [51], and/or hatching rate [11]. Even when the light/dark transition test is not used, general locomotor data are commonly collected to assess the behavioral response of zebrafish [3, 15, 52]. While we have not performed an exhaustive comparison of all statistical methods historically used on these types of data, we

argue that functional methods can and should be used. Any data that are collected repeatedly for the same subjects over a continuous domain can benefit from a functional approach.

Acknowledgments

We thank W. Baylor Steele for providing us with photolocomotor behavioral response data for used in the Real Data Example section. We also thank reviewers Kathryn S. Konrad and Adrian J. Green for their helpful feedback on this work.

Author Contributions

Conceptualization: Luke Durell, Bryan W. Brooks, Amanda S. Hering.

Data curation: Luke Durell.

Formal analysis: Natalie Mastin.

Funding acquisition: Bryan W. Brooks, Amanda S. Hering.

Methodology: Natalie Mastin, Amanda S. Hering.

Project administration: Amanda S. Hering.

Supervision: Amanda S. Hering.

Visualization: Natalie Mastin.

Writing – original draft: Natalie Mastin.

Writing – review & editing: Luke Durell, Bryan W. Brooks, Amanda S. Hering.

References

1. Saaristo M., Brodin T., Balshine S., Bertram M. G., Brooks B. W., Ehlman S. M., et al. (2018). Direct and indirect effects of chemical contaminants on the behaviour, ecology and evolution of wildlife. *Proceedings of the Royal Society B: Biological Sciences*, 285(1885):20181297. <https://doi.org/10.1098/rspb.2018.1297> PMID: 30135169
2. Rihel J., Prober D. A., Arvanites A., Lam K., Zimmerman S., Jang S., et al. (2010). Zebrafish behavioral profiling links drugs to biological targets and rest/wake regulation. *Science*, 327:348–351. <https://doi.org/10.1126/science.1183090> PMID: 20075256
3. Afrikanova T., Serruys A.-S. K., Buenafe O. E. M., Clinckers R., Smolders I., De Witte P. A. M., et al. (2013). Validation of the zebrafish pentylenetetrazol seizure model: Locomotor versus electrographic responses to antiepileptic drugs. *PLoS ONE*, 8(1):e54166. <https://doi.org/10.1371/journal.pone.0054166> PMID: 23342097
4. Parker T., Libourel P.-A., Hetheridge M. J., Cumming R. I., Sutcliffe T. P., Goonesinghe A. C., et al. (2014). A multi-endpoint in vivo larval zebrafish (*Danio rerio*) model for the assessment of integrated cardiovascular function. *Journal of Pharmacological and Toxicological Methods*, 69(1):30–38. <https://doi.org/10.1016/j.vascn.2013.10.002> PMID: 24140389
5. Spulber S., Kilian P., Wan Ibrahim W. N., Onishchenko N., Ulhaq M., Norrgren L., et al. (2014). PFOS induces behavioral alterations, including spontaneous hyperactivity that is corrected by dexamfetamine in zebrafish larvae. *PLoS ONE*, 9(4):e94227. <https://doi.org/10.1371/journal.pone.0094227> PMID: 24740186
6. Haigis A.-C., Ottermanns R., Schiwy A., Hollert H., and Legradi J. (2022). Getting more out of the zebrafish light dark transition test. *Chemosphere*, 295:133863. <https://doi.org/10.1016/j.chemosphere.2022.133863> PMID: 35124091
7. Lovin L. M., Kim S., Taylor R. B., Scarlett K. R., Langan L. M., Chambliss C. K., et al. (2021). Differential influences of ($\pm$) anatoxin-a on photolocomotor behavior and gene transcription in larval zebrafish and fathead minnows. *Environmental Sciences Europe*, 33(1):40. <https://doi.org/10.1186/s12302-021-00479-x> PMID: 34367861
8. Steele W. B., Kristofco L. A., Corrales J., Saari G. N., Haddad S. P., Gallagher E. P., et al. (2018). Comparative behavioral toxicology with two common larval fish models: Exploring relationships among

- modes of action and locomotor responses. *Science of The Total Environment*, 640–641:1587–1600. <https://doi.org/10.1016/j.scitotenv.2018.05.402> PMID: 30021323
9. Chen L., Huang C., Hu C., Yu K., Yang L., and Zhou B. (2012). Acute exposure to DE-71: Effects on locomotor behavior and developmental neurotoxicity in zebrafish larvae. *Environmental Toxicology and Chemistry*, 31(10):2338–2344. <https://doi.org/10.1002/etc.1958> PMID: 22833361
 10. Gyimah E., Zhu X., Zhang Z., Guo M., Xu H., Mensah J. K., et al. (2022). Oxidative stress and apoptosis in bisphenol AF–induced neurotoxicity in zebrafish embryos. *Environmental Toxicology and Chemistry*, 41(9):2273–2284. <https://doi.org/10.1002/etc.5412> PMID: 35723417
 11. Hua J., Vijver M. G., Ahmad F., Richardson M. K., and Peijnenburg W. J. (2014). Toxicity of different-sized copper nano- and submicron particles and their shed copper ions to zebrafish embryos: Toxicity of different-sized Cu particles to zebrafish embryos. *Environmental Toxicology and Chemistry*, 33(8):1774–1782. <https://doi.org/10.1002/etc.2615> PMID: 24839162
 12. Perkins E. J., To K. T., St. Mary L., Laber C. H., Bednar A. J., Truong L., et al. (2022). Developmental, behavioral and transcriptomic changes in zebrafish embryos after smoke dye exposure. *Toxics*, 10(5):210. <https://doi.org/10.3390/toxics10050210> PMID: 35622624
 13. Jarema K. A., Hunter D. L., Hill B. N., Olin J. K., Britton K. N., Waalkes M. R., et al. (2022). Developmental neurotoxicity and behavioral screening in larval zebrafish with a comparison to other published results. *Toxics*, 10(5):256. <https://doi.org/10.3390/toxics10050256> PMID: 35622669
 14. Rericha Y., Truong L., Leong C., Cao D., Field J. A., and Tanguay R. L. (2022). Dietary perfluorohexanoic acid (PFHxA) exposures in juvenile zebrafish produce subtle behavioral effects across generations. *Toxics*, 10(7):372. <https://doi.org/10.3390/toxics10070372> PMID: 35878277
 15. López-Patiño M. A., Yu L., Cabral H., and Zhdanova I. V. (2008). Anxiogenic effects of cocaine withdrawal in zebrafish. *Physiology & Behavior*, 93(1–2):160–171. <https://doi.org/10.1016/j.physbeh.2007.08.013>
 16. Zhao J., Xu T., and Yin D.-Q. (2014). Locomotor activity changes on zebrafish larvae with different 2,2',4,4'-tetrabromodiphenyl ether (PBDE-47) embryonic exposure modes. *Chemosphere*, 94:53–61. <https://doi.org/10.1016/j.chemosphere.2013.09.010> PMID: 24080000
 17. Ellis L. D. and Soanes K. H. (2012). A larval zebrafish model of bipolar disorder as a screening platform for neuro-therapeutics. *Behavioural Brain Research*, 233(2):450–457. <https://doi.org/10.1016/j.bbr.2012.05.043> PMID: 22677277
 18. Ellis L. D., Seibert J., and Soanes K. H. (2012). Distinct models of induced hyperactivity in zebrafish larvae. *Brain Research*, 1449:46–59. <https://doi.org/10.1016/j.brainres.2012.02.022> PMID: 22386495
 19. Zhang G., Truong L., Tanguay R. L., and Reif D. M. (2017). A new statistical approach to characterize chemical-elicited behavioral effects in high-throughput studies using zebrafish. *PLOS ONE*, 12(1): e0169408. <https://doi.org/10.1371/journal.pone.0169408> PMID: 28099482
 20. Legradi J., El Abdellaoui N., Van Pomeran M., and Legler J. (2015). Comparability of behavioural assays using zebrafish larvae to assess neurotoxicity. *Environmental Science and Pollution Research*, 22(21):16277–16289. <https://doi.org/10.1007/s11356-014-3805-8> PMID: 25399529
 21. Chen J., Huang C., Zheng L., Simonich M., Bai C., Tanguay R., et al. (2011). Trimethyltin chloride (TMT) neurobehavioral toxicity in embryonic zebrafish. *Neurotoxicology and Teratology*, 33(6):721–726. <https://doi.org/10.1016/j.ntt.2011.09.003> PMID: 21964161
 22. Irons T., MacPhail R., Hunter D., and Padilla S. (2010). Acute neuroactive drug exposures alter locomotor activity in larval zebrafish. *Neurotoxicology and Teratology*, 32(1):84–90. <https://doi.org/10.1016/j.ntt.2009.04.066> PMID: 19465114
 23. Kokoszka P. and Reimherr M. (2017). *Introduction to Functional Data Analysis*. Chapman and Hall/CRC, New York.
 24. Ramsay J. and Silverman B. (2005). *Functional Data Analysis*. Springer Series in Statistics. Springer, 2 edition.
 25. Fisher R. A. (1925). *Statistical Methods for Research Workers*. Oliver & Boyd: Edinburgh, 1 edition.
 26. Scariano S. M. and Davenport J. M. (1987). The effects of violations of independence assumptions in the one-way ANOVA. *The American Statistician*, 41(2):123–129. <https://doi.org/10.1080/00031305.1987.10475459>
 27. Davis C. S. (2002). *Statistical Methods for the Analysis of Repeated Measurements*. Springer Texts in Statistics. Springer, New York.
 28. Box G. (1954). Some theorems on quadratic forms applied in the study of analysis of variance problems, II. Effects of inequality of variance and of correlation between errors in the two-way classification. *The Annals of Mathematical Statistics*, 25(3):484–498. <https://doi.org/10.1214/aoms/1177728717>

29. Macey P. M., Schluter P. J., Macey K. E., and Harper R. M. (2016). Detecting variable responses in time-series using repeated measures ANOVA: Application to physiologic challenges. *F1000Res*, 5:563. <https://doi.org/10.12688/f1000research.8252.2> PMID: 27610219
30. Zhang J.-T. (2013). *Analysis of Variance for Functional Data*. Chapman and Hall/CRC, 1 edition.
31. Ramsay, J. O., Graves, S., and Hooker, G. (2022). fda: Functional data analysis R package version 6.0.5, <https://CRAN.R-project.org/package=fda>.
32. Shen Q. and Faraway J. (2004). An *F*-test for linear models with functional response. *Statistica Sinica*, 14(4):1239–1257.
33. Zhang J.-T. and Chen J. (2007). Statistical inferences for functional data. *The Annals of Statistics*, 35 (3). <https://doi.org/10.1214/009053606000001505>
34. Zhang J.-T. and Liang X. (2014). One-way ANOVA for functional data via globalizing the pointwise *F*-test: Globalizing the pointwise *F*-test. *Scandinavian Journal of Statistics*, 41(1):51–71. <https://doi.org/10.1111/sjost.12025>
35. Zhang J.-T., Cheng M.-Y., Wu H.-T., and Zhou B. (2019). A new test for functional one-way ANOVA with applications to ischemic heart screening. *Computational Statistics & Data Analysis*, 132:3–17.. <https://doi.org/10.1016/j.csda.2018.05.004>
36. Durell L., Lovin L. M., Steele W. B. IV, Mastin N., Brooks B. W., et al. (2023). Balanced two-way functional ANOVA for toxicological photolocomotor response studies. *Under Review*.
37. Mastin, N., Durell, L., Brooks, B. W., and Hering, A. S. (2023). Replication data for: Proper statistical treatment of photolocomotor behavioral response study data, Harvard Dataverse, <https://doi.org/10.7910/DVN/2XEL5H>.
38. Hubert M., Rousseeuw P. J., and Segae P. (2015). Multivariate functional outlier detection. *Statistical Methods & Applications*, 24(2): 177–202. <https://doi.org/10.1007/s10260-015-0297-8>
39. Sun Y. and Genton M. G. (2011). Functional boxplots. *Journal of Computational and Graphical Statistics*, 20(2): 316–34. <https://doi.org/10.1198/jcgs.2011.09224>
40. Qiu X., Nomichi S., Chen K., Honda M., Kang I. J., Shimasaki Y., et al. (2017). Short-term and persistent impacts on behaviors related to locomotion, anxiety, and startle responses of japanese medaka (*Oryzias latipes*) induced by acute, sublethal exposure to chlorpyrifos. *Aquatic Toxicology*, 192:148–154. <https://doi.org/10.1016/j.aquatox.2017.09.012> PMID: 28957716
41. Rice C., Ghorai J. K., Zalewski K., and Weber D. N. (2011). Developmental lead exposure causes startle response deficits in zebrafish. *Aquatic Toxicology*, 105(3–4):600–608. <https://doi.org/10.1016/j.aquatox.2011.08.014> PMID: 21955963
42. Kirshenbaum A. P., Chabot E., and Gibney N. (2019). Startle, pre-pulse sensitization, and habituation in zebrafish. *Journal of Neuroscience Methods*, 313:54–59.. <https://doi.org/10.1016/j.jneumeth.2018.12.017> PMID: 30586568
43. Spagnoli S., Xue L., and Kent M. L. (2015). The common neural parasite pseudoloma neurophilia is associated with altered startle response habituation in adult zebrafish (*Danio rerio*): Implications for the zebrafish as a model organism. *Behavioural Brain Research*, 291:351–360.. <https://doi.org/10.1016/j.bbr.2015.05.046> PMID: 26028515
44. Chanin S., Fryar C., Varga D., Raymond J., Kyzar E., Enriquez J., et al. (2012). *Assessing Startle Responses and Their Habituation in Adult Zebrafish*, volume 66 of *Neuromethods*, pages 287–300. Humana Press, Totowa, NJ.
45. Faria M., Prats E., Novoa-Luna K. A., Bedrossiantz J., Gómez-Canela C., Gómez-Oliván L. M., et al. (2019). Development of a vibrational startle response assay for screening environmental pollutants and drugs impairing predator avoidance. *Science of The Total Environment*, 650:87–96. <https://doi.org/10.1016/j.scitotenv.2018.08.421> PMID: 30196226
46. He J., Yang D., Wang C., Liu W., Liao J., Xu T., et al. (2011). Chronic zebrafish low dose decabrominated diphenyl ether (BDE-209) exposure affected parental gonad development and locomotion in F1 offspring. *Ecotoxicology*, 20(8):1813–1822. <https://doi.org/10.1007/s10646-011-0720-3> PMID: 21695510
47. Panlilio J. M., Jones I. T., Salanga M. C., Aluru N., and Hahn M. E. (2021). Developmental exposure to domoic acid disrupts startle response behavior and circuitry in zebrafish. *Toxicological Sciences*, 182 (2):310–326. <https://doi.org/10.1093/toxsci/kfab066> PMID: 34097058
48. Bertram M. G., Martin J. M., McCallum E. S., Alton L. A., Brand J. A., Brooks B. W., et al. (2022). Frontiers in quantifying wildlife behavioural responses to chemical pollution. *Biological Reviews*, 97 (4):1346–1364. <https://doi.org/10.1111/brv.12844> PMID: 35233915
49. Usenko C. Y., Robinson E. M., Usenko S., Brooks B. W., and Bruce E. D. (2011). PBDE developmental effects on embryonic zebrafish. *Environmental Toxicology and Chemistry*, 30(8):1865–1872. <https://doi.org/10.1002/etc.570> PMID: 21560146

50. Wang X., Dong Q., Chen Y., Jiang H., Xiao Q., Wang Y., et al. (2013). Bisphenol A affects axonal growth, musculature and motor behavior in developing zebrafish. *Aquatic Toxicology*, 142–143:104–113. PMID: [23994041](#)
51. Li Y., Liang H., and Qiu L. (2021). Enantioselective bioaccumulation of the chiral insecticide indoxacarb in zebrafish (*Danio rerio*). *Environmental Toxicology and Chemistry*, 40(4):1007–1016. <https://doi.org/10.1002/etc.4943> PMID: [33238038](#)
52. Lange M., Norton W., Coolen M., Chaminade M., Merker S., Proft F., et al. (2012). The ADHD-susceptibility gene lphn3.1 modulates dopaminergic neuron formation and locomotor activity during zebrafish development. *Molecular Psychiatry*, 17(9):946–954. <https://doi.org/10.1038/mp.2012.29> PMID: [22508465](#)