Check for updates

Received: 11 March 2025 | Revised: 30 September 2025 | Accepted: 17 November 2025

DOI: 10.1111/cobi.70224

Conservation Biology

CONTRIBUTED PAPER

An experimental approach to assess the combined effects of multiple stressors on a large vertebrate species

Enrico Pirotta[1] | Catriona M. Harris[1] | Lori H. Schwacke[2] | Len Thomas[1] | Jason B. Allen[3] | Aaron A. Barleycorn[3] | Marco Casoli[4,5] | Jonathan Crossman[3] | Laura Aichinger Dias[6,7] | Laura Engleby[8] | Erin LaBrecque[2] | Jenny A. Litz[6] | Katherine A. McHugh[3] | Patrick J. O. Miller[4] | Michael J. Moore[9] | Theoni Photopoulou[1] | Brian M. Quigley[10] | Cynthia R. Smith[10] | Brandon L. Southall[11] | Todd R. Speakman[10] | Ryan Takeshita[10] | Christina N. Toms[12,3] | Randall S. Wells[3] | Eric S. Zolman[10] | Peter L. Tyack[4]

[1]Centre for Research into Ecological and Environmental Modelling, University of St Andrews, St Andrews, UK

[2]Marine Mammal Commission, Bethesda, Maryland, USA

[3]Brookfield Zoo Chicago's Sarasota Dolphin Research Program, c/o Mote Marine Laboratory, Sarasota, Florida, USA

[4]Sea Mammal Research Unit, Scottish Oceans Institute, University of St Andrews, St Andrews, UK

[5]Morigenos – Slovenian Marine Mammal Society, Piran, Slovenia

[6]Marine Mammal and Turtle Division, Southeast Fisheries Science Center, National Oceanic and Atmospheric Administration, Miami, Florida, USA

[7]Cooperative Institute for Marine and Atmospheric Studies, University of Miami, Miami, Florida, USA

[8]Marine Mammal Branch, Southeast Region, National Oceanic and Atmospheric Administration Fisheries Service, St. Petersburg, Florida, USA

[9]Biology Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA

[10]National Marine Mammal Foundation, San Diego, California, USA

[11]Southall Environmental Associates, Inc., Aptos, California, USA

[12]New College of Florida, Sarasota, Florida, USA

Correspondence
Enrico Pirotta, Centre for Research into Ecological and Environmental Modelling, University of St Andrews, The Observatory, Buchanan Gardens, KY16 9LZ St Andrews, Fife, Scotland, UK. Email: ep343@st-andrews.ac.uk

Article impact statement: A multistressor experiment suggests that oil spills may exacerbate other stressor effects in coastal populations of long-lived vertebrates.

Funding information
Marine Mammal Commission; Office of Naval Research, Grant/Award Numbers: N000142012697, N000142112096; Strategic Environmental Research

Abstract

The design of experiments to investigate the combined effects of multiple stressors requires exposing target organisms to multiple combinations of stressor doses. Concurrent manipulation of stressors is often infeasible with wildlife, but long-lasting health effects allow individual health to be used as an integrator of prior stressor exposure. The population of bottlenose dolphins in Barataria Bay, Louisiana, experienced long-lasting health effects after exposure to oil from the *Deepwater Horizon* spill. We investigated whether compromised health status affects dolphins' ability to respond to other stressors, specifically vessel traffic, potentially leading to increased collision risk. We used this case study to develop a hypothesis-driven experimental approach to assess the combined effects of multiple stressors in a large, long-lived vertebrate species. We conducted controlled vessel approaches to test whether the health status of targeted dolphins affected their behavioral responses. Our results highlighted some effects of health, suggesting that oil spills may exacerbate the effects of other stressors in coastal populations. For example, lung dis-

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2026 The Author(s). *Conservation Biology* published by Wiley Periodicals LLC on behalf of Society for *Conservation Biology*.

and Development Program, Grant/Award Numbers: RC20-1097, RC20-7188, RC21-3091

ease was associated with a delayed dive response, which could affect the ability to vertically avoid vessels and lead to increased collision risk. However, health effects on response probability were overall smaller than anticipated, and other contextual variables (e.g., sex, age, calf presence, prior experience, and exposure context) contributed to response variability. Our work demonstrated the value of formalizing stressor interactions as multidimensional dose–response functions and showed the feasibility of an experimental, multiple-stressor study in a wild system in which individual health status can be used as an integrator of prior stressor exposure. This approach has broad implications for other species that are difficult to handle experimentally. The quantification and management of the cumulative risk from multiple stressors on wildlife will require a combination of empirical and mechanistic approaches to inform long-term, population consequences.

KEYWORDS
bottlenose dolphins, controlled exposure experiments, dose–response functions, health, multiple stressors, oil spill, vessel traffic

INTRODUCTION

The multiple stressors that human activities introduce into the environment may combine in unexpected ways to affect wildlife populations and their ecosystems (Crain et al., 2008; Orr et al., 2020; Steffen et al., 2011). Although environmental legislation mandates their evaluation under several jurisdictions, current evidence and methods are insufficient to support these assessments (National Academies, 2017; Tyack et al., 2022). The combined effects of stressors can be predicted using information on the mechanisms through which stressors operate and interact, or estimated from observational or experimental studies involving combinations of stressor doses. Pirotta et al. (2022) argue that all available approaches sit on a spectrum between data-driven and process-driven methods and that the position selected along the spectrum should be dictated by the availability of data, given the information required by management (e.g., the need to evaluate specific combinations of stressor doses). Among empirical approaches, methods that rely on experimentation are desirable if the confounding effects of contextual variables cannot be monitored or controlled in a purely observational study (Boyd et al., 2018). Targeted experiments can also support the testing of specific hypotheses about underlying interaction mechanisms (Boyd et al., 2018; Griffen et al., 2016).

Dose–response functions (also referred to as stressor- or exposure-response functions, response or reaction norms, or performance curves, depending on the context) have been used across disciplines to define the effects of single stressors (Calabrese, 2016; Rosenfeld et al., 2022; Segurado et al., 2022; van Moorsel et al., 2023). These functions can be empirical or informed by the underlying mechanisms, and they offer a unifying framework to quantify and predict a response of interest given the increasing level of a stressor (Rosenfeld et al., 2022). For example, controlled exposure experiments (CEEs) have been conducted to assess cetacean responses to anthropogenic noise (Miller et al., 2014; Southall et al., 2016). In these studies, the probability of showing a behavioral change has been modeled as a function of the increasing acoustic dose (e.g., sound pressure level) (Harris et al., 2018), in some cases capturing the influence of other contextual variables, such as behavioral state or prior exposure to the stressor (Miller et al., 2014; Southall et al., 2019). The dose–response framework has also been proposed as a way to evaluate interactions in multistressor research (Carmichael et al., 2025; Pirotta et al., 2022; Rosenfeld et al., 2022). In this context, the aim is to estimate how the shape of the dose–response function for one stressor is modified in the presence of a second stressor, resulting in a dose–response surface (Carmichael et al., 2025; Gennings et al., 2005; National Academies, 2017; Pirotta et al., 2022).

Experimental design for multistressor studies requires exposing target organisms to multiple combinations of doses of the stressors of interest along a gradient (Boyd et al., 2018). Although this can be feasible for smaller organisms that can be housed in a laboratory or in tractable field settings, manipulating the exposure of large vertebrate species to more than one stressor in the wild presents logistical and ethical challenges. However, when a stressor has a long-lasting effect on the health of individual animals, heterogeneity in health across experimental subjects can be used as a proxy for their differential responses to prior stressor exposure (Pirotta et al., 2022). *Health* can be defined as "the ability of an organism to adapt to and manage threats to survival and reproduction," and different health indicators can act as integrators of stressor effects operating along specific adverse outcome pathways (Tyack et al., 2022). Experimentally exposing animals in different health states to additional stressors can thus shed light on whether the consequences of having experienced a prior stressor modify the effect of subsequent stressor exposures. Examples of this approach are rare in the literature and are largely focused on the effect of nutritional condition on risk-taking behavior in terrestrial or small species (Moran et al., 2021). Among large marine vertebrates, Siegal et al. (2022) found that responses of northern bottlenose whales (*Hyperoodon ampullatus*) to experimental noise exposure vary as a function of individual body condition.

A resident population of common bottlenose dolphins (*Tursiops truncatus truncatus*, Fordyce & Perrin, 2025) inhabits Barataria Bay in Louisiana (USA) (Hayes et al., 2024; McDonald et al., 2017; Wells et al., 2017). Individuals in this population have experienced detrimental and long-lasting health effects

15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224 Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

from exposure to oil released following the explosion and subsequent sinking of the *Deepwater Horizon* drilling rig in 2010 (Schwacke et al., 2014, 2022; Smith et al., 2017, 2022). Health effects, including lung disease, decreased dolphin probability of survival (Schwacke et al., 2014; Smith et al., 2017). Effects on the hypothalamus–pituitary–adrenal (HPA) axis led to observations of adrenal cortex thinning in dead dolphins (Venn-Watson et al., 2015) and impaired stress response in live dolphins (Schwacke et al., 2014). Researchers conducted a series of catch-and-release health assessments, collecting data on these animals that have supported the development of a model for predicting survival probability from veterinary records (Schwacke et al., 2023) and the quantification of oil spill effects on population dynamics (Schwacke et al., 2017, 2022). Coastal dolphin populations are exposed to many other anthropogenic stressors in these habitats (Wells & Fahlman, 2024), including disturbance from and risk of being struck by vessels (Buckstaff, 2004; Nowacek et al., 2001; Pirotta et al., 2015; Wells & Scott, 1997; Wells et al., 2008). The relevance of this stressor is evidenced by the high proportion of scars from vessel interactions in this and other coastal dolphin populations (E. S. Zolman, unpublished data; Van Waerebeek et al., 2007). Poor health status resulting from oil exposure may affect the ability or propensity of an individual to respond to an oncoming vessel, which could in turn increase the risk of collision. In other words, there may be combined effects of the oil spill and vessel traffic on survival probability.

We used the population of bottlenose dolphins in Barataria Bay as a case study to develop a hypothesis-driven experiment to assess the combined effects of multiple stressors. Specifically, we approached individual dolphins with a vessel in a controlled fashion and evaluated their responses along with previously collected health data to investigate the effect of health status on the probability or type of their behavioral responses. To our knowledge, this represents the first time the combined effects of multiple stressors have been investigated experimentally in a large marine vertebrate species. Therefore, aside from addressing the management-relevant question of whether an oil spill might affect the risk of collision in a coastal dolphin population, we explored the feasibility of this kind of study in a wild system, which has broad implications for other species that are difficult to handle and for which only some of the contextual variables can be controlled.

METHODS

Ethical use of animal subjects

All data collection was conducted under National Oceanic and Atmospheric Administration (NOAA) Fisheries Marine Mammal Health and Stranding Response Program permit 932–1905 and was approved by the University of St Andrews Animal Welfare and Ethics Review Board. Initial development of protocols in Sarasota Bay was performed under NOAA Fisheries Scientific Research Permit 20455, with Institutional Animal Care and Use Committee approval by Mote Marine Laboratory.

Health assessments

We synthesized data from dolphin catch-and-release health assessments in Barataria Bay conducted from 2011 through 2018 and previously published by Smith et al. (2017) and Schwacke et al. (2023). We conducted additional health assessments in June 2023. We used catch-and-release methods and diagnostic sampling techniques described in Wells et al. (2004), Schwacke et al. (2014), and Barratclough et al. (2019) and summarized in Appendix S1.

Health metrics

We processed the data from the health assessments to generate 4 health metrics: lung disease, HPA dysfunction, and 2 overall health scores. We assigned a lung disease score (normal, mild, moderate, severe) to each individual based on ultrasound findings (Smith et al., 2022) and converted it to a binary variable: 0, normal or mild; 1, moderate or severe. We compared serum cortisol measures with a lower threshold (5th percentile) calculated from dolphins sampled from putatively healthy populations (Schwacke et al., 2023). The resulting binary flag corresponded to cases where serum cortisol measure was below the 5th percentile, suggesting dysfunction of the HPA axis based on a lack of expected cortisol increase after capture and restraint (Schwacke et al., 2023).

We used the veterinary expert system for outcome prediction (VESOP) model described by Schwacke et al. (2023) to estimate overall health scores. The VESOP model uses a suite of hematological and serum biochemical markers, including markers of inflammation (neutrophil count, albumin, globulin, alkaline phosphatase, hemoglobin) and metabolism (body mass index, cholesterol, glucose), along with hormone levels as indicators of neuroendocrine status (cortisol, triiodothyronine) and lung disease scores, as described above. The VESOP scores range from 0.0 to 1.0 and represent the predicted 1- and 2-year survival probability (hereafter, 1- and 2-year VESOP scores) for each dolphin.

Controlled exposure experiments

In April 2022 and June–July 2023, we carried out a series of CEEs in Barataria Bay that involved vessel approaches on exposure subjects (hereafter, *focal animals*). We used an observation vessel (7-m Grady-White with a single outboard Yamaha engine and an observation tower) to identify suitable experimental candidates. In 2022, these were animals in the Barataria Bay photoidentification catalog (maintained by the National Marine Mammal Foundation) that had their health assessed in previous years (2011–2018). In June 2023, we focused on the animals that had just been released after health assessment. In July 2023, we prioritized remaining animals that had their health assessed that year and subsequently selected known individuals with recent health data (Schwacke et al., 2023). In June 2023, we conducted

15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224, Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

10 CEEs on the same day as an animal's health assessment. In those instances, we waited at least 30 min after release from the health assessment before starting the collection of baseline behavioral data (see below) to allow time for any acute response from being restrained to subside (Casoli et al., 2022). Where possible, we targeted dolphin groups in shallow water (<2 m) because dolphins are more easily observed and more likely to respond to vessels in shallow water (Nowacek et al., 2001). We also aimed to avoid situations in which other boats were nearby, and socializing dolphin groups.

Five to 7 researchers were on the observation vessel, depending on the day. Once a focal animal was selected, a dedicated observer initiated behavioral data collection, which involved documenting the focal animal's group size, behavioral state, spread, and synchrony every 3 min (definitions in Appendix S2). Another dedicated observer monitored the surfacing events of the focal dolphin, recording the time, estimated distance, and magnetic bearing from the observation vessel (via handheld compass) at each surfacing event. Visual estimates were calibrated by checking distances of reference objects in the environment with a laser range finder prior to starting a follow and whenever possible during follows. Observers recorded surfacing events (including the estimated distance and bearing) with a voice recorder synchronized to GPS time and held by the boat driver to ensure precise timing. When available, the signal from the VHF transmitter helped track the focal animal and its surfacing events. Another person onboard the observation vessel launched and retrieved a drone to record video of the CEE. The observation vessel aimed to remain ~40–60 m from the focal animal throughout the approach, moving at a slow and steady speed to minimize any disturbance. Meanwhile, the approach vessel (a Sea-Doo Spark personal watercraft [PWC]) aimed to remain ~500 m from the focal animal during the baseline phase, while maintaining a low and steady speed. A person on the observation vessel was in contact with the driver of the PWC to ensure coordination of the approach. We collected ~10 min of baseline behavioral data before initiating the approach.

We developed the experimental approach protocol (see example in Figure 1) following expert discussions and pilot trials in 2021 near Sarasota, Florida, USA, where previous CEEs had been conducted (Nowacek et al., 2001), and subsequently in Barataria Bay, to ensure that it was consistent across experiments and that it resulted in a range of responses by targeted individuals. Nowacek et al. (2001) found that erratic approaches caused the strongest behavioral responses in bottlenose dolphins. Therefore, we chose this type of approach because our aim was to elicit clear responses and assess variation relative to health status. To achieve an erratic but repeatable approach, the PWC moved toward the focal animal at ploughing speed (fast enough to produce a wake, but slow enough to not plane on top of the water) in a continuous zigzag pattern. The PWC changed direction every ~15 s (which corresponded to ~100-m-long zigzag legs). The angle between consecutive legs was variable, but the PWC driver aimed to achieve an average of ~40–70°. If the animal changed direction, the PWC gradually adjusted its path to ensure that the closest point of approach (CPA) was as near as possible (<10 m) while avoiding risk of collision.

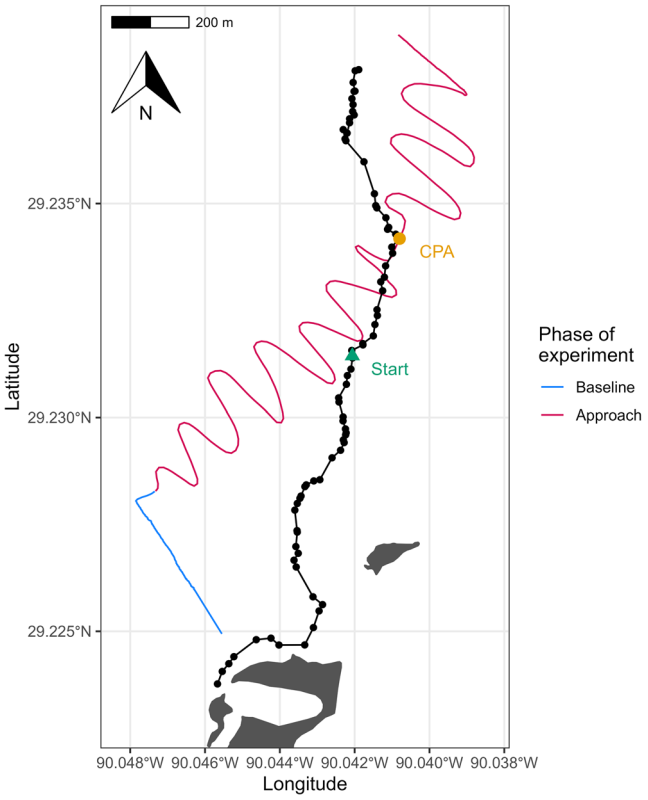


FIGURE 1 Example of an experiment in which a focal dolphin was exposed to a controlled vessel approach (controlled exposure experiment [CEE]) (continuous line, approach vessel; colors, phase of the CEE; black dots and line, track of the focal dolphin; green triangle, start of the approach on the dolphin track; orange dot, closest point of approach [CPA]).

After CPA, the PWC continued its path past the focal animal. Focal-follow data collection continued after CPA for ~10 min.

Data processing

We used the distance and bearing estimated by the researcher on the observation vessel and the observation vessel's GPS track to estimate the coordinates of the dolphin at each surfacing with geosphere (Hijmans, 2016) for R (R Core Team, 2024). We synched the drone footage with the voice recordings in Adobe Premiere Pro and used it to extract the time of any missed or dubious surfacings and to correct obvious errors (e.g., differing angle estimates from 2 close surfacings). We then used the time series of dolphin surfacing coordinates to compute 6 metrics: cumulative bearing change (degrees), turning angle (degrees), horizontal displacement rate (meters per second) (hereafter, *speed*), acceleration (meters per second squared), dive duration (seconds), and distance from the approach vessel (meters) at each surfacing (details in Appendix S3).

We also interpolated dolphin positions at every second and used them to identify salient points in the approach. The CPA marked the moment when the interpolated dolphin location was closest to the interpolated PWC location. After exploration of the PWC data, we defined the start of the approach as the

15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224, Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

first position where the PWC displacement rate was greater than 2.78 m/s within a 5- to 10-min window before CPA (the latter criterion was adjusted manually for different CEEs, if necessary). We validated these salient points visually based on the time series of distance between the dolphin and the PWC. Finally, we generated a simple animation of each CEE to visualize the approach vessel and dolphin interpolated locations, which we also used to validate the salient points in the approach.

Expert scoring

We took an expert elicitation approach to mark the occurrence of behavioral changes during the CEEs that might be related to the PWC approach (e.g., Miller et al., 2012; Southall et al., 2007, 2019). We selected a group of 8 researchers with expertise in dolphin behavior, data collection, or expert scoring. One of the experts had to be replaced after a first set of 13 scored CEEs. We provided experts with a set of anonymized plots that showed the time series of the metrics described above and a map for each CEE (Figure 2), the CEE animations, and a scoring protocol (Appendix S5). Each expert was then invited to independently score the first occurrence of any change in movement directionality, speed and acceleration, dive duration, or group metrics (spread, synchrony, and size) that they deemed to be associated with the PWC approach, as well as the occurrence of an orienting response or horizontal movement away from the PWC (definitions in Appendix S5). We asked experts to mark the phase of the CEE in which they thought the identified change had occurred. They were instructed to choose one among the following 5 phases: before the approach, at the start of the approach, during the approach, at CPA, or after CPA. The experts then discussed these scores during a group meeting, where they could volunteer the rationale of their choices and their associated confidence. After the meeting, the experts had an opportunity to revisit their scores and update them considering the group discussion, before submitting final scores. When an expert could not determine whether changes occurred or not (e.g., due to missing behavioral data), the score was marked as NA. Behavioral changes scored to have occurred before the start of the PWC approach could be interpreted as responses to the presence of the PWC in the environment. In contrast, changes scored after CPA are compatible with the idea that an animal's response threshold depends on the cumulative exposure since the start of the approach.

The final data included, for each CEE and each expert, a score for a change in each of the behavioral response metrics and the associated phase in the CEE. We treated the scoring data as an ordinal response variable with 6 categories, corresponding to the 5 CEE phases listed above plus a no-response category, which was assumed to follow the after-CPA category and so was assigned a score of 6. Therefore, we assumed the no-response category corresponded to a CEE that did not achieve a sufficient intensity of the disturbance source to elicit a response by the focal animal (similar to the treatment of right censored data in a standard dose-escalation framework [e.g., Miller et al., 2014]). In the analysis, we accounted for differences in experts' tendency to score responses with an expert-level random effect (see "Statistical analyses").

Statistical analyses

We developed our analytical approach with simulated data prior to seeing real data resulting from the experiments and expert scoring. We analyzed expert scores with an ordered probit model (e.g., Winship & Mare, 1984) fitted in a Bayesian framework. We let $y_{j,e}$ be the score given by the eth expert at the jth CEE, where $y_{j,e}$ is an ordinal value ranging from 1 to 6, and i be the dolphin that was the subject of the jth CEE. In this model, we assumed that, for each score, there exists a real-valued latent variable:

$$y^*_{j,e} \sim \text{Normal}\left(\mu_{i,e}, \psi_i\right) \quad (1)$$

(i.e., a normal distribution with mean $\mu_{i,e}$ and standard deviation ψ_i).

The mean $\mu_{i,e} = \alpha_1 + \beta_{1,i} + \beta_{2,e} + \beta_3 h_i$, where α_1 is the intercept, $\beta_{1,i}$ and $\beta_{2,e}$ are dolphin- and expert-level random effects on the mean, respectively, and β_3 is the fixed effect of individual health status, h_i. The dolphin-level random effect captured differences between dolphins in their propensity to respond that are not explained by health, and the expert-level random effect captured differences between experts in their propensity to score responses. Both random effects were assumed to be normally distributed (with mean = 0 and standard deviation $\sigma_{D,1}$ and σ_E, respectively). With this formulation, we assumed there is no within-dolphin, between-CEE variation in mean responsiveness, but see Appendix S9, where we show that substituting the dolphin-level random effect with a CEE-level random effect (conflating between- and within-dolphin variation) had negligible effects on the results. We hypothesized that health status (either a continuous VESOP score or the binary lung and cortisol scores) modifies the phase of the CEE when a dolphin is going to respond by causing a shift in the underlying threshold (i.e., responding either earlier or later). We standardized VESOP scores by subtracting the mean and dividing by the standard deviation to facilitate model convergence.

The standard deviation, ψ_i, is also assumed to have a dolphin-level random effect, reflecting both the consistency with which each individual responds and the clarity with which the moment of their response can be detected. When ψ_i is small, each dolphin will tend to respond and be detected to respond at the same approach phase; when ψ_i is large, any response or detection of a response may be inconsistent among CEEs. The random effect was modeled based on the assumption $\log(\psi_i) \sim \text{Normal}(\alpha_2, \sigma_{D,2})$, where α_2 is the mean and $\sigma_{D,2}$ is the standard deviation.

15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224 Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

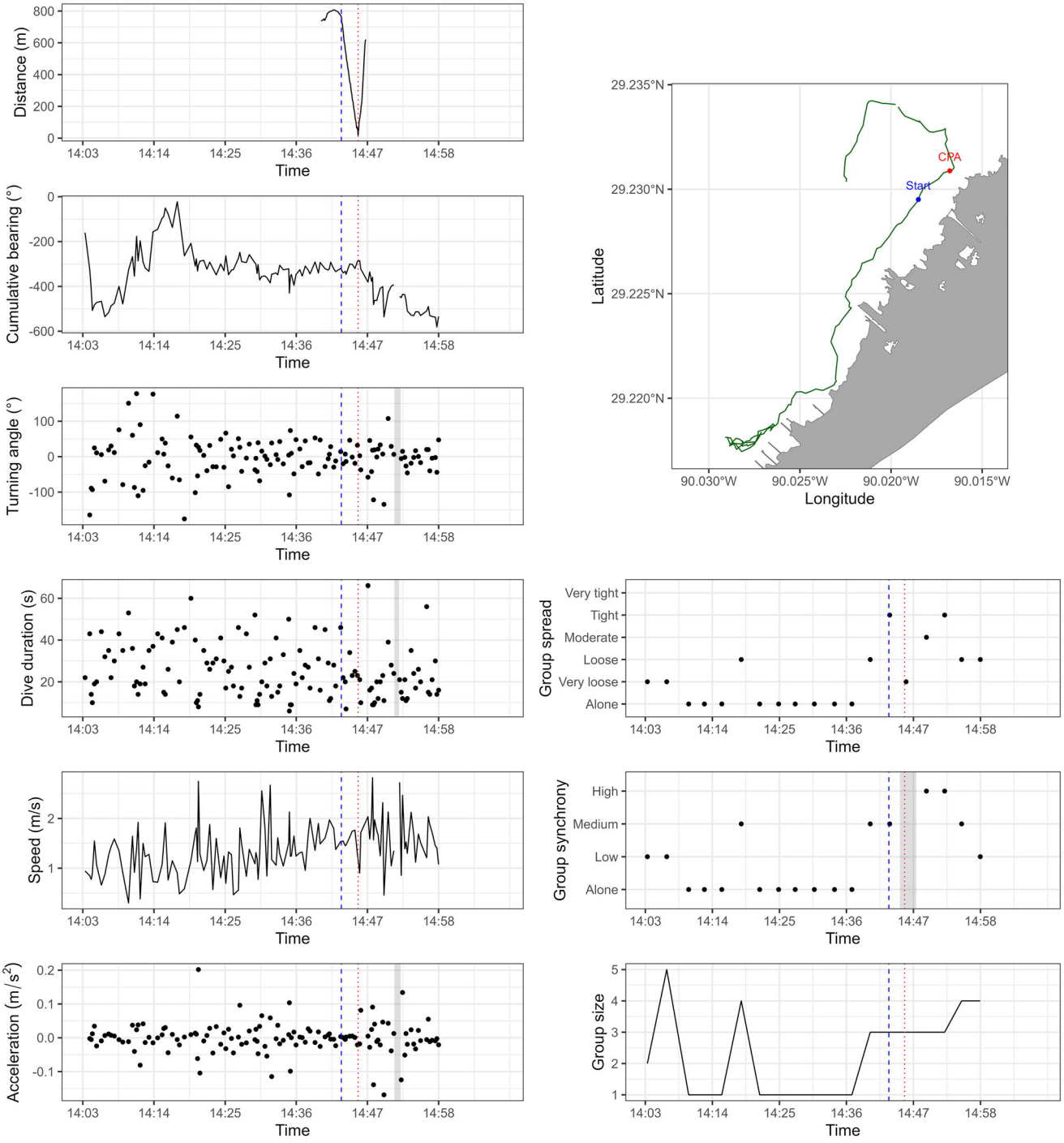


FIGURE 2 Example of time-series plots and map of an experiment in which a focal dolphin was exposed to a controlled vessel approach, as provided to experts for scoring the occurrence and phase of behavioral responses (dashed blue line, start of the approach; dotted red line, closest point of approach [CPA]). Periods with missing data or data scored as not applicable are indicated using gray shading in dot plots or gaps in the time series in line plots. The distance plot shows the distance between the dolphin and the approach vessel; all other plots report the behavioral metrics for the focal dolphin (described in the text).

The latent variable $y^*_{j,e}$ was mapped to the observed score $y_{j,e}$ as follows:

$$y_{j,e} = \begin{cases} 1, & \text{if } y^*_{j,e} \le \theta_1 \\ k, & \text{if } \theta_{k-1} \le y^*_{j,e} \le \theta_k \text{for } k = 2,3,4,5 \\ 6, & \text{if } \theta_5 \le y^*_{j,e} \end{cases} \quad (2)$$

where θ is a vector of thresholds with property $\theta_1 < \theta_2 < \ldots < \theta_5$, which divide the real line in 6 segments corresponding to the 6 response categories. If we define $\theta_0 = -\infty$ and $\theta_6 = +\infty$, the conditional distribution of $y_{i,e}$ can then be written as:

$$\Pr\left(y_{j,e} = k | b_i, j, e\right) = \Pr(\theta_{k-1} \le y^*_{j,e} \le \theta_k | b_i, j, e)$$

15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224, Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

$$= \Phi\left(\frac{\theta_k - \mu_{i,e}}{\psi_i}\right) - \Phi\left(\frac{\theta_{k-1} - \mu_{i,e}}{\psi_i}\right) \quad (3)$$

where Φ is the cumulative distribution function of the standard normal distribution. Because the models for the mean and standard deviation of the latent variable included an intercept (α_1 and α_2), we had to arbitrarily fix $\theta_1 = -3$ and $\theta_5 = 3$ to ensure identifiability (see discussion in Hirk et al. [2020]). We then defined the other thresholds as $\theta_k = \theta_{k-1} + \delta_k$ for $k = 2, 3, 4$ (see priors and constraints on δ_k in Appendix S6).

We fitted the models in a Bayesian framework with Markov chain Monte Carlo algorithms in NIMBLE (de Valpine et al., 2017) for R (details in Appendix S6).

A priori hypotheses for the effects of health on response probabilities

We hypothesized that health status modifies the phase of the CEE at which a dolphin tends to respond (depending on the metric, responding either earlier or later). Before running the analyses, we developed a series of hypotheses for how different health conditions (lung disease, HPA dysfunction, 1- and 2-year VESOP scores) might affect an animal's ability or propensity to respond. The full list of hypotheses for each of the 24 combinations of 6 response and 4 health metrics is in Appendix S4. In the Appendix S4 list, we do not distinguish between 1- and 2-year VESOP scores and instead refer to animals that were healthy or unhealthy overall. In 2 combinations, there was no a priori expectation of any difference between healthy and unhealthy individuals, so there were 22 hypothesized effects. Once analyses were run, we judged a hypothesis for a given combination of response and health metrics as being supported if the posterior 95% credible interval (CI) on the effect of the health metric did not include 0 and the posterior mean was in the predicted direction. We also highlighted model results where the posterior distribution indicated that there was a probability greater than 0.8 that the effect of the health metric was different from 0 (irrespective of the predicted direction). We acknowledge that, given multiple hypotheses were evaluated, there was an increased risk of making incorrect conclusions, and we interpreted results with appropriate caution.

We undertook power analyses to investigate the effect size that would be required to detect an effect of health status if it was there, given our sample size and the estimated random effect values. We also explored the sample size needed to consistently detect the mean estimated effect size, as well as the influence of different levels of individual dolphin replication across multiple CEEs. Details are in Appendix S8.

Effects of contextual variables on the probability of response

Following the assessment of the a priori hypotheses, we also evaluated the effects of a series of contextual variables on the threshold of response for an individual dolphin in each of the response metrics. These analyses were developed after seeing the experimental data and should be viewed as a way to generate hypotheses to test via future data collection. The contextual variables considered were group size at the start of the PWC approach (continuous), distance between the dolphin and the PWC at the start of the approach (continuous), sex (categorical: male or female), whether the animal was tagged with a Dtag at release from the health assessment (categorical: yes or no), whether the animal had been health assessed in the same sampling period (categorical: no, on the same day as the CEE, or on a different day than the CEE), whether the CEE took place in a marked navigation channel (categorical: yes or no), behavioral state of the individual or group immediately prior to the start of the approach (categorical: milling, slow travel, medium-fast travel, probable feeding), age class at the time of the CEE (categorical: subadult, adult, or older, corresponding to <9 years old, ≥9 and <25 years old, and ≥25 years old, respectively), previous fishing line entanglement assessed from scarring (categorical: yes or no), previous vessel strike assessed from scarring (categorical: yes or no), and whether the focal animal was a female with a dependent calf (categorical: female without calf, female with dependent calf, or male). When a contextual variable showed an effect on the threshold of response for any of the response metrics, we evaluated the potential interactions between that contextual variable and the 4 health metrics.

RESULTS

We carried out 13 successful CEEs in 2022 and 33 in 2023, for a total of 46 experimental approaches on 41 different subjects (5 animals were targeted twice) (Figure 3). Four additional CEEs were discarded because of malfunctioning of the instrumentation in the field or due to the large amount of missing behavioral information. Of the 13 CEEs in 2022, 5 were on individuals that underwent a health assessment in 2018, 3 in 2017, 2 in 2016, 1 in 2014, and 1 in 2011. One CEE was on an individual that subsequently had its health assessed in 2023, and its most recent health metrics (from 2023) were used. Of the 33 CEEs in 2023, 25 were on individuals health assessed that year, 4 in 2018, and 4 in 2017. The duration of the vessel approaches (from start to CPA) varied from 1.6 to 6.1 min (median 3 min).

Fifteen individuals had lung disease, and 20 did not. There was no information on lung condition for 6 individuals. Thirteen individuals showed evidence of HPA dysfunction, and 28 showed a normal cortisol response during handling. Five individuals showed evidence of both lung disease and HPA dysfunction. The 1-year VESOP score varied from 0.554 to 0.986 (10th and 90th quantiles: 0.789 and 0.980), whereas the 2-year VESOP score varied from 0.269 to 0.976 (10th and 90th quantiles: 0.671 and 0.976). Five individuals had a score lower than the 10th quantile (for both metrics). We used the 10th and 90th quantiles to indicate compromised versus good health when plotting the results below.

Across models, some dolphins tended to respond earlier or later in the approach, as suggested by the associated random effect (Appendix S7), although these differences were largely

15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224, Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

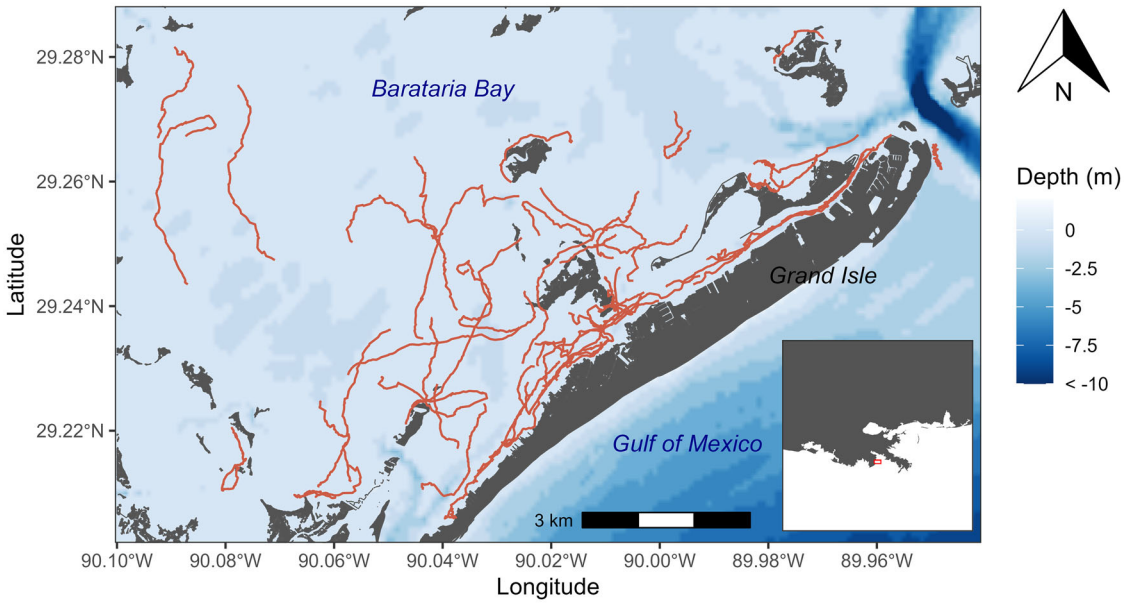


FIGURE 3 Study area in Barataria Bay, Louisiana (USA), and focal dolphin tracks across the 46 controlled exposure experiments (CEEs) (red). The maximum depth recorded from the observation vessel during CEEs carried out in the marked navigation channel northwest of Grand Isle range from 4.1 to 6.5 m, although this is not visible in the bathymetry data.

TABLE 1 For dolphins in Barataria Bay, Louisiana exposed to controlled vessel approaches, probabilities that the effect of the health metrics (i.e., health conditions and veterinary expert system for outcome prediction [VESOP] score) is different from 0 for all combinations of responses and health metrics.

Response metric	Health metric			
	Lung disease	HPA[a] dysfunction	1-year VESOP	2-year VESOP
Movement directionality	0.66	0.83	0.78[b]	0.75[b]
Speed/acceleration	0.70	0.77	0.71[b]	0.72[b]
Dive duration	0.96	0.58[c]	0.67[b]	0.67[b]
Group metrics	0.86	0.66	0.85	0.87
Orienting response	0.51[c]	0.70	0.73[b]	0.70[b]
Horizontal movement away	0.63	0.84	0.91	0.91

[a]Hypothalamus–pituitary–adrenal.
[b]Posterior mean sign opposite to the a priori hypothesis.
[c]No a priori expectation of any difference between healthy and unhealthy individuals.

inconsistent across different behavioral metrics. Experts also had a different tendency of scoring responses (Appendix S7). The 95% CI of the effect of health overlapped 0 across all 22 combinations of response and health metrics where we hypothesized an effect. However, in 8 combinations, the probability that the effect of the health metric was different from 0 was greater than 0.8 (Table 1; Figure 4). Specifically, there was a 0.83 probability that cortisol-flagged individuals showed a change in movement directionality later in the CEE compared with non-flagged individuals. There was also a 0.96 probability that lung disease resulted in a reduced or delayed dive response. We found a 0.85 and a 0.87 probability of an earlier change in group metrics when the focal animal had poorer 1- or 2-year VESOP scores, respectively, and a coherent effect of lung condition on group metrics with a posterior probability of 0.86. Finally, there was a 0.91 probability that animals with lower 1- and 2-year VESOP scores showed earlier horizontal movement away from the PWC, and there was a 0.84 probability that cortisol-flagged individuals responded less or later in this metric.

The power analysis for a subset of behavioral and health metric combinations provided an indication of the size of the effect of health on the phase in which a response occurred that would have been required to consistently detect such an effect in our analyses with the realized sample size (Appendix S8). In general, there was a low proportion of simulations in which a response was reliably detected given the effect sizes estimated in our analysis. This suggested that the true effect sizes may be too small to ensure consistent detection, in light of our sample size and the observed degree of variation in responsiveness within and between individuals. Given the estimated mean size

15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224 Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

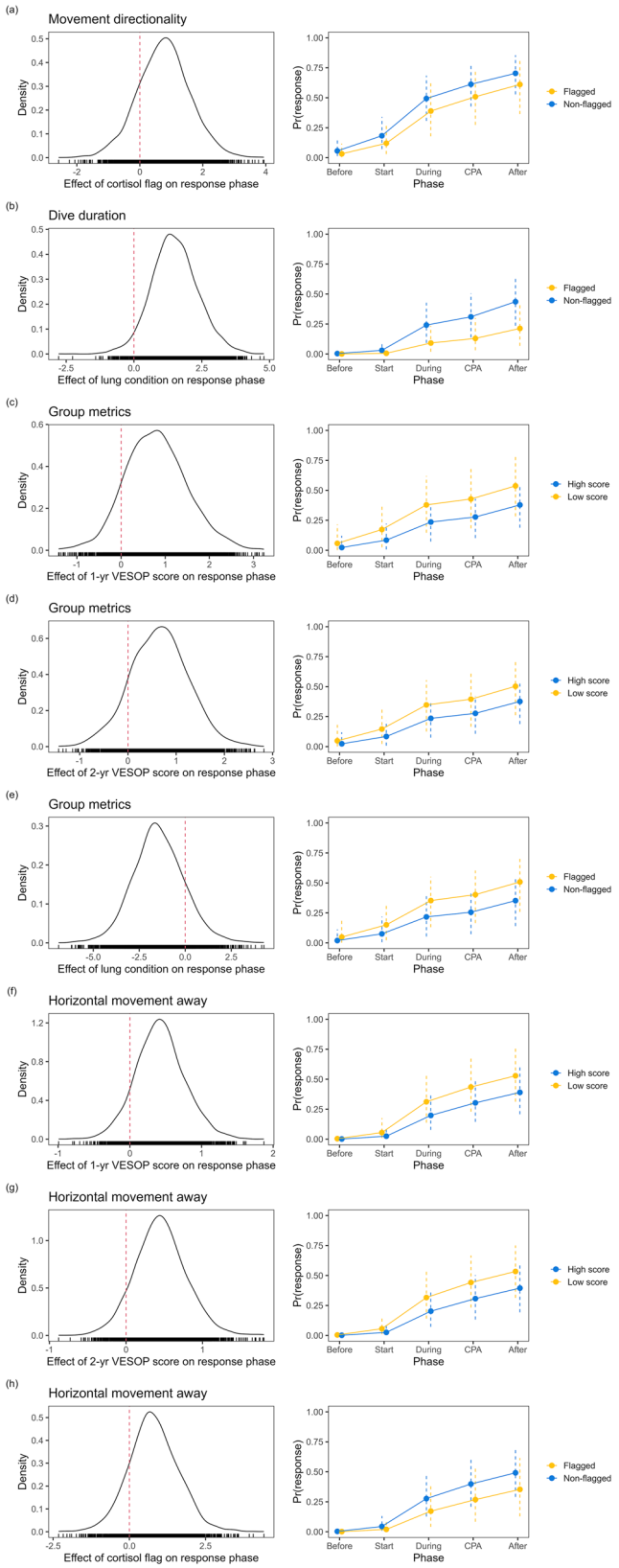


FIGURE 4 Posterior density distribution of the effect of a health metric (lung condition, veterinary expert system for outcome prediction [VESOP] score, and hypothalamus–pituitary–adrenal dysfunction [i.e., cortisol flag]) on the underlying probability of exhibiting a behavioral response (i.e., parameter β_3) (left column) and cumulative probability of responding over the course of the controlled exposure experiment for dolphins in different health states (right column). Results are reported for the 8 combinations of response and health metrics for which the probability that the effect of the health metric was different from 0 was greater than 0.8: (a) movement directionality and cortisol flag, (b) dive duration and lung disease, (c) group metrics and 1-year VESOP score, (d) group metrics and 2-year VESOP score, (e) group metrics and lung disease, (f) horizontal movement away and 1-year VESOP score, (g) horizontal movement away and 2-year VESOP score, and (h) horizontal movement away and cortisol flag. Low versus high scores are defined as the 10th and 90th quantiles of the 1- and 2-year VESOP scores (right column graphs: dots, medians; dashed lines, 95% credible intervals).

(Continues)

of the effects of the health metrics on the threshold of response, between 250 and >500 CEEs may be required for consistent detection (Appendix S8). We also found that targeting new individuals for each CEE granted more power than repeating the approach on the same individuals over multiple experiments (Appendix S8).

Full results of the ancillary analyses for the effects of contextual variables are in Appendices S10 and S11; findings are summarized here. The results suggested that distance between the focal dolphin and the PWC at the start of the approach was associated with the probability of showing responses, responding either earlier or later depending on the behavioral metric. Moreover, males tended to respond more or earlier across several metrics, whereas individuals that were Dtagged or health assessed on the same day as the CEE tended to respond less or later in the approach. Responses to CEEs that were carried out in a marked channel occurred earlier for some metrics. Scarring from entanglements was associated with some earlier responses, whereas there was a weak indication that scarring from vessel strikes was related to a delayed response in some metrics. Finally, older animals showed some delayed responses, and having a dependent calf altered responses in focal females, but the direction varied by metric. For some combinations of metrics, these contextual variables interacted with the effect of the health metrics on the threshold of response (Appendix S11), but the small sample size when splitting the data into different contexts suggests caution in the interpretation of the emerging dose–response relationships. Moreover, some of these effects could be confounded among contextual variables because there was overlap between sets of individuals across levels of these variables (Appendix S10).

DISCUSSION

We developed an experimental approach to assess the combined effects of multiple stressors in a large, free-ranging vertebrate species. This approach used health as an integrator of the effects of an individual's prior exposure to stressors and evaluated its influence on responses to subsequent stressors. Even though our results aligned with the a priori hypotheses for some combinations of behavioral and health metrics, the analysis of dolphin behavior during controlled PWC approaches did not highlight a strong effect of health status on the probabil-

15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224, Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ity of response under the given experimental conditions. We demonstrated the importance of several contextual variables in affecting dolphin response thresholds across metrics and their potential interactions with the effects of health.

Lung disease was associated with a delayed or reduced probability of changing dive duration, which aligns with the mechanistic hypothesis that lung condition may affect the ability of an individual to dive effectively (Smith et al., 2022) and thus vertically avoid approaching vessels (Lusseau, 2003; Nowacek et al., 2001). This result provides an example of how poor health status may limit the response options available to an individual or alter their effectiveness, which could ultimately lead to the adoption of more costly or risky behaviors (Bejder et al., 2009; Burslem et al., 2022). We also found some indication that HPA dysfunction, as evidenced by a low cortisol response, was associated with a delayed or reduced tendency to change directionality and move away from the vessel. If an animal is unable to recruit a stress response promptly by secreting adrenal hormones, this can result in reduced responsiveness to a threat (Rosol et al., 2001). There was also some evidence of this effect when considering an individual's overall health score. Finally, lung disease and the overall health score were weakly associated with an earlier change in group metrics. It is difficult to predict how the health status of the focal individual may affect the behavior of nearby dolphins and vice versa. However, in a highly social species, it is reasonable to assume that individual and group behavior may be linked and that the health of one individual may influence the decision-making of other group members (Caldwell & Caldwell, 1966). Altogether, these changes in the propensity or ability to respond to the PWC suggest that individuals in poorer health may experience an increased probability of getting struck by vessels if their evasive tactics become less effective.

Despite the results for this subset of behavioral and health metric combinations, the effects of health on the probability of responding to the vessel were generally smaller than anticipated. Even though using an experimental approach allowed us to control for some aspects of experimental context (e.g., we could use a broadly repeatable stimulus across experiments), this uncertainty may result from the confounding effects of other variables. For example, males were generally more responsive than females, which aligns with findings across taxa (Dantzer et al., 2014), including cetaceans (Lusseau, 2003; Pirotta et al., 2023). Having a dependent calf also appeared to affect the type and timing of responses, conforming with the expected influence of reproductive state on vigilance and risk tolerance in mammals (e.g., Burger & Gochfeld, 1994; Marneweck et al., 2021). In addition, the context of the exposure (e.g., distance of the PWC at the start of the approach, and presence of topographical features that may alter the subject's behavioral options) appeared to influence dolphin responsiveness, and we hypothesize that other differences among CEEs may have also contributed (e.g., variation in approach angle, bathymetry, occurrence of other boats in the area, among others) (Tablado & Jenni, 2017). Importantly, previous experiences of an individual with stressors (as evidenced by scarring from entanglements and vessel collisions) and recent capture and handling during the health assessment and tagging procedure affected the response threshold (Tablado & Jenni, 2017), which can inform the future design of similar exposure experiments (e.g., allowing for longer time between handling and exposure). Some of these contextual variables interacted with health and thus could mask the effects of health when not accounted for, but our sample size was too small to partition the CEEs across many combinations of context and health conditions. In general, our results emphasize that the influence of internal and external contextual variables in many cases may exceed the effect of health status on behavioral responsiveness and, ultimately, collision risk.

Aside from the effect of context, the power analyses suggested that the effects of the health metrics on responsiveness in our experimental setting were likely small. Given our sample size, we would have been able to detect stronger effects, but our power to detect the estimated mean effect sizes was limited overall. As a result, we estimated that hundreds of CEEs would be required for reliable detection of these subtle effects, which is relevant for the design of future studies in comparable systems. For safety reasons, we used a PWC to approach focal dolphins during the experiments, but the noise generated by the jet drive of the vessel we used differed from that of a boat with a propeller and may have resulted in weaker responses. The low temporal resolution of surfacing data and inability to observe the animals under water could have also contributed to the low power to detect responses and the associated effects of health on behavior measured at this coarse scale. The murkiness of the water in Barataria Bay prevented us from using drone footage to characterize behavior of the focal animal between surfacings (e.g., Hao et al., 2024). In contrast, we deployed high-resolution sound and movement recording tags (Dtags) that could illuminate additional responses that may not have been observable in the surface data (Johnson & Tyack, 2003), but only 9 tagged individuals were the subject of successful CEEs. The subtle nature of the behavioral changes that could be detected in the surfacing data forced us to turn to expert scoring to mark the occurrence and timing of responses, which could also have introduced some level of subjectivity. Our analytical approach for these data was based on the assumption that the no-response category corresponds to a CEE that did not achieve a sufficient intensity to elicit a response by the focal dolphin (similar to Miller et al. [2014]). In some CEEs, some experts scored the occurrence of an early response and others scored a no response, which questions this assumption. These instances could correspond to individuals that showed early, subtle responses, and in our analyses, they could be partly captured via the estimation of a large standard deviation in an individual's response threshold. In general, increased individual replication may be needed to appropriately apportion the variation in responses to the different random effects in the model. In particular, we assumed that the variation in scores among experts reflected the clarity of a dolphin's response and that the model could estimate between-dolphin differences in response threshold and consistency. However, our explorations (Appendix S9) suggested that variation among dolphins was

15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224, Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

potentially conflated with differences between CEEs because of the small number of individuals that were targeted twice.

Our study has broader implications for the assessment of the combined effects of multiple stressors in a large-vertebrate system. We demonstrated the value of using an animal's health status as the integrator of its prior stressor exposure (Kophamel et al., 2022; National Academies, 2017; Tyack et al., 2022), which makes it possible to test for combined effects without having to manipulate the doses of multiple stressors concurrently (e.g., Maurer et al., 2025). However, health may also change over time and be influenced by other intervening and undocumented stressors (Appendix S12), requiring intensive monitoring of individuals in a population. In many marine vertebrate species, routine health assessments and the collection of detailed longitudinal data are expensive, logistically challenging, or outright infeasible, but new technology is improving the capacity to monitor wildlife health (e.g., Barratclough et al., 2024; Hunt et al., 2013; Kophamel et al., 2022).

We also found that, despite the advantages granted by an experimental approach, many contextual variables did influence the results but could not be controlled. Any large-scale experiment on wildlife species will likely encounter similar difficulties because animal responses to external sources of disturbance are highly context dependent (Beale, 2007; Ellison et al., 2012; Madliger & Love, 2014; Moran et al., 2021; Tablado & Jenni, 2017). Behavioral changes, in particular, result from complex processes that involve the detection of a disturbance stimulus and a subsequent trade-off between perceived costs and benefits of different responses (Sih, 2013). Experiments in the wild can be designed to cover a range of realistic contexts that match the complexity of the environment relevant for management, but sufficient replication will continue to be difficult. Further understanding of the mechanisms that drive an animal's decision making can help guide the development of analytical approaches to tackle this complexity (Pirotta et al., 2022). In general, context dependency will challenge the estimation of multidimensional dose–response functions in the case of most stressor mixtures, but we support the conclusion that being able to incorporate both empirical and mechanistic knowledge makes this framework particularly effective for studying combined effects (Rosenfeld et al., 2022; Segurado et al., 2022; van Moorsel et al., 2023). We also note that multidimensional dose–response functions can help quantify the cumulative risk to a target population irrespective of whether the interactions among stressors are categorized as additive, synergistic, or antagonistic (Pirotta et al., 2022).

More than a decade after the *Deepwater Horizon* oil spill, this catastrophic event continues to have cascading effects for the population of bottlenose dolphins in Barataria Bay (Smith et al., 2017, 2022). Our study provides some evidence that, in addition to their degraded health status, individuals in this population may be at greater risk from other stressors due to their reduced ability to elicit and engage in a behavioral response (e.g., the effect of lung disease on diving abilities). However, experimental results also suggested that internal and external contextual variables play a key role in driving individual responses. Ultimately, estimating the combined effects of vessel interactions and protracted poor health following the oil spill is the first step toward the development of effective management solutions to ensure the viability of affected coastal populations (Tyack et al., 2022). For example, the modified dose–response functions characterizing responses to vessels (and thus collision risk) under different health status could inform simulations to identify which stressors (among the ones that can be reduced or mitigated) should be the focus of ongoing restoration efforts (e.g., Runge et al., 2023). The quantification and management of the cumulative risk from multiple stressors on wildlife populations will thus require a combination of empirical and mechanistic approaches, which could involve targeted experiments, such as the one we presented here, to address specific linkages in the cascade between stressor exposure and consequences for population dynamics. Communication with managers could guide the design of experiments with sufficient power to assess whether the combined effects of 2 stressors exceed management goals for a population of interest.

AUTHOR CONTRIBUTIONS

Lori H. Schwacke, Peter L. Tyack, Randall S. Wells, and Len Thomas originally conceived the study. All authors contributed to the development of the experimental design and field protocols. Lori H. Schwacke, Jason B. Allen, Aaron A. Barleycorn, Marco Casoli, Jonathan Crossman, Katherine A. McHugh, Michael J. Moore, Brian M. Quigley, Cynthia R. Smith, Todd R. Speakman, Ryan Takeshita, Christina N. Toms, Randall S. Wells, Eric S. Zolman, and Peter L. Tyack collected the data. Enrico Pirotta, Marco Casoli, Katherine A. McHugh, Christina N. Toms, and Theoni Photopoulou processed the field data. Ryan Takeshita, Cynthia R. Smith, and Lori H. Schwacke provided health data and ancillary dolphin information. Eric S. Zolman, Todd R. Speakman, and Ryan Takeshita processed scarring data. Catriona M. Harris developed the list of hypotheses to test with support from Peter L. Tyack and Enrico Pirotta. Catriona M. Harris organized and led the expert scoring with support from Enrico Pirotta, Patrick J. O. Miller, and Brandon L. Southall. Jason B. Allen, Erin LaBrecque, Laura Engleby, Laura Aichinger Dias, Jenny Litz, Patrick J. O. Miller, Brandon L. Southall, and Eric S. Zolman conducted the expert scoring. Enrico Pirotta and Len Thomas developed the analytical approach. Enrico Pirotta ran the analyses. Enrico Pirotta led the writing of the manuscript. All authors contributed critically to the drafts and gave approval for publication.

ACKNOWLEDGMENTS

This work was supported by the Office of Naval Research (N000142012697, N000142112096) and the Strategic Environmental Research and Development Program (RC20-1097, RC20-7188, RC21-3091). The Marine Mammal Commission also provided financial support for field efforts and analyses. We thank the PCoMS working group for guidance; B. Jones, C. Thompson, A. Allen, J. Shearer, and R. Martinson for help with data collection; Louisiana Department of Wildlife and Fisheries and numerous other organizations and collaborators for

15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224 Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

support of field operations; Brookfield Zoo Chicago's Sarasota Dolphin Research Program for help set up controlled approach protocols; T. Spradlin for contributing to the expert scoring; and the health assessment team, including F. Gomez, F. Townsend, A. Barratclough, B. Bauer, K. Christman, A. Cheever, V. Cendejas, J. Eichman, W. Musser, and K. Ross.

DATA AVAILABILITY STATEMENT

All data and code to run these analyses are available from the Open Science Framework (https://osf.io/kx7sj).

ORCID

Enrico Pirotta https://orcid.org/0000-0003-3541-3676
Catriona M. Harris https://orcid.org/0000-0001-9198-2414
Len Thomas https://orcid.org/0000-0002-7436-067X
Laura Aichinger Dias https://orcid.org/0000-0001-9378-4577
Jenny A. Litz https://orcid.org/0000-0002-7211-1063
Peter L. Tyack https://orcid.org/0000-0002-8409-4790

REFERENCES

Barratclough, A., Takeshita, R., Thomas, L., Photopoulou, T., Pirotta, E., Rosel, P. E., Wilcox Talbot, L. A., Vollmer, N. L., Wells, R., Smith, C. R., Rowles, T. K., Horvath, S., & Schwacke, L. (2024). Estimating age and investigating epigenetic changes related to health across multiple bottlenose dolphin populations. *Biological Conservation*, *293*, Article 110570.

Barratclough, A., Wells, R. S., Schwacke, L. H., Rowles, T. K., Gomez, F. M., Fauquier, D. A., Sweeney, J. C., Townsend, F. I., Hansen, L. J., Zolman, E. S., Balmer, B. C., & Smith, C. R. (2019). Health assessments of common bottlenose dolphins (*Tursiops truncatus*): Past, present, and potential conservation applications. *Frontiers in Veterinary Science*, *6*, Article 444.

Beale, C. M. (2007). The behavioral ecology of disturbance responses. *International Journal of Comparative Psychology*, *20*, 111–120.

Bejder, L., Samuels, A., Whitehead, H., Finn, H., & Allen, S. (2009). Impact assessment research: Use and misuse of habituation, sensitisation and tolerance in describing wildlife responses to anthropogenic stimuli. *Marine Ecology Progress Series*, *395*, 177–185.

Boyd, P. W., Collins, S., Dupont, S., Fabricius, K., Gattuso, J. P., Havenhand, J., Hutchins, D. A., Riebesell, U., Rintoul, M. S., Vichi, M., Biswas, H., Ciotti, A., Gao, K., Gehlen, M., Hurd, C. L., Kurihara, H., McGraw, C. M., Navarro, J. M., Nilsson, G. E., … Pörtner, H. O. (2018). Experimental strategies to assess the biological ramifications of multiple drivers of global ocean change—A review. *Global Change Biology*, *24*, 2239–2261.

Buckstaff, K. (2004). Effects of watercraft noise on the acoustic behavior of bottlenose dolphins, *Tursiops truncatus*, in Sarasota Bay, Florida. *Marine Mammal Science*, *20*, 709–725.

Burger, J., & Gochfeld, M. (1994). Vigilance in African mammals: Differences among mothers, other females, and males. *Behaviour*, *131*, 153–169.

Burslem, A., Isojunno, S., Pirotta, E., & Miller, P. J. O. (2022). Modelling the impact of condition-dependent responses and lipid-store availability on the consequences of disturbance in a cetacean. *Conservation Physiology*, *10*, Article coac069.

Calabrese, E. J. (2016). The emergence of the dose-response concept in biology and medicine. *International Journal of Molecular Sciences*, *17*, Article 2034.

Caldwell, M. C., & Caldwell, D. K. (1966). Epimeletic (care-giving) behavior in Cetacea. In K. S. Norris (Ed.), *Whales, dolphins, and porpoises* (pp. 755–789). University of California Press.

Carmichael, H., Warfield, R., & Yvon-Durocher, G. (2025). Reconciling variability in multiple stressor effects using environmental performance curves. *Ecology Letters*, *28*, Article e70065.

Casoli, M., Johnson, M., McHugh, K. A., Wells, R. S., & Tyack, P. L. (2022). Parameterizing animal sounds and motion with animal-attached tags to study acoustic communication. *Behavioral Ecology and Sociobiology*, *76*, Article 59.

Crain, C. M., Kroeker, K., & Halpern, B. S. (2008). Interactive and cumulative effects of multiple human stressors in marine systems. *Ecology Letters*, *11*, 1304–1315.

Dantzer, B., Fletcher, Q. E., Boonstra, R., & Sheriff, M. J. (2014). Measures of physiological stress: A transparent or opaque window into the status, management and conservation of species? *Conservation Physiology*, *2*, Article cou023.

de Valpine, P., Turek, D., Paciorek, C. J., Anderson-Bergman, C., Temple Lang, D., & Bodik, R. (2017). Programming with models: Writing statistical algorithms for general model structures with NIMBLE. *Journal of Computational and Graphical Statistics*, *26*, 403–413.

Ellison, W. T., Southall, B. L., Clark, C. W., & Frankel, A. S. (2012). A new context-based approach to assess marine mammal behavioral responses to anthropogenic sounds. *Conservation Biology*, *26*, 21–28.

Fordyce, E., & Perrin, W. F. (2025). *World Cetacea Database. Tursiops truncatus truncatus Hershkovitz, 1966*. World Register of Marine Species. https://www.marinespecies.org/aphia.php?p=taxdetails&id=384565

Gennings, C., Carter, W. H., Carchman, R. A., Teuschler, L. K., Simmons, J. E., & Carney, E. W. (2005). A unifying concept for assessing toxicological interactions: Changes in slope. *Toxicological Sciences*, *88*, 287–297.

Griffen, B. D., Belgrad, B. A., Cannizzo, Z. J., Knotts, E. R., & Hancock, E. R. (2016). Rethinking our approach to multiple stressor studies in marine environments. *Marine Ecology Progress Series*, *543*, 273–281.

Hao, X., Hamel, H., Grandjean, C. H., Fedutin, I., Wahlberg, M., Frankish, C. K., & Nabe-Nielsen, J. (2024). Harbour porpoises respond to recreational boats by speeding up and moving away from the boat path. *Ecology and Evolution*, *14*, Article e11433.

Harris, C. M., Thomas, L., Falcone, E. A., Hildebrand, J., Houser, D., Kvadsheim, P. H., Lam, F.-P. A., Miller, P. J. O., Moretti, D. J., Read, A. J., Slabbekoorn, H., Southall, B. L., Tyack, P. L., Wartzok, D., & Janik, V. M. (2018). Marine mammals and sonar: Dose-response studies, the risk-disturbance hypothesis and the role of exposure context. *Journal of Applied Ecology*, *55*, 396–404.

Hayes, S., Josephson, E., Maze-Foley, K., Rosel, P., & McCordic, J. (2024). *US Atlantic and Gulf of Mexico Marine Mammal Stock Assessments 2023* (Technical Memorandum NMFS-NE-321). National Oceanic and Atmospheric Administration.

Hijmans, R. J. (2016). *geosphere: Spherical trigonometry*. R package version 1.5-5. https://CRAN.R-project.org/package=geosphere

Hirk, R., Hornik, K., & Vana, L. (2020). mvord: An R package for fitting multivariate ordinal regression models. *Journal of Statistical Software*, *93*, 1–41.

Hunt, K. E., Moore, M. J., Rolland, R. M., Kellar, N. M., Hall, A. J., Kershaw, J., Raverty, S. A., Davis, C. E., Yeates, L. C., Fauquier, D. A., Rowles, T. K., & Kraus, S. D. (2013). Overcoming the challenges of studying conservation physiology in large whales: A review of available methods. *Conservation Physiology*, *1*, Article cot006.

Johnson, M. P., & Tyack, P. L. (2003). A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE Journal of Oceanic Engineering*, *28*, 3–12.

Kophamel, S., Illing, B., Ariel, E., Difalco, M., Skerratt, L. F., Hamann, M., Ward, L. C., Méndez, D., & Munns, S. L. (2022). Importance of health assessments for conservation in noncaptive wildlife. *Conservation Biology*, *36*, Article e13724.

Lusseau, D. (2003). Male and female bottlenose dolphins *Tursiops spp*. have different strategies to avoid interactions with tour boats in Doubtful Sound, New Zealand. *Marine Ecology Progress Series*, *257*, 267–274.

Madliger, C. L., & Love, O. P. (2014). The need for a predictive, context-dependent approach to the application of stress hormones in conservation. *Conservation Biology*, *28*, 283–287.

Marneweck, C. J., Louis Van Schalkwyk, O., Marneweck, D. G., Beverley, G., Davies-Mostert, H. T., & Parker, D. M. (2021). Reproductive state influences the degree of risk tolerance for a seasonally breeding mesopredator. *Behavioral Ecology*, *32*, 717–727.

Maurer, N., Boyi, J. O., Schick, L. A., Nachtsheim, D. A., Schaffeld, T., Gross, S., Teilmann, J., Schnitzler, J., & Siebert, U. (2025). Unsealing behaviour: Variation in harbour seal (*Phoca vitulina*) responses to anthropogenic sound in relation to individual health. *Marine Pollution Bulletin*, *214*, Article 117777.

15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224, Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

McDonald, T. L., Hornsby, F. E., Speakman, T. R., Zolman, E. S., Mullin, K. D., Sinclair, C., Rosel, P. E., Thomas, L., & Schwacke, L. H. (2017). Survival, density, and abundance of common bottlenose dolphins in Barataria Bay (USA) following the *Deepwater Horizon* oil spill. *Endangered Species Research*, *33*, 193–209.

Miller, P. J. O., Antunes, R. N., Wensveen, P. J., Samarra, F. I. P., Alves, A. C., Tyack, P. L., Kvadsheim, P. H., Kleivane, L., Lam, F.-P. A., Ainslie, M. A., & Thomas, L. (2014). Dose-response relationships for the onset of avoidance of sonar by free-ranging killer whales. *The Journal of the Acoustical Society of America*, *135*, 975–993.

Miller, P. J. O., Kvadsheim, P. H., Lam, F. P. A., Wensveen, P. J., Antunes, R., Alves, A. C., Visser, F., Kleivane, L., Tyack, P. L., & Sivle, L. D. (2012). The severity of behavioral changes observed during experimental exposures of killer (*Orcinus orca*), long-finned pilot (*Globicephala melas*), and sperm (*Physeter macrocephalus*) whales to naval sonar. *Aquatic Mammals*, *38*, 362–401.

Moran, N. P., Sánchez-Tójar, A., Schielzeth, H., & Reinhold, K. (2021). Poor nutritional condition promotes high-risk behaviours: A systematic review and meta-analysis. *Biological Reviews*, *96*, 269–288.

National Academies. (2017). *Approaches to understanding the cumulative effects of stressors on marine mammals*. The National Academies Press. http://www.nap.edu/catalog/23479

Nowacek, S. M., Wells, R. S., & Solow, A. R. (2001). Short-term effects of boat traffic on bottlenose dolphins, *Tursiops truncatus*, in Sarasota Bay, Florida. *Marine Mammal Science*, *17*, 673–688.

Orr, J. A., Vinebrooke, R. D., Jackson, M. C., Kroeker, K. J., Kordas, R. L., Mantyka-Pringle, C., van den Brink, P. J., de Laender, F., Stoks, R., Holmstrup, M., Matthaei, C. D., Monk, W. A., Penk, M. R., Leuzinger, S., Schäfer, R. B., & Piggott, J. J. (2020). Towards a unified study of multiple stressors: Divisions and common goals across research disciplines. *Proceedings of the Royal Society B: Biological Sciences*, *287*, Article 20200421.

Pirotta, E., Fernandez Ajó, A., Bierlich, K., Bird, C. N., Buck, C. L., Haver, S. M., Haxel, J. H., Hildebrand, L., Hunt, K. E., Lemos, L. S., New, L., & Torres, L. G. (2023). Assessing variation in faecal glucocorticoid concentrations in gray whales exposed to anthropogenic stressors. *Conservation Physiology*, *11*, Article coad082.

Pirotta, E., Merchant, N. D., Thompson, P. M., Barton, T. R., & Lusseau, D. (2015). Quantifying the effect of boat disturbance on bottlenose dolphin foraging activity. *Biological Conservation*, *181*, 82–89.

Pirotta, E., Thomas, L., Costa, D. P., Hall, A. J., Harris, C. M., Harwood, J., Kraus, S. D., Miller, P. J., Moore, M., Photopoulou, T., Rolland, R., Schwacke, L., Simmons, S. E., Southall, B. L., & Tyack, P. (2022). Understanding the combined effects of multiple stressors: A new perspective on a longstanding challenge. *Science of the Total Environment*, *821*, Article 153 322.

R Core Team. (2024). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Rosenfeld, J., Gonzalez-Espinosa, P., Jarvis, L., Enders, E., Bayly, M., Paul, A., MacPherson, L., Moore, J., Sullivan, M., Ulaski, M., & Wilson, K. (2022). Stressor-response functions as a generalizable model for context dependence. *Trends in Ecology & Evolution*, *37*, 1032–1035.

Rosol, T. J., Yarrington, J. T., Latendresse, J., & Capen, C. C. (2001). Adrenal gland: Structure, function, and mechanisms of toxicity. *Toxicologic Pathology*, *29*, 41–48.

Runge, M. C., Linden, D. W., Hostetler, J. A., Borggaard, D. L., Garrison, L. P., Knowlton, A. R., Lesage, V., Williams, R., & Pace, R. M. I. (2023). *A management-focused population viability analysis for North Atlantic right whales* (NOAA Technical Memorandum NMFS-NE-307). US Department of Commerce, Northeast Fisheries Science Centre. https://repository.library.noaa.gov/view/noaa/61490

Schwacke, L. H., Marques, T. A., Thomas, L., Booth, C. G., Balmer, B. C., Barratclough, A., Colegrove, K., De Guise, S., Garrison, L. P., Gomez, F. M., Morey, J. S., Mullin, K. D., Quigley, B. M., Rosel, P. E., Rowles, T. K., Takeshita, R., Townsend, F. I., Speakman, T. R., Wells, R. S., … Smith, C. R. (2022). Modeling population effects of the *Deepwater Horizon* oil spill on a long-lived species. *Conservation Biology*, *36*, Article e13878.

Schwacke, L. H., Smith, C. R., Townsend, F. I., Wells, R. S., Hart, L. B., Balmer, B. C., Collier, T. K., De Guise, S., Fry, M. M., Guillette, L. J., Lamb, S. V., Lane, S. M., McFee, W. E., Place, N. J., Tumlin, M. C., Ylitalo, G. M., Zolman, E. S., & Rowles, T. K. (2014). Health of common bottlenose dolphins (*Tursiops truncatus*) in Barataria Bay, Louisiana, following the *Deepwater Horizon* oil spill. *Environmental Science and Technology*, *48*, 93–103.

Schwacke, L. H., Thomas, L., Wells, R. S., McFee, W. E., Hohn, A. A., Mullin, K. D., Zolman, E. S., Quigley, B. M., Rowles, T. K., & Schwacke, J. H. (2017). Quantifying injury to common bottlenose dolphins from the *Deepwater Horizon* oil spill using an age-, sex- and class-structured population model. *Endangered Species Research*, *33*, 265–279.

Schwacke, L. H., Thomas, L., Wells, R. S., Rowles, T. K., Bossart, G., Townsend, F., Mazzoil, M., Allen, J. B., Balmer, B. C., Barleycorn, A. A., Barratclough, A., Burt, M. L., De Guise, S., Fauquier, D., Gomez, F. M., Kellar, N. M., Schwacke, J. H., Speakman, T. R., Stolen, E., … Smith, C. R. (2023). An expert-based system to predict population survival rate from health data. *Conservation Biology*, *38*, Article e14073.

Segurado, P., Gutiérrez-Cánovas, C., Ferreira, T., & Branco, P. (2022). Stressor gradient coverage affects interaction identification. *Ecological Modelling*, *472*, Article 110089.

Siegal, E., Hooker, S. K., Isojunno, S., & Miller, P. J. O. (2022). Beaked whales and state-dependent decision-making: How does body condition affect the trade-off between foraging and predator avoidance? *Proceedings of the Royal Society B: Biological Sciences*, *289*(1967), Article 20212539.

Sih, A. (2013). Understanding variation in behavioural responses to human-induced rapid environmental change: A conceptual overview. *Animal Behaviour*, *85*, 1077–1088.

Smith, C. R., Rowles, T. K., Gomez, F. M., Ivančić, M., Colegrove, K. M., Takeshita, R., Townsend, F. I., Zolman, E. S., Morey, J. S., Cendejas, V., Meegan, J. M., Musser, W., Speakman, T. R., Barratclough, A., Wells, R. S., & Schwacke, L. H. (2022). Poor pulmonary health in Barataria Bay dolphins in the eight years following the *Deepwater Horizon* oil spill. *Frontiers in Marine Science*, *9*, Article 975006.

Smith, C. R., Rowles, T. K., Hart, L. B., Townsend, F. I., Wells, R. S., Zolman, E. S., Balmer, B. C., Quigley, B., Ivančić, M., McKercher, W., Tumlin, M. C., Mullin, K. D., Adams, J. D., Wu, Q., McFee, W., Collier, T. K., & Schwacke, L. H. (2017). Slow recovery of Barataria Bay dolphin health following the *Deepwater Horizon* oil spill (2013-2014), with evidence of persistent lung disease and impaired stress response. *Endangered Species Research*, *33*, 127–142.

Southall, B., Bowles, A., Ellison, W., Finneran, J. J., Gentry, R. L., Greene, C. R. J., Kastak, D., Ketten, D. R., Miller, J. H., Nachtigall, P. E., Richardson, W. J., Thomas, J. A., & Tyack, P. L. (2007). Marine mammal noise-exposure criteria: Initial scientific recommendations. *Aquatic Mammals*, *33*(4), 411–521.

Southall, B. L., DeRuiter, S. L., Friedlaender, A., Stimpert, A. K., Goldbogen, J. A., Hazen, E., Casey, C., Fregosi, S., Cade, D. E., Allen, A. N., Harris, C. M., Schorr, G., Moretti, D., Guan, S., & Calambokidis, J. (2019). Behavioral responses of individual blue whales (*Balaenoptera musculus*) to mid-frequency military sonar. *Journal of Experimental Biology*, *222*, Article 190637.

Southall, B. L., Nowacek, D. P., Miller, P. J. O., & Tyack, P. L. (2016). Experimental field studies to measure behavioral responses of cetaceans to sonar. *Endangered Species Research*, *31*, 293–315.

Steffen, W., Persson, Å., Deutsch, L., Zalasiewicz, J., Williams, M., Richardson, K., Crumley, C., Crutzen, P., Folke, C., Gordon, L., Molina, M., Ramanathan, V., Rockström, J., Scheffer, M., Schellnhuber, H. J., & Svedin, U. (2011). The Anthropocene: From global change to planetary stewardship. *Ambio*, *40*, 739–761.

Tablado, Z., & Jenni, L. (2017). Determinants of uncertainty in wildlife responses to human disturbance. *Biological Reviews*, *92*, 216–233.

Tyack, P. L., Thomas, L., Costa, D. P., Hall, A. J., Harris, C. M., Harwood, J., Kraus, S. D., Miller, P. J., Moore, M., Photopoulou, T., Pirotta, E., Rolland, R. M., Schwacke, L. H., Simmons, S. E., & Southall, B. L. (2022). Managing the effects of multiple stressors on wildlife populations in their ecosystems: Developing a cumulative risk approach. *Proceedings of the Royal Society B: Biological Sciences*, *289*, Article 20222058.

van Moorsel, S. J., Thébault, E., Radchuk, V., Narwani, A., Montoya, J. M., Dakos, V., Holmes, M., De Laender, F., & Pennekamp, F. (2023). Predicting effects of multiple interacting global change drivers across trophic levels. *Global Change Biology*, *29*, 1223–1238.

Van Waerebeek, K., Baker, A. N., Félix, F., Gedamke, J., Iñiguez, M., Sanino, G. P., Secchi, E., Sutaria, D., Van Helden, A., & Wang, Y. (2007). Vessel colli-

15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224, Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

sions with small cetaceans worldwide and with large whales in the Southern Hemisphere, an initial assessment. *Latin American Journal of Aquatic Mammals*, *6*, 43–69.

Venn-Watson, S., Colegrove, K. M., Litz, J., Kinsel, M., Terio, K., Saliki, J., Fire, S., Carmichael, R., Chevis, C., Hatchett, W., Pitchford, J., Tumlin, M., Field, C., Smith, S., Ewing, R., Fauquier, D., Lovewell, G., Whitehead, H., Rotstein, D., … Rowles, T. (2015). Adrenal gland and lung lesions in Gulf of Mexico common bottlenose dolphins (*Tursiops truncatus*) found dead following the *Deepwater Horizon* oil spill. *PLoS ONE*, *10*, Article e0126538.

Wells, R. S., Allen, J. B., Hofmann, S., Bassos-Hull, K., Fauquier, D. A., Barros, N. B., DeLynn, R. E., Sutton, G., Socha, V., & Scott, M. D. (2008). Consequences of injuries on survival and reproduction of common bottlenose dolphins (*Tursiops truncatus*) along the west coast of Florida. *Marine Mammal Science*, *24*, 774–794.

Wells, R. S., & Fahlman, A. (2024). Human impacts on dolphins: Physiological effects and conservation. In A. Fahlman & S. K. Hooker (Eds.), *The physiology of dolphins* (pp. 267–284). Academic Press. http://doi.org/10.1016/B978-0-323-90516-9.00007-5

Wells, R. S., Rhinehart, H. L., Hansen, L. J., Sweeney, J. C., Townsend, F. I., Stone, R., Casper, D. R., Scott, M. D., Hohn, A. A., & Rowles, T. K. (2004). Bottlenose dolphins as marine ecosystem sentinels: Developing a health monitoring system. *EcoHealth*, *1*, 246–254.

Wells, R. S., Schwacke, L. H., Rowles, T. K., Balmer, B. C., Zolman, E., Speakman, T., Townsend, F. I., Tumlin, M. C., Barleycorn, A., & Wilkinson, K. A. (2017). Ranging patterns of common bottlenose dolphins *Tursiops truncatus* in Barataria Bay, Louisiana, following the *Deepwater Horizon* oil spill. *Endangered Species Research*, *33*, 159–180.

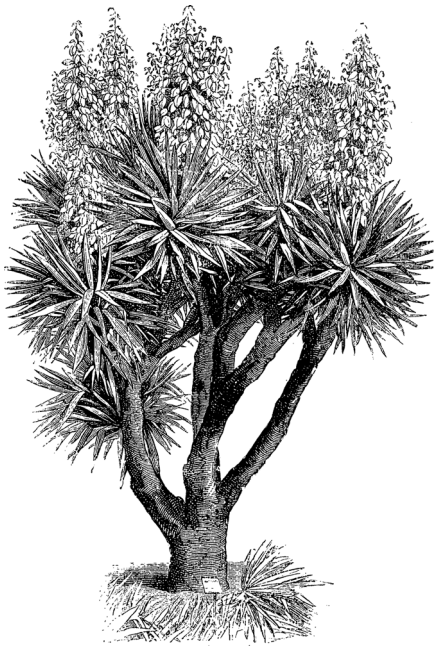
Wells, R. S., & Scott, M. D. (1997). Seasonal incidence of boat strikes on bottlenose dolphins near Sarasota, Florida. *Marine Mammal Science*, *13*, 475–480.

Winship, C., & Mare, R. D. (1984). Regression models with ordinal variables. *American Sociological Review*, *49*(4), 512–525.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Pirotta, E., Harris, C. M., Schwacke, L. H., Thomas, L., Allen, J. B., Barleycorn, A. A., Casoli, M., Crossman, J., Aichinger Dias, L., Engleby, L., LaBrecque, E., Litz, J., McHugh, K. A., Miller, P. J. O., Moore, M. J., Photopoulou, T., Quigley, B. M., Smith, C. R., Southall, B. L., … Tyack, P. L. (2026). An experimental approach to assess the combined effects of multiple stressors on a large vertebrate species. *Conservation Biology*, *40*, e70224. https://doi.org/10.1111/cobi.70224



15231739, 2026, 4, Downloaded from https://conbio.onlinelibrary.wiley.com/doi/10.1111/cobi.70224, Wiley Online Library on [21/09/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License