Measuring critical thermal maximum in aquatic ectotherms: a practical guide

Graham D. Raby[1,2]*, Rachael Morgan[3], Anna H. Andreassen[4], Erin M.C. Stewart[2], Jeremy de Bonville[5], Elizabeth C. Hoots[6], Luis Kuchenmüller[6], Moa Metz[7], Lauren E. Rowsey[8], Leon Green[9], Robert A. Griffin[10], Sidney Martin[10], Heather Bauer Reid[2], Rasmus Ern[7], Eirik Ryvoll Åsheim[11], Zara-Louise Cowan[9,12], Robine H.J. Leeuwis[7], Tamzin A. Blewett[10], Ben Speers-Roesch[8], Timothy D. Clark[6], Sandra A. Binning[5], Josefin Sundin[11], Fredrik Jutfelt[7,9]

1 - Department of Biology, Trent University, ON, Canada
2 - Environmental & Life Sciences Graduate Program, Trent University, ON, Canada
3 - Department of Biological Sciences, University of Bergen, Bergen, Norway
4 - DTU Aqua, National Institute of Aquatic Resources, Technical University of Denmark, Kongens Lyngby, Denmark
5 - Groupe de recherche interuniversitaire en limnologie (GRIL), Département de sciences biologiques, Université de Montréal, QC, Canada
6 - School of Life and Environmental Sciences, Deakin University, Geelong, VIC, Australia
7 - Department of Biology, Norwegian University of Science and Technology, Trondheim, Norway
8 - Department of Biological Sciences, University of New Brunswick, Saint John, NB, Canada, E2L 4L5
9 - Department of Biology and Environmental Sciences, Kristineberg Center, University of Gothenburg, Fiskebäckskil, Sweden
10 - Department of Biological Sciences, University of Alberta, Edmonton, AB, Canada.
11 - Department of Aquatic Resources, Swedish University of Agricultural Sciences, Drottningholm, Sweden
12 - Natural Resources Institute Finland (Luke), Oulu, Finland

*author for correspondence: grahamraby@trentu.ca

Abstract

Critical thermal limits, commonly quantified as CT_{max} (maximum) or CT_{min} (minimum), are core metrics in the thermal biology of aquatic ectotherms. CT_{max}, in particular, has recently surged in popularity due to its various applications, including understanding and predicting the responses of animals to climate warming. Despite its growing popularity, there is a limited literature aimed at establishing best practices for designing, running, and reporting CT_{max} experiments. This lack of standardisation and insufficiently detailed reporting in the literature creates challenges when designing CT_{max} studies or comparing results across studies. Here, we provide a comprehensive, practical guide for designing and conducting experiments to measure critical thermal limits, with an emphasis on CT_{max}. Our recommendations cover 12 topic areas including apparatus design, masking (blinding), warming rates, endpoints, replication, and reporting. We include diagrams and photos for designing and building critical thermal limit arenas for field or lab applications. We also provide a reporting checklist as a reference for researchers when carrying out experiments and preparing manuscripts. Future studies incorporating critical thermal limits would benefit from transparent reporting of warming/cooling rates (raw data, supplementary graphs) and photo/video evidence showing arena designs and critical thermal limit endpoints. We also provide directions for empirical research that will help further inform the measurement of critical thermal limits, including on biotic factors like stress and digestion, warming/cooling rates, the effects of body mass on heat transfer, and the physiological mechanisms underlying thermal tolerance.

Keywords: climate change, heat stress, heat shock, heat wave, experimental design, global warming, fish

Introduction

As climate change causes more frequent and severe aquatic heat waves (e.g., Tassone et al., 2023; Woolway et al., 2021), scientists are increasingly focused on understanding the thermal biology of aquatic animals. Tolerance to acute temperature changes, as experienced during heat waves, is typically assessed through experiments that involve gradual increases in temperature (e.g., degrees per minute or per hour). Experimentally testing acute thermal tolerance has more direct ecological relevance for some species and contexts than others (e.g., rockpools can undergo extreme changes over a timescale of minutes; Desforges et al., 2023; Smit & Glassom, 2017). Nevertheless, standardised testing of tolerance to acute heat stress has been useful across a range of disciplines and aquatic ecosystems (Portner et al., 2005; Sunday et al., 2012; Terblanche et al., 2011).

Critical thermal maximum (CT_{max}) is the most widely used metric to quantify the acute upper thermal tolerance of ectotherms (Beitinger et al., 2000; Desforges et al., 2023; Ern et al., 2023). CT_{max} is typically reported as the temperature at which an animal loses motor control – often following a distinct period of agitation – during thermal ramping. This method is useful across a range of applications and research questions, such as predicting species distributions (Sunday et al., 2012; Woodin et al., 2013), investigating impacts of stressors (Haney & Walsh, 2003; Patra et al., 2009), quantifying acclimation and adaptation (Fangue et al., 2006; Ruthsatz et al., 2024; Comte & Olden, 2017) and investigating physiological mechanisms underlying thermal tolerance (Cuenca Cambronero et al., 2018; Ern et al., 2023). For aquatic ectotherms, CT_{max} experiments involve placing animals into an arena for observation, warming the water in the arena at a steady rate, and then recording the temperature at which each individual reaches its endpoint (loss of motor control, e.g., loss of equilibrium [LOE], state of immobility, or onset of spasms; Cowles & Bogert, 1944). As a result of its simplicity, CT_{max} has long been popular among ecologists and physiologists. However, over the past 15 years, interest in the CT_{max} of

aquatic ectotherms has grown dramatically (Fig. 1), suggesting that many researchers are now trying the method for the first time.

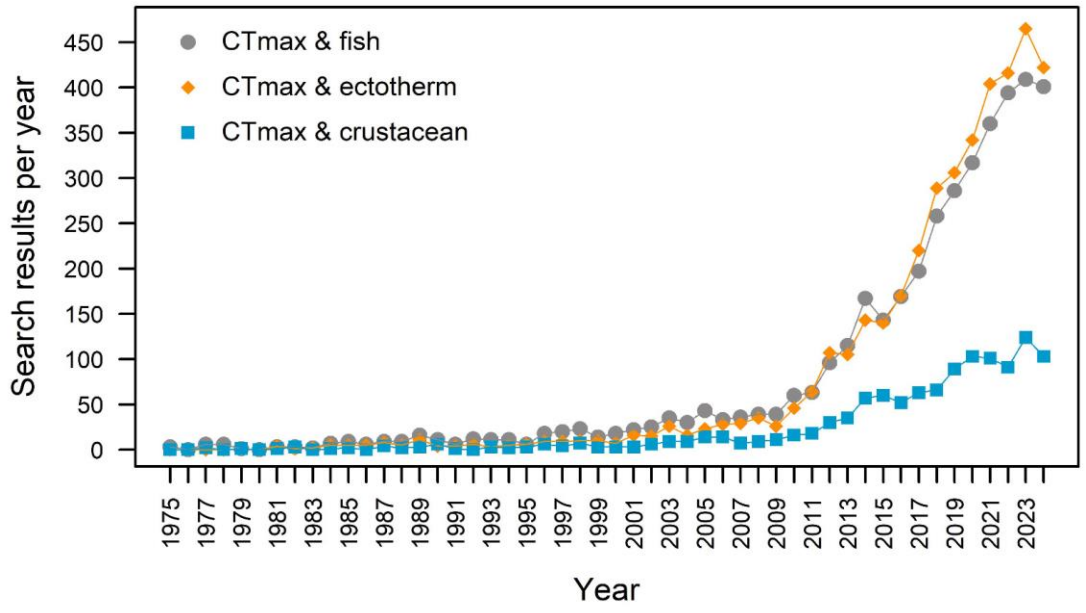


Fig. 1. The number of scientific articles mentioning CT_{max} were collected for each year between 1975 and 2024 using Google Scholar (2025-01-21 search terms: "CTmax & fish", "CTmax & ectotherm", "CTmax & crustacean"), indicating a surge in interest in CT_{max} over the past *ca*. 15 years.

Despite its popularity and potential for standardisation, CT_{max} experiments vary greatly in how they are conducted and reported. As a result, the data from CT_{max} experiments likely vary in how interoperable and reusable they are. Papers have focused on a range of aspects of CT_{max} methodology including warming rates (Galbreath et al., 2004; Mora & Maya, 2006), within-individual repeatability (Morgan et al., 2018; Grinder et al., 2020), and behavioural endpoints (Lutterschmidt & Hutchison, 1997; Cowan et al., 2023). Lacking, however, is a comprehensive practical guide for measuring CT_{max} across a range of aquatic ectotherms.

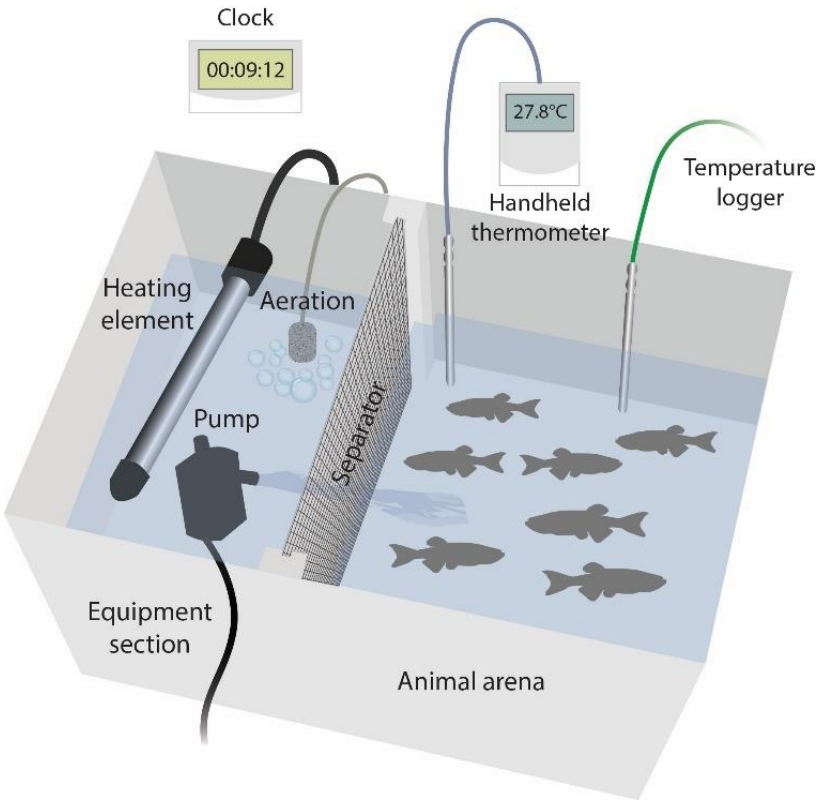
This paper builds on the authors' experience with CT_{max} methodology and presents a set of best practices for conducting CT_{max} trials on aquatic ectotherms. With the guidance

presented here, we hope to standardise the methodology and reporting of CT_{max} experiments to enhance replicability, reliability, and comparability of CT_{max} data. While there are other useful measures of maximum thermal limits (e.g., upper incipient lethal temperature [UILT], Thermal Death Time; Hasnain et al., 2013; Jørgensen et al., 2021) including field data on thermal occupancy (Bear et al., 2007; Dallas & Ketley, 2011; Challice et al., 2019; Webb et al., 2020), CT_{max} is our focus here. Most of our recommendations are transferable to other methods for estimating thermal limits like median lethal temperature (LT_{50}) experiments, critical thermal minimum (CT_{min}), or CT_{max} measurements of terrestrial ectotherms. Our set of best practices is primarily based on experiments with fishes, but we include insights from studies on aquatic invertebrates (section 5).

1. Overview of CT_{max} apparatus design

In a CT_{max} trial, water is continuously heated in an arena where animals can be visually tracked and scored. There is no standard design for a CT_{max} apparatus; a customized aquarium tailored to the experiment's context and species is often necessary. Here we describe a template that can be modified as needed. Our design (Fig. 2) features a tank with a mesh separator dividing the animal arena from a small section of the tank (the 'equipment section') containing the heating element(s), a submersible pump for mixing (section 4), and an air stone for aeration. Most of the tank is reserved for the animals ('animal arena'), providing sufficient water volume for horizontal and vertical movement. The animal arena should have a background colour that contrasts with the animals (e.g., Fig. 3b-d), to facilitate manual or automated visual tracking. The tank should contain at least one temperature logger in the animal arena (see section 3 and 12), and at least one temperature sensor connected to a digital display so that temperature can be read (section 4) and recorded by experimenters (and/or on video) in real time as animals reach their endpoints (sections 6 and 7). Temperature sensors should be positioned so that

123 they are not in contact with the walls or bottom of the tank as these may warm at a different
124 rate than the water. The mesh separator used to partition the animal arena from the equipment
125 arena should be fine enough to prevent animals from getting through, while allowing ample
126 water circulation. This design allows for the animal arena to be free from most equipment,
127 tubing, and other obstructions, which helps with observing and netting individuals. It also
128 prevents animals from hiding behind equipment and reduces the risk of harmful contact with a
129 heating element or a pump.
130



131

132 **Fig. 2.** Diagram of a template design for a CT_{max} apparatus. Aeration, mixing mechanisms
133 (pump), and heating elements are in an 'equipment section' that is physically separated via a
134 screen from the part of the arena containing the animals ('animal arena'; here with fish). A
135 submerged temperature sensor (or data logger) is used to log temperatures, in addition to a
136 temperature sensor with a display that is used to score CT_{max} in real time. The time and
137 temperature should only be visible to the data recorder (and/or video camera), not to the person
138 observing animals for endpoints (see section 6).
139

In our design, a uniform heating rate among replicate trials (section 3) is achieved by using consistent heating power (i.e., number and type of heaters), a fixed water volume, and, ideally, a fixed surrounding air temperature (but see section 10). As such, the CT_{max} tank should be filled with a standardised volume of water between trials (e.g., by having a 'fill to' water level mark) (section 3) and the room temperature kept constant. Ideally, the tank should be made of thick or insulated material to reduce heat exchange with the surrounding environment. For example, CT_{max} tanks can be constructed using insulated containers, or by adding insulating foam to the outside of the tank (Fig. 3a-c).

Modifications to this suggested template are often necessary to accommodate factors like animal size. For example, if the mesh size needs to be small due to small animals tested, it may be useful to pump water via a tube from the equipment section to the animal arena to promote mixing and thermal homogeneity (Fig. 3b). It is important to ensure that water mixing is not so vigorous as to force the animals to swim actively to maintain position (section 4). Alternatively, a second tank (sump) can be connected by tubing to the animal arena with constant water exchange between the two – the sump tank then serves as the 'equipment section' (Fig. 3a,c). This latter design may allow a higher total water volume and necessitate more heating power, along with more powerful submersible pumps to ensure sufficient mixing between the sump and the animal arena. Submersible pumps do generate some heat and sound, which can influence heating rate and potentially stress animals. Pumps should therefore be kept consistent (position, make, model) among replicate trials.

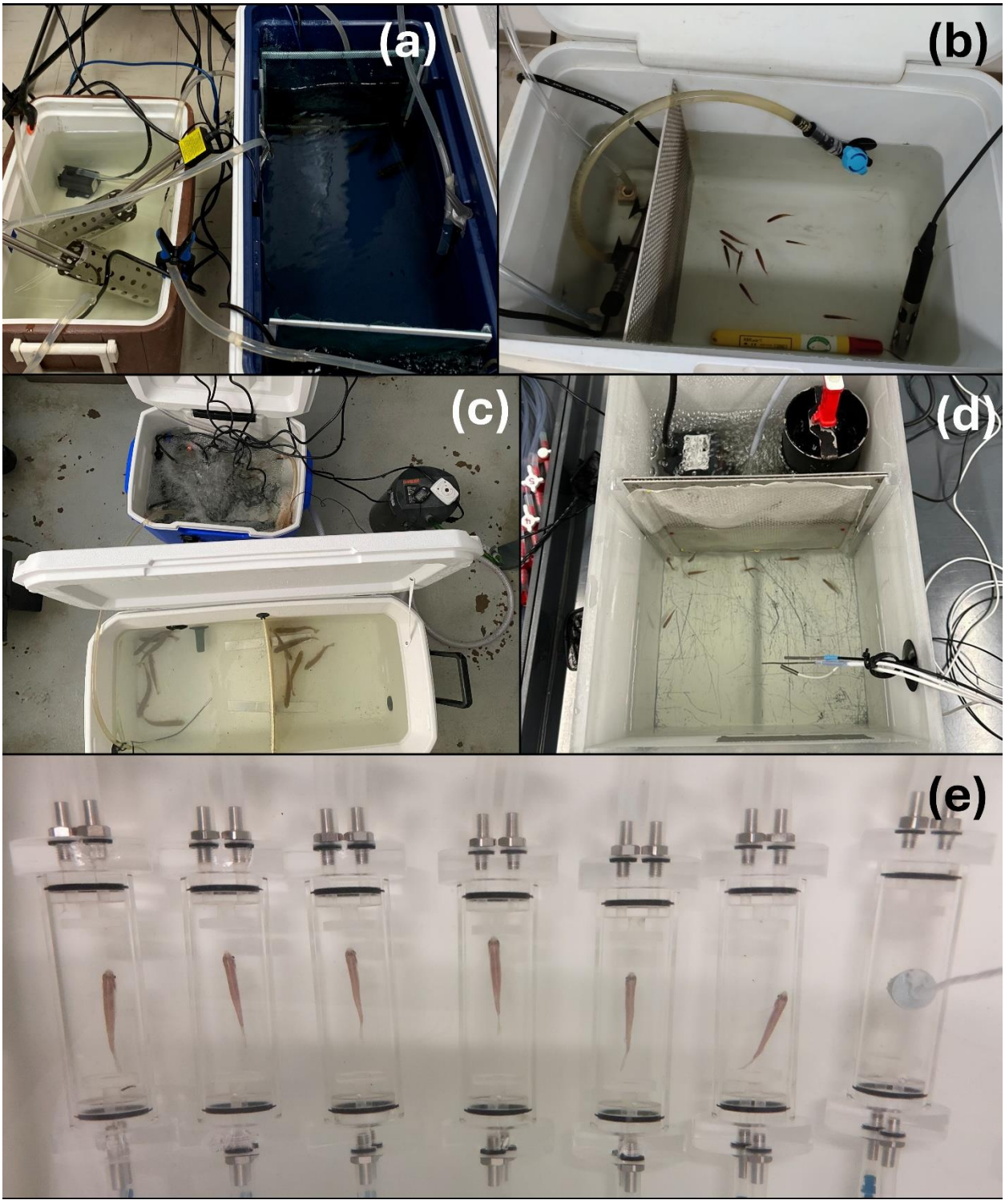


Fig. 3. Images of CT_{max} apparatus designs. **(a)** A sump connected design, with the sump tank (left) containing heating elements and a pump to recirculate water to the animal arena (right) used for pumpkinseed (*Lepomis gibbosus*). Note the dark background of the arena in this example makes it more difficult to observe the fish (a contrasting background is preferred). **(b)**

CT_{max} apparatus similar to our template design (Fig. 2). A YSI Pro-solo oxygen and temperature sensor is placed in the animal arena (right) along with an additional temperature logger (RBRsolo^3 T, https://rbr-global.com/). The hose and valve going from the left side of the partition to the right is connected to a submersible pump to facilitate water mixing. Setup used for several species, here bluntnose minnow (*Pimephalus notatus*). **(c)** A split design similar to **(a)**, in which a sump (top) contains heaters, aeration, and pumps to constantly exchange water with the animal arena (bottom). In this example, there is a partition in the animal arena to keep two replicate groups of brook trout (*Salvelinus fontinalis*) separated. The template design in Fig. 2 is based on the arena in **(d)** (from Morgan et al., 2019). In addition to a temperature sensor connected to a handheld display (not shown), there is a second thermistor and an oxygen sensor from a FireSting-O_2 meter (PyroScience, https://www.pyroscience.com/). The black metal cylinder in the top section of the arena in **(d)** is designed to hold a 300 W heating element and connected to the submersible pump to ensure water flow over the heating element and then out of the metal cylinder towards the animal arena without melting and damaging the tank. **(d)** shows a CT_{max} trial on a group of zebrafish (*Danio rerio*) while **(e)** displays zebrafish housed individually in flow-through respirometry chambers. The setup includes six trial chambers and one control chamber equipped with a temperature probe, each receiving water from a flush pump.

Another CT_{max} apparatus design is to house animals individually or in groups within replicate chambers (i.e., multiple animal arenas) that are immersed in a common temperature-controlled water bath that contains the heater and mixing pump (e.g., Healy & Schulte, 2012). With the latter design, each chamber is fitted with at least one temperature sensor and the chambers allow water to flow through, using a pump via tubing. Housing animals individually allows easier quantification of individual agitation temperatures from video (Wells et al., 2016; Firth et al., 2021), a task that can be more difficult when many individuals are moving around in the animal arena. Furthermore, individual housing chambers may allow for the CT_{max} trial to be combined with measurements of other physiological responses. For example, animals tethered via catheters or sensors typically require confinement.

Measurement of CT_{max} can also be done with animals confined individually in respirometry chambers so that oxygen uptake can be estimated (Ern et al., 2016, 2017; Fig. 3e). Depending on the size of the respirometry chambers, this technique might restrict movement with unknown (likely species-specific) consequences for CT_{max}. Fish exposed to water oxygen

levels (hypoxic, e.g., <70% air saturation) during typical CT_{max} trials with an open water surface may exhibit aquatic surface respiration (Rutledge & Beitinger, 1989). Aquatic surface respiration could potentially affect CT_{max} and can be prevented using submerged chambers. Recording the number of aquatic surface respiration events can be used to document changes in frequency of this behaviour during thermal ramping as an additional variable to explore (Francispillai & Chapman, 2025).

A heating mantle design can be useful for testing CT_{max} in very small organisms such as aquatic larvae or embryos (Andreassen et al., 2022; Cowan et al., 2023). In this design, the water in the animal arena is not exchanged, but heated from a surrounding heating mantle (Fig. 4). The setup requires aeration and recording of temperature and oxygen level directly in the animal arena but has the advantage of maintaining a still water surface, enabling clear video recording.

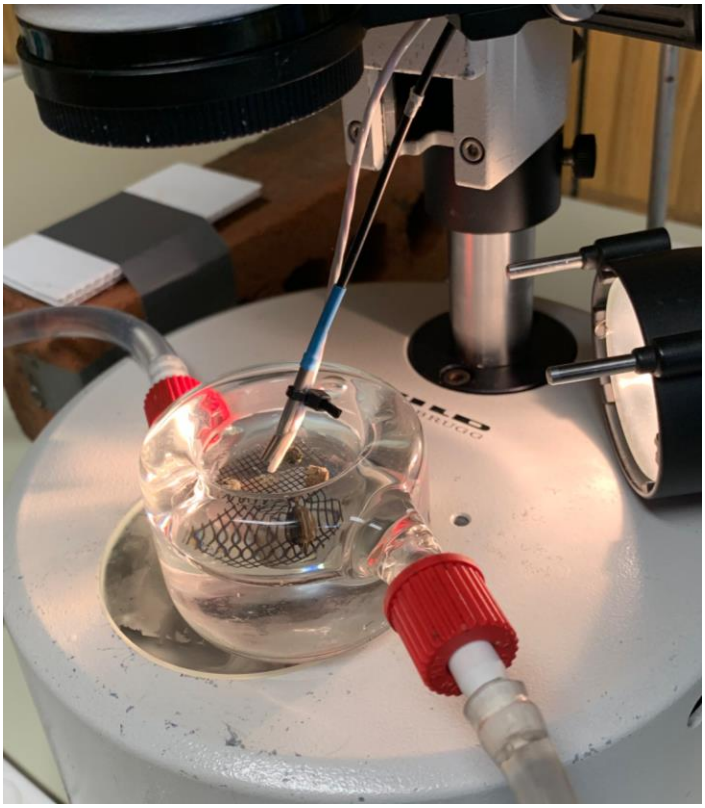


Fig. 4. Custom-built glass chamber including a shallow well (20 mL volume) to assess CT_{max} of aquatic embryos or larvae. The chamber is a well in an enclosed water bath (heating mantle) heated by water pumped through the red nozzles. Embryos are at the bottom of the well, inside a fine mesh, kept in place by small pebbles. Larvae can be tested in the same arena without a mesh. The heating mantle is placed on a stereomicroscope, attached to a camera, to monitor individual movement. An optical oxygen probe and temperature sensor are placed in the animal arena without blocking the view of the embryos or larvae.

2. Considerations for selecting a warming rate

Warming rates can affect the CT_{max} values obtained from a trial (Elliott & Elliott, 1995), and thus an appropriate warming rate must be identified before the experiments begin. There are two key considerations for selecting a warming rate: 1) The rate of warming should be slow enough to allow internal body temperature to remain close (ideally, within 0.2°C) to the temperature of the surrounding water as it warms, and 2) since thermal acclimation can manifest within hours of heat exposure in some species (De Bonville et al., 2025), the rate should be fast enough to avoid unwanted (partial) acclimation to intermediate temperatures along the temperature ramp (Åsheim et al., 2020; Jutfelt et al., 2019; Penman et al., 2023). Regarding point 1, it is notable that the processes and tissues that govern CT_{max} remain incompletely understood (Ern et al., 2023), so consideration must be given to the warming rate and the thermal inertia of different sized animals and the tissues within them (e.g., thermal inertia may be higher, and the lag time greater, for the body cavity and deep muscle of a fish in comparison with its brain or heart).

Optimal rates of warming in CT_{max} experiments may differ across taxa, life stages, body sizes, acclimation temperatures, research questions, and in relation to the natural environment of the species (Mora and Maya, 2006; Moyano et al., 2017; Vinagre et al., 2015). A warming rate of 0.3°C min^{-1} is commonly used and generally accepted as a standard for small fish (Becker & Genoway, 1979). In Atlantic salmon parr (*Salmo salar*), warming rates ranging from 1 to 18°C h^{-1} (0.02 to 0.3°C min^{-1}) produced no differences in CT_{max} (Elliott & Elliott, 1995). Becker and Genoway (1979) found that in juvenile coho salmon (*Oncorhynchus kisutch*) and pumpkinseed

(*Lepomis gibbosus*), warming rates of 0.5 and 1°C min^{-1} produced higher temperatures for LOE than did 0.3°C min^{-1}. Since rates higher than 0.3°C min^{-1} can produce delayed physiological responses (see below), when appropriate, we recommend using 0.3°C min^{-1} to allow for comparison across studies and species.

A faster warming rate may not allow a test animal's internal body temperature to equilibrate with the water temperature, especially in larger animals. Internal temperature lag (thermal inertia) depends on body size: larger animals have more thermal inertia than do smaller animals (Kitagawa & Kimura, 2006; Nakamura et al., 2020). Thus, larger animals typically require slower warming rates if the goal is to ensure core tissues remain in thermal equilibrium with water temperature (Sandblom et al., 2016; Morgan et al. 2018; Jutfelt et al., 2019; Fig. 5). For example, in ~122 g juvenile Atlantic cod (*Gadus morhua*) warmed at 0.3°C min^{-1}, the difference between the water temperature and that of the deep muscle tissue was around 1.5°C (Jutfelt et al., 2019). Here, CT_{max} may be overestimated due to the discrepancy between the water temperature and internal temperature of the fish (Fig. 5). Similarly, at a 2°C h^{-1} warming rate, the internal temperature of 15.1–120.4 g brook trout (*Salvelinus fontinalis*) did not lag behind the surrounding water, but that during an instantaneous 8°C temperature change, lag time to equilibrium increased with fish size at rate of 7.2 s per mm of fish length (O'Donnell et al., 2024).

As noted, thermal inertia likely differs among tissues in most ectotherms. For example, the brain is highly vascularised and generally close to the body surface, allowing it to become equilibrated to external temperature considerably faster than deep dorsal muscle (Jutfelt et al., 2019). Furthermore, thermal inertia may be influenced by the body shape (e.g., spherical *vs*. flat) and colour (e.g., dark *vs*. light) of the animal (Stevenson, 1985). Hence, we discourage attempts at *post-hoc* correction of CT_{max} data based on heat transfer time, given that species-specific mechanisms (and relevant tissues) involved in CT_{max} are complex and not fully understood (Ern

et al., 2023). The lag between the internal and external ramping rates can be quantified with a pilot experiment (preferably on anaesthetised animals) in which thermocouples are implanted into the deep muscle tissue (Mora and Maya, 2006; Sandblom et al., 2016;). Pilot experiments should be run before any new CT_{max} study, especially when measuring CT_{max} in a species for which an appropriate ramping rate has not already been identified.

Lastly, it is important to strive for consistent and homogenous warming rates within and across studies, as CT_{max} can vary as a result of inconsistencies in warming rate (e.g., Becker & Genoway, 1979; Mora and Maya, 2006; Moyano et al., 2017; Åsheim et al., 2020;). Although it has been suggested that discrepancies in warming rates should be standardised by modelling 'thermal death time' or cumulative exposure to thermal stress (Jørgensen et al., 2021; Ørsted et al., 2022), the potential for animals to partially acclimate during slower warming rates may make the comparison of thermal death time estimates inconsistent and unreliable, plus there are ethical issues with using death as an endpoint. Until this method has been more thoroughly validated, we recommend adhering to a consistent 0.3°C min^{-1} heating rate or, for larger animals, the fastest warming rate that results in minimal thermal lag of the tissues of interest and provide a rationale for the warming rate used.

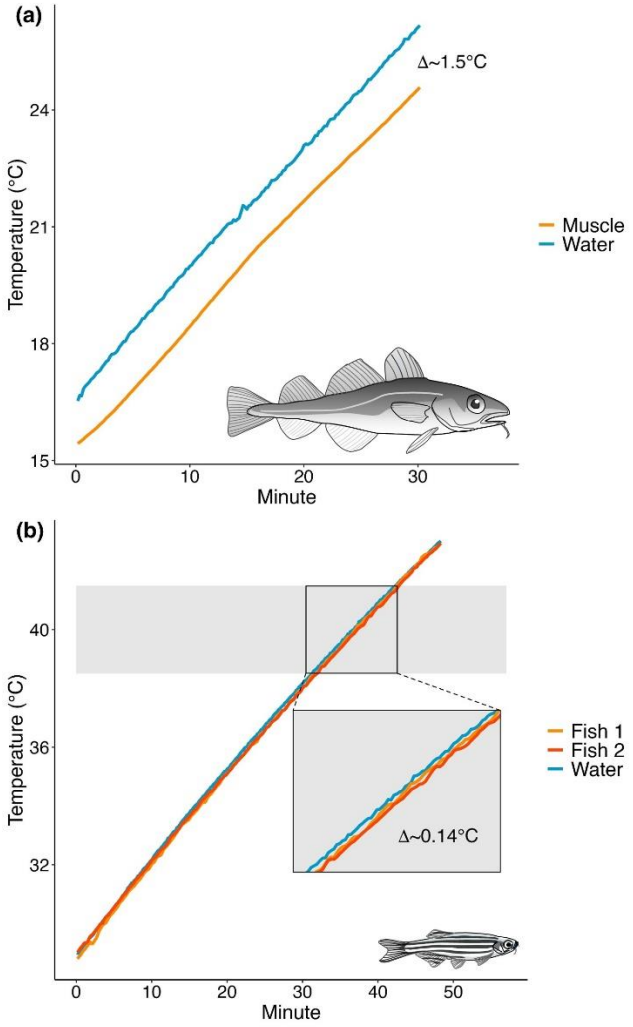


Fig. 5. During 0.3°C min^{-1} warming, internal body temperature lags behind environmental temperature with the magnitude of the lag dependent on body size. **(a)** Body temperature (°C) vs. water temperature (°C) during CT_{max} trials in the deep dorsal muscle of a 25.1 cm and 122.3 g Atlantic cod (*Gadus morhua*) (data from Jutfelt et al., 2019). The difference between the cod dorsal muscle and the water was consistently around 1.5°C once the 0.3°C min^{-1} warming was underway. **(b)** Similar data for two zebrafish (Fish 1: 2.8 cm and 0.4 g; Fish 2: 2.6 cm and 0.4 g , data from Morgan et al. 2018). The grey shaded area indicates the range of expected CT_{max} for zebrafish (*Danio rerio*). Zebrafish internal temperatures were, on average, 0.14°C lower than the water temperature during the range of temperatures when CT_{max} was most likely to occur (inset). In each case, the animal was anaesthetised throughout.

3. How to achieve the desired warming rate

There are a few strategies that can be used to achieve a constant warming rate within the animal arena. First, the power needed to heat the chosen water volume in the CT_{max} apparatus can be estimated as follows:

$$P = mc\Delta T \quad \text{(Eq. 1)},$$

where P is the power of the heater (in W), c is the specific heat capacity of water (approximately 4186 J kg^{-1} $°C^{-1}$), m is the mass of water (in kg), and ΔT is the desired rate of heating (in °C s^{-1}). This equation does not account for heat exchange with the surrounding environment; pilot trials and adjustments are often needed. However, once the heating power and water volume are determined via pilot experiments, achieving the desired warming rate consistently for each trial is as simple as filling the chamber to the correct volume and turning on the heater(s). A drawback with this approach is that the warming rate may be somewhat different at high or low absolute temperatures, depending on the surrounding air temperature (Fig. 6, and see section 10).

A more labour-intensive approach is to manually monitor and adjust the water temperature throughout the CT_{max} trial (e.g., Åsheim et al., 2020; Stewart et al., 2023, 2024). This might include adding or removing heaters in a stepwise manner and/or adjusting the water volume. An automated proportional–integral–derivative controller can be used to precisely control the rate of warming throughout the trial (e.g., McDonnell & Chapman, 2015; Ern & Jutfelt, 2024), including automating increases to heating power at higher temperatures to maintain warming rates (this requires equipping the system with overcapacity heating relative to calculations from Equation 1). Regardless of the warming rate, researchers should log and report the raw temperature data from inside the CT_{max} arena (Fig. 6, section 12).

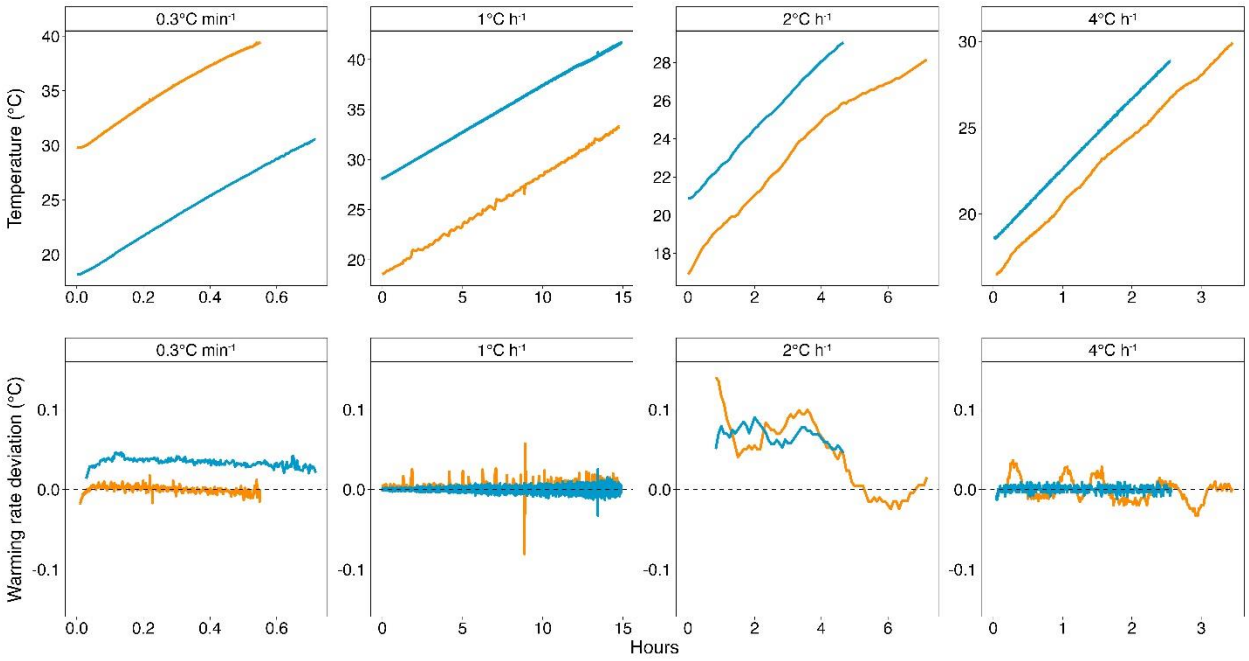


Fig. 6. Warming rate data from CT_{max} trials at four warming rates: 0.3°C min^{-1} (18°C h^{-1}), 1°C h^{-1}, 2°C h^{-1}, and 4°C h^{-1} (4°C h^{-1} data from Stewart et al., 2023 & Stewart et al., 2024, all other data are Raby et al., unpublished data) that varied in how well they adhered to the desired warming rates. The top row depicts actual warming rates, and the bottom row the deviation from the desired warming rate in degrees Celsius per unit of time (unit of time matching the panel label, i.e., per minute for the left panel, per hour for the others). Each panel shows the temperature change during two different trials, distinguished by colour. Blue lines depict examples of warming rates that were more consistent throughout the duration of the trial. Orange lines depict examples of less optimal or more variable warming rates, with the 2°C h^{-1} example being the only shown here that is truly problematic. Black dashed lines indicate a warming rate deviation of 0°C (i.e., the ideal warming rate). Warming rate deviation data are smoothed using a rolling mean over 10 recording intervals. Recording intervals varied by trial: from left to right, temperature was recorded at 3 s, 1 s, 5 min, and 15 s intervals. See Table S1 and S2 for warming rates (slopes and R^2 values).

4. Thermal homogeneity

Thermal homogeneity across the animal arena (e.g., <0.1°C in variation) is a critical assumption of CT_{max} because the animals will not consistently be in the direct vicinity of temperature probe(s). Commonly used approaches to heating (e.g., titanium heating rods) can lead to thermal heterogeneity (with insufficient mixing of the water). In turn, some areas of the arena may become warmer than others, which could cause over- or under-estimation of CT_{max} among

342 individuals based on their positioning. Aquatic ectotherms behaviourally thermoregulate and
343 thus are likely to seek out cooler areas of an arena once the water temperature becomes supra-
344 optimal.

345 Thermal homogeneity within the testing arena can be maintained with a submersed
346 recirculating water pump. Vigorous aeration also helps promote thermal mixing. However, the
347 rate of water movement within the arena should not be so high that the animals have to swim
348 actively against a current, which could cause exhaustion prior to CT_{max} (e.g., Blasco et al., 2020).
349 Thus, the flow rate of the pump should be minimised to the lowest level needed to ensure
350 thermal homogeneity (see below). Baffles can be used to disperse the flow, and observations of
351 animal behaviour can be used to determine if water velocities are too high. In arenas with high
352 rates of water mixing, placing fish in chambers with a lower flow, continuously running flush
353 pump can protect them from excessive water movement and ensure uniform flow rates for all
354 fish (Fig. 3e). Assessment of thermal homogeneity can be made easier by using multi-probe
355 temperature logging systems (e.g., PicoTech TC-08, Pico Technology, Cambridgeshire, UK).
356 Confirmation of thermal homogeneity should occur with every new CT_{max} arena setup.
357

358 **5. Determining a repeatable endpoint**

359 Though some approaches use lethal endpoints (e.g., IULT or LT_{50}), the prevailing definition of
360 upper or lower thermal tolerance is an 'ecological death' of the organism due to loss of motor
361 functions and, consequently, an inability to escape the harmful (lethal) conditions or avoid
362 predators (Lutterschmidt & Hutchison, 1997; Vernon, 1899). When selecting an endpoint, it is
363 important to remember that CT_{max}, unlike LT_{50}, is a non-lethal assay of an organism's thermal
364 tolerance; the animal should recover and survive upon being returned to cooler water. Rates of

365 survival after CT_{max} should therefore be reported. Post-CT_{max} mortality can occur minutes to
366 days following a trial (De Bonville et al., 2024).
367 Loss of motor function is typically assessed through behavioural endpoints: LOE
368 (Beitinger et al., 2000) is the most common (at least in fishes), but loss of swimming (Geerts et
369 al., 2015) and loss of response to touch (Andreassen et al., 2022) are also used. Cessation of
370 movement of embryos (Cowan et al., 2023) and loss of responsiveness of larval fish
371 (Andreassen et al., 2022) are comparable alternatives to LOE. Assessing embryo CT_{max} requires
372 that embryos are mature enough to have spontaneous activity and exhibit movement in
373 response to a thermal stimulus. While differences in methodological approaches can make it
374 difficult to compare CT_{max} across studies, behavioural and morphological differences among
375 species or life stages can demand different endpoints. For example, during warming, flatfish
376 and many invertebrates often lay still on the bottom, gastropods such as abalone attach
377 themselves to the arena walls, and pipefish tend to coil themselves around each other or the
378 temperature sensor (personal observations by the authors). Due to variation among species and
379 life stages in endpoints, it is important to: 1) run pilot trials to map out species-specific
380 behavioural responses to the warming; 2) accurately determine and define the endpoint (i.e., the
381 behavioural response and duration threshold to identify CT_{max}); 3) ensure the endpoint definition
382 is clearly communicated and agreed upon among all researchers performing experiments (or,
383 ideally, only use one observer for all replicate trials); and 4) report details on the behaviour
384 observed at the agitation temperatures and the endpoint (section 12). Video recording can also
385 help to determine consistent endpoints.
386 The LOE or loss of responsiveness endpoints are often preceded by hyperactivity (i.e.
387 agitation), then progressive loss of movement and/or increasing lethargy that eventually leads
388 to loss of equilibrium/responsiveness (e.g., Friedlander et al., 1976, Lutterschmidt and
389 Hutchison, 1997; Kochhann et al., 2021). Some animals also appear to lose equilibrium then

quickly regain motor functions when startled (Åsheim et al., 2024). Thus, there may be a trade-off between accuracy of the CT_{max} measurement and the risk of recording a premature LOE when determining the duration of the behavioural endpoint for CT_{max}. Whether or not the behaviour of individuals close to the endpoint differs among experimental treatment groups should be assessed during pilot tests. Different time thresholds for LOE to trigger an animal's endpoint (e.g., 1 *vs* 3 *vs* 10 s of LOE) have been used. This detail should be reported.

The sample size within a CT_{max} trial should be small enough to allow the observer to carefully monitor the behaviour of all the animals. At a 0.3°C min^{-1} warming rate, it is common for all animals in a trial to reach CT_{max} within as little as three minutes (i.e., a range of 0.9°C in CT_{max} values). In such a scenario, observing too many animals at once makes it difficult to precisely assess endpoints for each individual, in turn making CT_{max} times and temperatures less accurate. When using 0.3°C min^{-1}, we recommend limiting sample sizes to approximately 8–10 animals per trial. If the behavioural endpoint requires the observer to flip over or test responsiveness in animals, a lower sample size per trial is likely necessary (e.g., 4–6 per trial). With slower warming rates, animals typically reach CT_{max} over a longer period, effectively allowing for a higher within-trial sample size while maintaining the same level of careful observation. As a result, slower warming rates may mean that 15–20 animals can be accurately tested per trial. As many as 30 animals was feasible at a warming rate of 4°C $hour^{-1}$ (Stewart et al., 2024) because the slower rate meant that fish reached their CT_{max} endpoints across a span of ~20 min.

CT_{max} protocols can be established for slow moving or sessile organisms. For example, a study on abalone (*Haliotis rubra* × *H. laevigata*) used a customised tank to ensure animals remained attached to vertical walls (Holland et al., 2024). The CT_{max} endpoint was taken as the temperature when animals lost pedal adherence (Holland et al., 2024). The endpoints of less mobile invertebrates that attach to the edges of tanks can be observed as the animal suddenly

falling off a vertical surface to which it was attached (Giomi et al., 2019). Mobile macroinvertebrates such as decapods often swim, allowing similar CT_{max} endpoints as fish. In rusty crayfish (*Faxonius rusticus*), CT_{max} was typically preceded by the animal bursting up off the bottom of the arena then drifting back to the bottom with negative equilibrium (Chasse et al., 2025). For some species, such as flatfish or some crab species, it may be necessary to intermittently turn animals' upside down to assess their righting response. However, doing so too frequently could lead to exhaustion and earlier CT_{max} and should only be done at predetermined intervals for consistency.

6. Minimising observer bias and maximising scoring consistency

Predetermined expectations about the outcome of an experiment can lead to conscious or unconscious observer bias (Tuyttens et al., 2014). Evidence that observer bias affects results and interpretation has been documented broadly across the life sciences, and is likely a problem in much of experimental biology (e.g., Tuyttens et al. 2014; Holman et al., 2015). Masking (also known as blinding) the observer from the treatments is commonly used as a technique for reducing observer bias (Holman et al., 2015). Without masking, false positive findings are more likely, and treatment effects tend to be overestimated (Holman et al., 2015). To apply masking to CT_{max} trials, the observing researcher scoring the animals should ideally be unaware of the temperature in the animal arena by ensuring the display temperature (Fig. 2) is not visible to them. This can be achieved by having two researchers performing the experiments: one observing the animals, and one recording data. With this setup, the animal observer monitors the arena for behavioural endpoints and calls out when an animal is removed and placed into a numbered, individual recovery tank. Ideally, the animal observer is also masked to any treatments of the animals being tested.

If having two researchers running the CT_{max} trials is a limitation, an alternative is to film the trials and assess the temperature and time at the endpoint from masked videos. There may be other concerns or issues that prevent researchers from incorporating masking (Karp et al., 2022). Information revealing the temperature or treatment may be visible (both visually and auditorily, e.g., from seeing heaters or other equipment turning on and off). This can be counteracted by ensuring transparent recording of data, for example by video recording the setup and trial (Clark, 2017).

7. Replication and inter-trial variability

CT_{max} trials are usually conducted on groups of animals (e.g., 10 animals at a time), giving rise to the possibility that animals within a trial are not independent. Trial effects, where replicate CT_{max} trials on animals from the same treatment show markedly different mean temperatures or different variance do occur. However, testing animals individually requires many parallel arenas or more time to achieve similar sample sizes. We recommend assessing the number of replicate trials that are necessary to reliably detect a treatment effect given that most studies will likely continue to test groups of animals together for efficiency. Studies that use only one CT_{max} trial per treatment may report erroneous treatment effects (type I errors) simply due to variation among trials (trial effects, analogous to tank effects). To illustrate this point, we re-examined data from previous CT_{max} experiments to answer the following question: how many replicate trials are needed to reliably detect a difference (or lack thereof) in CT_{max} between two groups of animals?

We modelled the likelihood of detecting differences in the mean CT_{max} among replicate trials using Bayesian linear regression models (see Supplementary Information for detailed methods and results). In a dataset consisting of four replicate CT_{max} trials under control

conditions for twelve species of aquatic ectotherms (n = 5–10 individuals per trial; overall N = 417; Raby et al., unpublished data), we found that in seven of the twelve cases, there was no evidence of a trial effect. That is, trials 2-4 did not differ in their mean CT_{max} values from the first trial. In these cases, there appeared to be a low risk of a type I error due to trial effects, and thus experiments involving only one or two CT_{max} trials per treatment would have reached the same conclusions as experiments testing four trials per treatment. Conversely, in three of the 12 case studies (25%), the trial mean CT_{max} in at least one of trials 2-4 differed from the initial trial mean (i.e., there was a 95% likelihood that a subsequent trial mean would not equal the initial trial's mean). In two cases, two or more subsequent trials differed from trial 1 (mean replicate differences = 0.21°C; Fig. 7). These scenarios illustrate that type I errors (false positives) can occur when conducting only one CT_{max} trial group.

In another dataset consisting of twelve replicate CT_{max} trials from eight treatment groups in zebrafish (*Danio rerio*; n = 9–18 individuals per trial; overall N = 1192), there was only one treatment in which there was no evidence of any trial effects. In the other seven treatments, an average of 3 of the 11 subsequent trials differed from the first trial (range: 1–6). Again, these results illustrate that even with well-controlled experiments on the same animals (no treatment effects), statistically significant 'unexplained' trial differences can occur. Therefore, researchers should avoid a design in which a single CT_{max} trial per treatment group is used. Determining the total sample size needed per treatment and how animals are divided among trials can be based on power analyses where effect sizes and variance can be estimated from pilot trials and data from previous research. In general, we recommend at least three replicate trials per treatment because it allows for an assessment of inter-trial variability; if the three trials show meaningful inter-trial variability then more trials would likely be warranted.

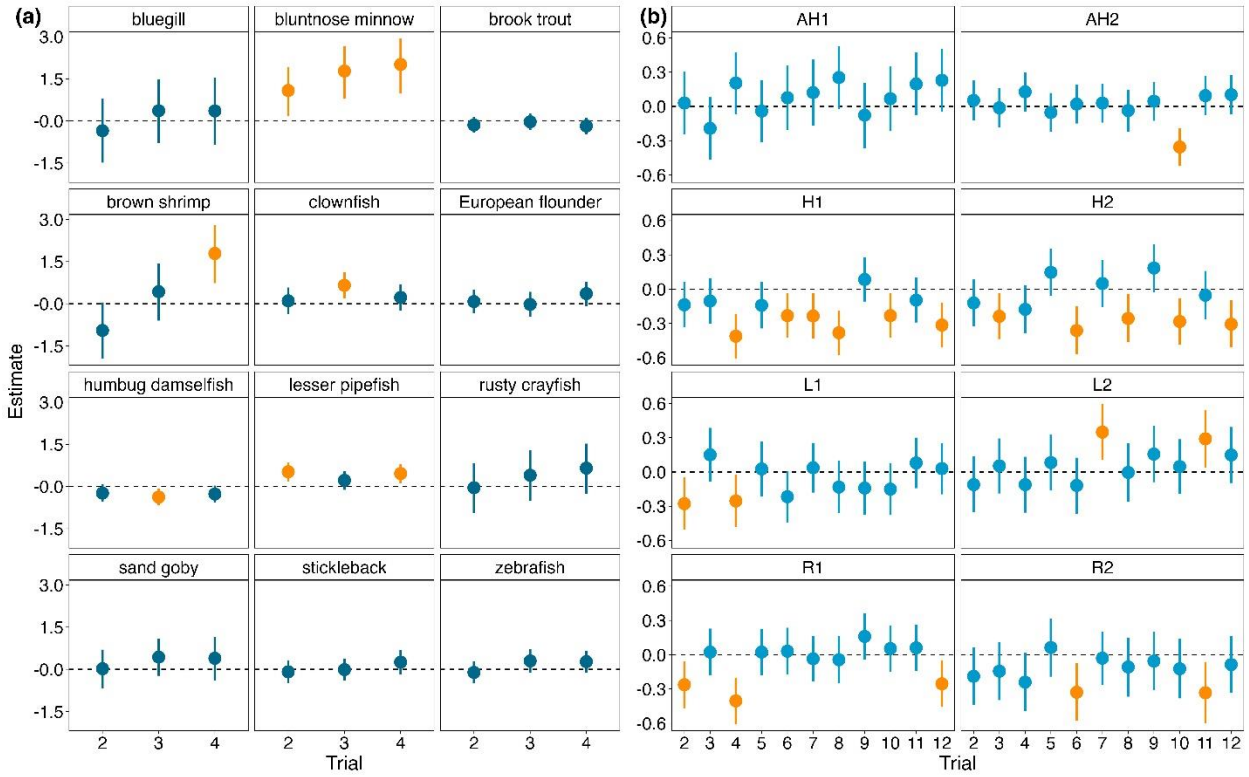


Fig. 7. Modelled posterior predicted mean (estimate; points) and posterior distributions (95% credible intervals; lines) of replicate trial CT_{max} in reference to the initial trial in **(a)** twelve species of aquatic ectotherms tested under controlled conditions (Raby et al., unpublished data: bluegill *Lepomis macrochirus*, bluntnose minnow *Pimephalus notatus*, brook trout *Salvelinus fontinalis*, brown shrimp *Crangon crangon*, clownfish *Amphiprion anemonefish*, European flounder *Platichthys flesus*, humbug damselfish *Dascyllus aruanus*, lesser pipefish *Syngnathus rostellatus*, rusty crayfish *Faxonius rusticus*, sand goby *Pomatoschistus rostellatus*, three-spined stickleback *Gasterosteus aculeatus*, zebrafish *Danio rerio*), and **(b)** zebrafish tested under eight treatment conditions (Morgan et al., 2020). Where posterior distributions do not cross 0, strong evidence of a given subsequent trial (replicate) effect exists (orange points), indicating a difference in mean from the initial trial. See Supplementary Information for methods and all model parameters.

8. Biotic confounds to consider

Of the many considerations for designing CT_{max} experiments, confounding effects of biotic factors (covariates) are poorly understood in most cases. However, we highlight some known aspects that researchers should consider. Infection with parasites and pathogens has been associated with reduced thermal tolerance in juvenile brown trout (*Salmo trutta*; Bruneaux et al.,

2017), pumpkinseed (De Bonville et al., 2024), bluegill (*L. macrochirus)*, and longear sunfish (*L. megalotis*; Lutterschmidt et al., 2007). If parasites are not a covariate of interest, researchers could consider treating wild animals for parasites upon arrival to the laboratory, quantifying parasite load following tests, and/or excluding noticeably infected individuals (Chrétien et al., 2023). However, removing infected hosts from a study could lead to an overestimation of the CT_{max} of the natural population.

Food quality, ration size, and the time elapsed since the last meal could all affect CT_{max} (Verhille et al., 2016; Gomez Isaza et al., 2019). Juvenile barramundi (*Lates calcarifer*) fed a 20% fat diet for four weeks had lower CT_{max} than those fed a diet containing 10% fat (both groups were fasted for 40–48 h prior to CT_{max}; Gomez Isaza et al., 2019). Through one anecdotal observation, we noted that juvenile European flounder (*Platichthys flesus*) that were mistakenly fed 6 h prior to a CT_{max} trial exhibited markedly lower CT_{max} than did fish fasted for 24 h (De Bonville et al., unpublished data). Similarly, juvenile white sturgeon (*Acipenser transmontanus*) fed larger rations had lower CT_{max} (Verhille et al., 2016). We recommend fasting laboratory animals for an appropriate and consistent duration (species and temperature dependent, typically 24 h minimum) before quantifying CT_{max}. However, we caution readers that further research is needed to assess how consistent and large an effect digestion has on CT_{max}. Independent of digestion, photoperiod and time of day could also affect CT_{max} based on evidence in common killifish (Healy and Schulte, 2012).

Differences in the upper and lower thermal tolerance can occur between different size classes of animals (McKenzie et al., 2021). For example, Di Santo and Lobel (2017) measured CT_{max} and CT_{min} of two goby congeners (*Elacatinus lobeli* and *E. oceanops*) and found larger fish had narrower thermal tolerance (lower CT_{max} and higher CT_{min}). Size effects can vary among life stages. In juvenile zebrafish, larger individuals had a lower CT_{min}, but there was no effect of mass on CT_{min} in adults (De Bonville et al., 2025). Overall, there are numerous examples in the

literature in which body size did or did not affect CT_{max} (e.g., Clark et al., 2017; Messmer et al., 2017; Morgan et al., 2018; Recsetar et al., 2012). Thus, we recommend carefully controlling for body size (statistically or experimentally; also see section 2, Fig. 5).

Social dynamics within trial groups may alter the physiology and thermal tolerance of animals (Gilmour et al., 2005; Currie & Tattersall, 2018). For example, subordinate salmonids had elevated plasma cortisol and reduced CT_{max} compared to dominants when tested in groups (LeBlanc et al., 2011; Bard et al., 2021). This effect may be species-specific: in gregarious lake sturgeon, isolation or conspecific grouping did not alter CT_{max} (Yusishen et al., 2020). Therefore, the social structure of the species should be considered when choosing the appropriate testing method. Pilot experiments can be run to assess the effects of isolation vs. groups on CT_{max} results.

Habituation time to the experimental arena can help reduce stress following animal capture, transport, or transfer from holding tanks. Indeed, Bard et al. (2021) found a link between elevated cortisol and reduced CT_{max} in subordinate rainbow trout (*Oncorhynchus mykiss*). Habituation times reported in the literature range from hours (e.g., overnight) to minutes, with the optimal habituation time likely to be species dependent. Animals should be placed in the animal arena for a standardised duration before thermal ramping begins (and this duration reported in the methods). Similarly, the starting temperature of the trial could impact CT_{max} values (Cicchino et al., 2024). In most cases, the starting temperature should be the same as the holding temperature of the organisms. In studies with different thermal acclimation groups, this either means that different groups must start their CT_{max} trials at different temperatures, or they could start at a common temperature with the understanding that some groups will undergo an abrupt temperature change (and potentially commence a new thermal acclimation trajectory) immediately prior to the CT_{max} trial (e.g., De Bonville et al., 2024; 2025;

555 Stewart et al., 2023). As a result, the time spent in the trial and the accumulated thermal injury
556 during the trial may differ between treatment groups.
557 Reproductive status could affect CT_{max} (Auer et al., 2021; Dahlke et al., 2020; Pörtner &
558 Peck, 2010) but evidence for reduced CT_{max} in spawning fish is the subject of ongoing debate
559 (Pottier et al., 2022). A reduction in CT_{max} did occur in guppies (*Poecilia reticulata*), a live-bearing
560 fish, at a later gestational stage compared to those at an earlier or non-gestational stage (Auer
561 et al., 2021). Thus, it may be worth considering or controlling for variation in reproductive status
562 during experimental design.

563

564 **9. Abiotic factors to consider**

565 Oxygen-limitation has been proposed as a mechanism involved in thermal tolerance, and there
566 are examples in the literature of hypoxia and hyperoxia affecting CT_{max} (Verberk et al., 2016;
567 McArley et al., 2022; Reemeyer & Chapman, 2024). If manipulation of dissolved oxygen is not of
568 interest to the study, a CT_{max} arena should be equipped with sufficient aeration to ensure oxygen
569 remains close to 100% air saturation throughout the trial. Dissolved oxygen should be
570 monitored as % air saturation, because the concentration in mg L^{-1} changes dramatically with
571 warming. Towards the end of a CT_{max} trial, both temperature and stress-induced increases in
572 metabolism can increase the rate of oxygen consumption by animals, potentially leading to
573 declines in dissolved oxygen in the arena if the water volume is low and aeration is insufficient.
574 Conversely, when warming water at 0.3°C min^{-1}, arena water can become supersaturated with
575 oxygen (e.g., 110-120% air saturation). The solubility of oxygen in water decreases as
576 temperature increases, but the rate at which oxygen diffuses out of water is too slow to
577 maintain air saturation at or below 100% when the water is heated quickly. Aeration should
578 minimise both issues, and researchers should use a reliable O_2 sensor to confirm that the O_2

level remains stable (e.g., between 95-105% air saturation) throughout trials in pilot experiments.

CO_2 could build up in arena water during CT_{max} experiments because of animal respiration. This phenomenon has rarely been measured, and thus its potential impacts on CT_{max} estimates are not well understood. A few studies have tested the effect of high CO_2 levels (acute or chronic) on CT_{max}, and typically report no impact (Clark et al., 2017; Frommel et al., 2020; Montgomery et al., 2024). For example, coral reef fish acclimated to an end-of-century CO_2 level of ~1000 µatm had a similar CT_{max} to fish acclimated to a current-day coral reef CO_2 level of ~500 µatm (Clark et al. 2017). CO_2 is >20 times more soluble in water than is O_2, so vigorous aeration is unlikely to prevent its buildup during CT_{max} trials (or the buildup of other waste products like ammonia with similarly high solubility). We encourage more research to understand whether metabolic waste products interact with CT_{max} (e.g., in relation to animal biomass:water volume).

Other abiotic considerations include the water salinity where varying effects have been found on CT_{max} (see Åsheim et al., 2024 for a review of available salinity effects). Some natural habitats have large fluctuations in conductivity (e.g., estuarine and coastal systems), which should be considered. Dissolved organic matter or carbon might influence thermal tolerance through modifications of ion movements across the gills or other epithelial surfaces. Humic substances, for instance, make up a substantial portion of dissolved organic carbon and can bind to biological surfaces (Campbell et al., 1997), altering membrane permeability and transepithelial potential in fish and invertebrate species (Glover & Wood, 2005), which may confound or alter CT_{max}. Similarly, sedimentary turbidity could impact CT_{max} by reducing the efficiency of gill oxygen uptake (Francispillai et al., 2024; Fortin-Hamel and Chapman, 2024).

Toxicants have the potential to act as invisible confounds via contaminated field or facility water, or from the animal arena. Toxicants that could potentially modify CT_{max} include

604 inorganic and organic contaminants such as pharmaceuticals, metals, micro- and nanoplastics,
605 pesticides, nanopesticides, legacy and emerging contaminants, oil and oil co-contaminants, and
606 potential mixtures of all of the above (Lydy & Wissing, 1988; Patra et al., 2009; Khursigara et al.,
607 2019; Chang et al., 2023).
608

609 **10. Field vs. laboratory considerations**

610 The principles involved in designing and setting up a CT_{max} experiment in the field (e.g.,
611 streamside; Leclair et al., 2020; Firth et al., 2021; Stewart et al., 2024) are no different than doing
612 so in the lab in terms of the arena design, warming rates, or endpoints. In the field, a petroleum-
613 powered generator may be used as the power source to run equipment (Fig. S1). Depending on
614 the species and arena water volume, researchers could even conceivably use more powerful
615 propane water heaters (Clark et al., 2012) to provide a supply of warmed water to a larger CT_{max}
616 arena (instead of electric heaters). Warming rates in the field can be affected by variation in air
617 temperature (weather) and site conditions (shade, ground surface). Most of our
618 recommendations above about achieving good warming rates apply here (see section 3).
619 The advantages to the *in situ* CT_{max} approach are threefold: 1) it allows for assessment
620 of different genetically- or phenotypically-distinct wild populations of animals without the need
621 to transport animals to the lab, 2) animals can be released (alive) after the experiment, which
622 can be important for obtaining scientific collection permits for some species, and 3) it may
623 provide 'realism' in terms of animals being in a more physiologically 'wild' state and by exposing
624 animals to the natural, ambient water chemistry to which they are acclimated (see section 9). In
625 pumpkinseed (*L. gibbosus*), hypoxia tolerance changes when wild fish are acclimatised to
626 captivity (Borowiec et al., 2016). Unfortunately, there are few studies explicitly testing how CT_{max}
627 changes with captivity (but see Morgan et al., 2019; Kraskura et al., 2024).

In the field, a key difference from the laboratory (arguably a disadvantage) is that the animals are (typically) caught directly from the wild and then exposed to a CT_{max} test minutes or hours later (Morgan et al., 2019). As a result, the animals are likely to be in more variable physiological states, or states of stress, than animals acclimatised to captivity. Wild-caught animals also have unknown individual thermal and ecological history and may be in varying states of digestion and gut fullness, differ in parasite burden (see section 8) or health status. This added variation may necessitate larger sample sizes to accurately estimate CT_{max}. Similarly, field sites may vary in their water chemistry or toxicants. For example, Mottola et al. (2022) simulated a storm, resulting in relevant increases to environmental copper within a field collection site, which in turn increased thermal tolerance male three-spined stickleback by 1.5°C (*Gasterosteus aculeatus*). Researchers assessing thermal tolerance across a wide range of locations may therefore want to consider partnering with environmental chemists to quantify water chemistry and consider surrounding land use (e.g., agricultural pesticides).

CT_{max} data should always be reported relative to the acclimation temperature of the animals (see section 12). Acclimation temperatures are more difficult to quantify in the field than in the lab. We recommend deploying temperature loggers at each field site at least five days prior to CT_{max} (ideally 1-2 months earlier; Reemeyer et al., 2024; Stewart et al., 2024). The logged temperatures can then be used to estimate mean acclimation temperatures for some time window preceding the date of CT_{max} (e.g., 5-10 days). If animals are caught from a thermally heterogeneous environment (e.g., a thermally stratified lake), a single temperature logger deployed may not represent the diversity of acclimation temperatures among the animals tested. Deployment of multiple loggers could help understand the thermal habitats available, capturing the range of possible acclimation temperatures.

11. Measuring CT_{min}

Critical thermal minimum (CT_{min}) is less commonly measured than is CT_{max} but is conceptually and methodologically similar. However, there are methodological considerations specific to CT_{min}. The assessment of the LOE endpoint (section 5) at CT_{min} is complicated by the cold-induced reductions in activity common among aquatic ectotherms, especially lethargic or winter-dormant species (Reeve et al., 2022). Some species do lose equilibrium at temperatures approaching CT_{min}. In others, a response to touch stimuli at low temperatures when animals become minimally active can be used as an endpoint, especially for animals with CT_{min} values close to 0°C (Ford & Beitinger, 2005). Polar species may not have a CT_{min}. Instead, they function down to the temperature at which ice forms inside their bodies. Antarctic fishes maintain high activity and aerobic scope with a body temperature equal to that of the freezing point of seawater (-1.86°C; Brijs et al., 2020; DeVries & Cheng, 2005).

Cooling water is more difficult than warming it and chillers are more costly than are heaters (the cost per watt for cooling is up to 10× that of heating). Water chillers available in the consumer aquatics industry are not capable of cooling water below 2-3°C, which is typically not an issue for warm-water species with CT_{min} above this range. However, many temperate animals are likely to have a CT_{min} approaching 0°C, especially when acclimated to moderate or cool temperatures. The cooling power of the chiller must account for the volume of the test arena and the temperature difference between the arena and environment. These are the same considerations as with heating, but more challenging in practice. Extra insulation of the test arena may be needed, and setting up the CT_{min} arena in a temperature-controlled room with cold air can help.

For species with very low CT_{min}, even more complex recirculating water bath chillers are needed (potentially paired with conventional chillers to more easily reach 2-3°C). A glycol-water bath can be ultra-cooled (e.g., -20°C) to achieve temperatures reaching the freezing point of

fresh or seawater (though freezing of live vertebrates should be avoided). If animals are small, they may be cooled directly in small open-topped arenas submerged in a glycol-water bath (the arenas must be closed because ethylene glycol is toxic to animals; Hymel et al., 2002). However, this limits throughput given the relatively small working area of recirculating water baths. These setups allow good control over thermal ramping by simply changing the setpoint of the water bath. However, there is a risk of contamination of the CT_{min} arena water with the toxic glycol-water mixture, and they are prone to freezing the water without sufficient stirring.

To avoid contamination, water from the test arena can be circulated through a heat exchange coil within an external glycol bath (Fangue et al., 2006). The flow rate should be high enough to avoid water freezing within the heat exchange coil, and slow enough to ensure sufficient heat exchange (i.e., water returning to the test arena should be as close to freezing as possible). Note that switching on and off the recirculating pump is problematic because water inside the coil will freeze without flow; control over cooling rate can instead be achieved by using a pump with adjustable flow. Decreasing the setpoint on the recirculating water bath chiller is also possible, but it may not cool quickly enough to match the required thermal ramp. Pre-frozen ice blocks (or crushed ice made using non-chlorinated water) can be added to the CT_{min} equipment section to keep cooling rates more stable as temperatures approach freezing. For saltwater species, freshwater ice would have to be kept physically separated (e.g., in sealed bags) from the arena water to avoid changing the arena's salinity.

12. Reporting

To improve reproducibility, transparency, and to facilitate evidence synthesis, researchers should strive for detailed and consistent reporting of their methods (Percie du Sert et al., 2020). Inspired by Killen et al. (2021), who provided a detailed guide on reporting methods for aquatic

701 respirometry, here we provide a checklist for reporting on CT_{max} experiments, along with a

702 downloadable (fillable) version that authors can use (Table 1; see supplementary file).

703

704 **Table 1.** Considerations when preparing for, conducting, and reporting on critical thermal limit
705 experiments. A fillable version of this table has been included as a downloadable,
706 supplementary file.

	Description	Rationale
Pre-trial		
1	Information on feeding history of animals (fasted or not, duration of fasting)	Whether or not animals were fasted prior to testing should be explicit. For field CT_{max} where the satiation states of animals are unknown and likely to vary, that caveat should be noted. Most laboratory studies fast animals for *ca*. 24 h prior to CT_{max}.
2	Habituation time to the laboratory (or time since capture for field studies)	A change of environment could cause endocrine responses with unknown consequences for CT_{max}. Behavioural changes with lab habituation (e.g., establishment of social hierarchies) could conceivably affect CT_{max}. Report whether animals were raised in captivity or fully domesticated.
3	Thermal acclimation (temperature and duration)	Thermal acclimation has strong effects on CT_{max} (Chrétien & Chapman, 2016; Fangue et al., 2014), so authors should clearly report the temperature (and duration) of acclimation, including whether temperatures were stable or fluctuating. Additional information about long term prior thermal exposures is also valuable. If known, provide basic context about the thermal biology of the species or population (e.g., optimal temperature).
4	CT_{max} arena dimensions and water volume	Arena size and water volume can influence ramping rate, accumulation of CO_2 and nitrogenous waste, and subject behaviour (Stewart & Allen, 2014).
5	Body mass of animals at time of CT_{max}	CT_{max} is correlated with body mass in some animals (Ospina & Mora, 2004; Bartlett et al., 2022). Studies should measure body mass of each animal immediately before or after CT_{max} (mass values linked to each CT_{max} value) and not simply report a mean mass for the whole sample.
6	Total animal biomass: water volume ratio	The ratio of animal biomass to water volume will influence the accumulation of CO_2 and nitrogenous waste in the arena. Low animal biomass:volume ratios are especially prone to increases in DO at fast ramping rates.
7	Details on source water	Studies should indicate whether the water used was filtered or treated in any way, as well as the source (i.e., natural lake, river, or seawater, or dechlorinated municipal tap water). For field-

		based CT_{max}, indicate whether the water was collected directly from the sampling site.
8	Type of temperature probe and details on data logging	Different probe types have different levels of precision and accuracy (provide these values). Indicate the equipment used for scoring CT_{max} and the equipment used for logging arena temperatures (if these were separate devices).
9	Type, number, and wattage of heaters	Different heater types have different efficacies based on their design. Indicate if and how water was circulated through/around the heating element(s).
10	Life stage and sex (if known) of the test animals	CT_{max} may differ across life stages and/or between the sexes (Cowan et al., 2023; Wheeler et al., 2022).
11	Information on treatments/health metrics of wild animals	Parasites and pathogens can influence thermal tolerance and should be considered (Chrétien et al., 2023; De Bonville et al., 2024). Studies should mention whether animals are treated, if infected/sick individuals are excluded or if parasite load was quantified. If health metrics are unknown, that should be stated.
12	Time of the day that trials were run, and photoperiod	Diurnal cycles could potentially influence thermal tolerance (Healy and Schulte, 2012).
13	Habituation time in the CT_{max} arena before trial start	Times can vary between studies and could affect stress and behaviour of the animals depending on density.
14	Temperature at which the trials were started	Variations in starting temperature can affect the accumulation of thermal injury over the course of a CT_{max} trial (Cicchino et al., 2024).
15	Details of pilot experiments used to refine main trials	Indicate whether pilot trials performed on subsets of animals to refine thermal ramping profiles, whether intratissue thermocouples were used determine thermal inertia and appropriate thermal ramping rates.
During trial		
16	Warming rate with measures of variation (e.g., standard deviation or range).	Warming rates should be monitored and calculated throughout the trial. Best practice is to use temperature loggers in the test arena that can be compared with digital thermometer recordings during trial runs (e.g., De Bonville et al., 2024). The logger outputs can be plotted (Fig. 6) and included (for each trial) as supplemental material (with R^2 values for warming rates).

17	Details on how consistent ramping rates were achieved	In some cases, heaters are simply switched on and left on for the entire trial, in others, they are manually switched on/off throughout the trial (Morgan et al., 2018; Stewart et al., 2023; Ern and Jutfelt, 2024). In either case, supplementary plots of warming rates (i.e., raw data; Fig. 6) will allow readers to assess the consistency in warming rates.
18	Indicate whether CT_{max} arena was visually shielded from external disturbance.	External disturbance may cause agitation and stress, potentially altering CT_{max} (McDonnell & Chapman, 2015).
19	Endpoints	Critical thermal endpoints assessed behaviourally can vary qualitatively among study organisms (Ziegeweid et al., 2008; Morgan et al., 2018). Providing a clear description of the endpoint used in the experiment in terms of behaviours observed (i.e., loss of equilibrium (LOE), erratic swimming, loss of righting response [LRR]), as well as the duration of this behaviour prior to removal of the individual (i.e., LOE for > 5 s) can facilitate replication.
20	Water quality monitoring	Indicate whether water in the test arena was aerated during the trials, and whether dissolved oxygen or other water parameters (e.g., salinity, conductivity, pH, ammonia) were measured.
21	Numbers of animals and replicate trials	State the number of animals per CT_{max} trial, the number of replicate trials per treatment, and the resulting total n per treatment. If animals were kept in separate tanks for treatments prior to the trial, state which tank(s) they were kept in.
After trial		
22	Duration of all trials combined.	Should ideally be reported as the start and end dates of data collection for the study. Animals may habituate or acclimate to the laboratory setting, which could affect behaviour and CT_{max}. Body mass or life-stage may also change over time.
23	Recovery duration and temperature	Survival is typically monitored for 30 min - 1 h following the trial, but can be reported over longer durations as well (24 h or more). Recovery temperatures could potentially affect survival if they are either too low (cold shock) or too high (continued thermal stress).
24	Rates of survival	CT_{max} is defined as a non-lethal procedure. The percent recovery of all individuals in a trial following their return to cooler temperatures should be verified and reported.
25	Arena cleaning/ maintenance	CT_{max} arenas should be fully drained and refilled between trials. Indicate if any other cleaning or maintenance tasks took place between trials.
Data handling and statistics		

26	Assignment of a CT_{max} value	The source of the CT_{max} value chosen must be specified, particularly when multiple pieces of equipment are used to record temperature, such as a handheld thermometer and logger(s) in the water bath.
27	Observer bias	Report whether any measures were taken to minimise observer bias (e.g., any form of masking/blinding and if so how that was achieved and if it applied to masking/blinding of both treatment and/or temperature and ramping), and whether CT_{max} trials were video recorded (and if videos are uploaded to a repository). Blinding increases the reliability of the experiment (Holman et al., 2015).
28	Approach to ensuring scoring consistency	Ideally a single observer is used across treatments, after having gone through one or two 'test-run' pilot experiments. Otherwise, multiple observers need to agree on endpoint criteria that are as objective as possible, and should be trained together during pilot trials. Video footage can accompany publications to enable replication.

Summary and knowledge gaps

Here we have provided methodological advice for designing, carrying out, and reporting on CT_{max} (and CT_{min}) experiments. A common theme is that researchers should always conduct careful pilot experiments after thinking through recommendations in sections 1-12 (above). Pilot experiments are especially important whenever using a new CT_{max} arena, new species, new life stage, or new warming rate. Our hope is that the guidance will be particularly useful for researchers developing a new interest in CT_{max}.

While there is ample literature on thermal biology to provide evidence-based recommendations for CT_{max} experiments, knowledge gaps remain that we hope to see addressed. These include, but are not limited to:

a) The tissue-specific physiological mechanism(s) that cause loss of motor function at extreme temperatures at fast (e.g., 0.3°C min^{-1}) and slow (e.g., 1°C $hour^{-1}$) rates of warming.

b) The confounding roles of biotic factors including stress caused by capture and handling, confinement, or agonistic social interactions, as well as the impacts of digestion and reproduction — these factors could have implications for interpreting CT_{max} data in an ecological context.

c) The effect of body mass on the temperature lag of different tissues (i.e., heat transfer time). A robust dataset collected across a range of body sizes and species is lacking. Such data could be used by researchers to make decisions on what warming rates are appropriate for the range of body sizes in their study.

d) The effects of varying warming rates on CT_{max} (and CT_{min}). Thermal biologists would benefit from more empirical data (across species) from experiments focused on this question. This could include interactions between warming rate, body mass, and acclimation temperature.

e) Acclimation dynamics. Surprisingly few studies have focused on the dynamics of thermal acclimation such as quantifying the time needed for CT_{max} to exhibit 'full' acclimation to a new temperature or the duration after which CT_{max} starts changing (but see Bennett et al., 1998; Fangue et al., 2014; De Bonville et al., 2025; Fu et al., 2018; Stewart et al., 2023), and it seems likely that acclimation is quicker in some species than in others (Burton & Einum, 2025). For designing lab experiments and quantifying acclimation in field studies, more data about acclimation dynamics would be useful.

f) The physiological difference and relative ecological meaning of 'typical' CT_{max} and CT_{min} experiments when compared against alternative approaches in which animals are forced to exercise during thermal ramping, such as the critical thermal maximum for swimming (CT_{swim}; Blasco et al., 2020).

g) More controlled studies that assess the difference in CT_{max} between animals measured in situ (field) vs. in animals acclimatised to the laboratory for varying durations.

747 Ultimately, addressing these knowledge gaps will help improve future CT_{max} studies and help
748 interpret the vast literature on CT_{max}. CT_{max} will likely continue to be a popular metric in animal
749 biology (Fig. 1), providing value for comparative ecophysiology, evolutionary biology, risk
750 assessments, and species distribution modelling. Ensuring CT_{max} data are as precise and
751 reliable as possible, through sound experimental approaches and reporting, will help optimise
752 their value.
753
754 **Acknowledgments**
755 We thank the 2024 FIN club for the opportunity to write this paper. This work was funded
756 primarily by an ERC consolidator grant (CLIMEVOLVE) to Jutfelt. Raby, Blewett, Binning and
757 Speers-Roesch are supported by Natural Sciences and Engineering Research Council of Canada
758 (NSERC) Discovery Grants. Stewart and Reid were supported by NSERC doctoral scholarships.
759 Hoots and Kuchenmüller were supported by PhD scholarships through Deakin University,
760 Australia. Ern was supported by Marie Skłodowska-Curie Actions (893895). Clark was
761 supported by an Australian Research Council Future Fellowship (FT180100154) funded by the
762 Australian Government.
763
764 **Author Contributions**
765 Raby conceived of the idea for this paper; Stewart conducted the data analysis; all authors
766 contributed to drafting and revising the manuscript.
767
768 **Data availability**
769 Data and analysis code are publicly archived on figshare:
770 https://doi.org/10.6084/m9.figshare.28319774.v1
771

References

Andreassen, A. H., Hall, P., Khatibzadeh, P., Jutfelt, F., & Kermen, F. (2022). Brain dysfunction during warming is linked to oxygen limitation in larval zebrafish. *Proceedings of the National Academy of Sciences*, *119*(39), e2207052119. https://doi.org/10.1073/pnas.2207052119

Åsheim, E. R., Andreassen, A. H., Morgan, R., Silvestre, M., & Jutfelt, F. (2024). Acute warming tolerance (CT_{max}) in zebrafish (*Danio rerio*) appears unaffected by changes in water salinity. *PeerJ*, *12*, e17343. https://doi.org/10.7717/peerj.17343

Åsheim, Er., Andreassen, Ah., Morgan, R., & Jutfelt, F. (2020). Rapid-warming tolerance correlates with tolerance to slow warming but not growth at non-optimal temperatures in zebrafish. *Journal of Experimental Biology*, jeb.229195. https://doi.org/10.1242/jeb.229195

Auer, S. K., Agreda, E., Chen, A. H., Irshad, M., & Solowey, J. (2021). Late-stage pregnancy reduces upper thermal tolerance in a live-bearing fish. *Journal of Thermal Biology*, *99*, 103022. https://doi.org/10.1016/j.jtherbio.2021.103022

Bard, B., Dodge, A., Joyce, W., Lawrence, M., Cooke, S. J., & Gilmour, K. M. (2021). Elevated cortisol lowers thermal tolerance but results in limited cardiac remodelling in rainbow trout (*Oncorhynchus mykiss*) experiencing chronic social stress. *Journal of Experimental Biology*, jeb.238683. https://doi.org/10.1242/jeb.238683

Bear, E. A., McMahon, T. E., & Zale, A. V. (2007). Comparative Thermal Requirements of Westslope Cutthroat Trout and Rainbow Trout: Implications for Species Interactions and Development of Thermal Protection Standards. *Transactions of the American Fisheries Society*, *136*(4), 1113–1121. https://doi.org/10.1577/T06-072.1

Becker, C. D., & Genoway, R. G. (1979). Evaluation of the critical thermal maximum for determining thermal tolerance of freshwater fish. *Environmental Biology of Fishes*, *4*(3),

245–256. https://doi.org/10.1007/BF00005481

Beitinger, T. L., Bennett, W. A., & McCauley, R. W. (2000). Temperature Tolerances of North American Freshwater Fishes Exposed to Dynamic Changes in Temperature. *Environmental Biology of Fishes*, *58*(3), 237–275. https://doi.org/10.1023/A:1007676325825

Bennett, W. A., McCauley, R. W., & Beitinger, T. L. (1998). Rates of Gain and Loss of Heat Tolerance in Channel Catfish. *Transactions of the American Fisheries Society*, *127*(6), 1051–1058. https://doi.org/10.1577/1548-8659(1998)127<1051:ROGALO>2.0.CO;2

Blasco, F. R., Esbaugh, A. J., Killen, S. S., Rantin, F. T., Taylor, E. W., & McKenzie, D. J. (2020). Using aerobic exercise to evaluate sub-lethal tolerance of acute warming in fishes. *Journal of Experimental Biology*, *223*(9), jeb218602. https://doi.org/10.1242/jeb.218602

Borowiec, B. G., Crans, K. D., Khajali, F., Pranckevicius, N. A., Young, A., & Scott, G. R. (2016). Interspecific and environment-induced variation in hypoxia tolerance in sunfish. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *198*, 59–71. https://doi.org/10.1016/j.cbpa.2016.04.010

Bruneaux, M., Visse, M., Gross, R., Pukk, L., Saks, L., & Vasemägi, A. (2017). Parasite infection and decreased thermal tolerance: Impact of proliferative kidney disease on a wild salmonid fish in the context of climate change. *Functional Ecology*, *31*(1), 216–226. https://doi.org/10.1111/1365-2435.12701

Burghardt, G. M., Bartmess-LeVasseur, J. N., Browning, S. A., Morrison, K. E., Stec, C. L., Zachau, C. E., & Freeberg, T. M. (2012). Perspectives – Minimizing Observer Bias in Behavioral Studies: A Review and Recommendations. *Ethology*, *118*(6), 511–517. https://doi.org/10.1111/j.1439-0310.2012.02040.x

Burton, T., & Einum, S. (2025). High capacity for physiological plasticity occurs at a slow rate in ectotherms. *Ecology Letters*, *28*(1), e70046. https://doi.org/10.1111/ele.70046

822 Campbell, P. G. C., Twiss, M. R., & Wilkinson, K. J. (1997). Accumulation of natural organic
823 matter on the surfaces of living cells: Implications for the interaction of toxic solutes
824 with aquatic biota. *Canadian Journal of Fisheries and Aquatic Sciences*, *54*, 2543-2554.
825 Challice, A. R., Milne, S. W., & Ridgway, M. S. (2019). Does habitat occupancy by lake trout and
826 lake whitefish in large lakes match published thermal habitat envelopes? *Ecology of*
827 *Freshwater Fish*, *28*(4), 611–623. https://doi.org/10.1111/eff.12479
828 Chang, M., Li, M., Xu, W., Li, X., Liu, J., Stoks, R., & Zhang, C. (2023). Microplastics increases the
829 heat tolerance of Daphnia magna under global warming via hormetic effects.
830 *Ecotoxicology and Environmental Safety*, *249*, 114416.
831 https://doi.org/10.1016/j.ecoenv.2022.114416
832 Chasse, S. D., Stewart, E. M. C., & Raby, G. D. (2025). Repeatability of critical thermal maximum
833 (CT_{max}) in two freshwater ectotherms across contexts. *Journal of Thermal Biology*, *127*,
834 104049. https://doi.org/10.1016/j.jtherbio.2025.104049
835 Chrétien, E., De Bonville, J., Guitard, J., Binning, S. A., Melis, É., Kack, A., Côté, A., Gradito, M.,
836 Papillon, A., Thelamon, V., Levet, M., & Barou-Dagues, M. (2023). Few studies of wild
837 animal performance account for parasite infections: A systematic review. *Journal of*
838 *Animal Ecology*, *92*(4), 794–806. https://doi.org/10.1111/1365-2656.13864
839 Clark, T. D., Donaldson, M. R., Pieperhoff, S., Drenner, S. M., Lotto, A., Cooke, S. J., Hinch, S. G.,
840 Patterson, D. A., & Farrell, A. P. (2012). Physiological Benefits of Being Small in a
841 Changing World: Responses of Coho Salmon (*Oncorhynchus kisutch*) to an Acute
842 Thermal Challenge and a Simulated Capture Event. *PLOS ONE*, *7*(6), e39079.
843 https://doi.org/10.1371/journal.pone.0039079
844 Clark, T. D., Roche, D. G., Binning, S. A., Speers-Roesch, B., & Sundin, J. (2017). Maximum
845 thermal limits of coral reef damselfishes are size-dependent and resilient to near-future
846 ocean acidification. *Journal of Experimental Biology*, jeb.162529.

https://doi.org/10.1242/jeb.162529

Comte, L., & Olden, J. D. (2017). Evolutionary and environmental determinants of freshwater fish thermal tolerance and plasticity. *Global Change Biology*, *23*(2), 728–736. https://doi.org/10.1111/gcb.13427

Cowan, Z.-L., Andreassen, A. H., De Bonville, J., Green, L., Binning, S. A., Silva-Garay, L., Jutfelt, F., & Sundin, J. (2023). A novel method for measuring acute thermal tolerance in fish embryos. *Conservation Physiology*, *11*(1), coad061. https://doi.org/10.1093/conphys/coad061

Cowles, R. B., & Bogert, C. M. (1944). A preliminary study of the thermal requirements of desert reptile. *Bulletin of the American Museum of Natural History*, *83*, 261–296.

Cuenca Cambronero, M., Beasley, J., Kissane, S., & Orsini, L. (2018). Evolution of thermal tolerance in multifarious environments. *Molecular Ecology*, *27*(22), 4529–4541. https://doi.org/10.1111/mec.14890

Currie, S., & Tattersall, G. J. (2018). Social cues can push amphibious fish to their thermal limits. *Biology Letters*, *14*(10), 20180492. https://doi.org/10.1098/rsbl.2018.0492

Dahlke, F. T., Wohlrab, S., Butzin, M., & Pörtner, H.-O. (2020). Thermal bottlenecks in the life cycle define climate vulnerability of fish. *Science*, *369*(6499), 65–70. https://doi.org/10.1126/science.aaz3658

Dallas, H. F., & Ketley, Z. A. (2011). Upper thermal limits of aquatic macroinvertebrates: Comparing critical thermal maxima with 96-LT50 values. *Journal of Thermal Biology*, *36*(6), 322–327. https://doi.org/10.1016/j.jtherbio.2011.06.001

De Bonville, J., Andreassen, A. H., Cowan, Z.-L., Silva-Garay, L., Leeuwis, R. H. J., Åsheim, E. R., Speers-Roesch, B., Raby, G. D., Binning, S. A., & Jutfelt, F. (2025). Dynamics of thermal tolerance plasticity across fish species and life stages. *Journal of Thermal Biology*, *127*, 104024. https://doi.org/10.1016/j.jtherbio.2024.104024

De Bonville, J., Côté, A., & Binning, S. A. (2024). Thermal tolerance and survival are modulated by a natural gradient of infection in differentially acclimated hosts. *Conservation Physiology*, *12*(1), coae015. https://doi.org/10.1093/conphys/coae015

Desforges, J. E., Birnie-Gauvin, K., Jutfelt, F., Gilmour, K. M., Eliason, E. J., Dressler, T. L., McKenzie, D. J., Bates, A. E., Lawrence, M. J., Fangue, N., & Cooke, S. J. (2023). The ecological relevance of critical thermal maxima methodology for fishes. *Journal of Fish Biology*, *102*(5), 1000–1016. https://doi.org/10.1111/jfb.15368

Di Santo, V., & Lobel, P. S. (2017). Body size and thermal tolerance in tropical gobies. *Journal of Experimental Marine Biology and Ecology*, *487*, 11–17. https://doi.org/10.1016/j.jembe.2016.11.007

Elliott, J. M., & Elliott, J. A. (1995). The effect of the rate of temperature increase on the critical thermal maximum for parr of Atlantic salmon and brown trout. *Journal of Fish Biology*, *47*(5), 917–919. https://doi.org/10.1111/j.1095-8649.1995.tb06014.x

Ern, R., Andreassen, A. H., & Jutfelt, F. (2023). Physiological Mechanisms of Acute Upper Thermal Tolerance in Fish. *Physiology*, *38*(3), 141–158. https://doi.org/10.1152/physiol.00027.2022

Ern, R., & Jutfelt, F. (2024). The OptoReg system: A simple and inexpensive solution for regulating water oxygen. *Conservation Physiology*, *12*(1), coae024. https://doi.org/10.1093/conphys/coae024

Fangue, N. A., Hofmeister, M., & Schulte, P. M. (2006). Intraspecific variation in thermal tolerance and heat shock protein gene expression in common killifish, *Fundulus heteroclitus*. *Journal of Experimental Biology*, *209*(15), 2859–2872. https://doi.org/10.1242/jeb.02260

Fangue, N. A., Wunderly, M. A., Dabruzzi, T. F., & Bennett, W. A. (2014). Asymmetric thermal acclimation responses allow sheepshead minnow *Cyprinodon variegatus* to cope with

rapidly changing temperatures. *Physiological and Biochemical Zoology*, *87*(6), 805–816. https://doi.org/10.1086/678965

Firth, B. L., Drake, D. A. R., & Power, M. (2021). Seasonal and environmental effects on upper thermal limits of eastern sand darter (*Ammocrypta pellucida*). *Conservation Physiology*, *9*(1), coab057. https://doi.org/10.1093/conphys/coab057

Ford, T., & Beitinger, T. L. (2005). Temperature tolerance in the goldfish, *Carassius auratus*. *Journal of Thermal Biology*, *30*(2), 147–152. https://doi.org/10.1016/j.jtherbio.2004.09.004

Fortin-Hamel, L., & Chapman, L. J. (2024). Interactive effects of sedimentary turbidity and elevated water temperature on the Pugnose Shiner (*Miniellus anogenus*), a threatened freshwater fish. *Conservation Physiology*, *12*(1), coae053. https://doi.org/10.1093/conphys/coae053

Fu, K.-K., Fu, C., Qin, Y.-L., Bai, Y., & Fu, S.-J. (2018). The thermal acclimation rate varied among physiological functions and temperature regimes in a common cyprinid fish. *Aquaculture*, *495*, 393–401. https://doi.org/10.1016/j.aquaculture.2018.06.015

Galbreath, P. F., Adams, N. D., & Martin, T. H. (2004). Influence of heating rate on measurement of time to thermal maximum in trout. *Aquaculture*, *241*(1–4), 587–599. https://doi.org/10.1016/j.aquaculture.2004.08.004

Geerts, A. N., Vanoverbeke, J., Vanschoenwinkel, B., Van Doorslaer, W., Feuchtmayr, H., Atkinson, D., Moss, B., Davidson, T. A., Sayer, C. D., & De Meester, L. (2015). Rapid evolution of thermal tolerance in the water flea Daphnia. *Nature Climate Change*, *5*(7), 665–668. https://doi.org/10.1038/nclimate2628

Gilmour, K. M., Wilson, R. W., & Sloman, K. A. (2005). The integration of behaviour into comparative physiology. *Physiological and Biochemical Zoology*, *78*(5), 669–678. https://doi.org/10.1086/432144

Giomi, F., Barausse, A., Duarte, C. M., Booth, J., Agusti, S., Saderne, V., Anton, A., Daffonchio, D., & Fusi, M. (2019). Oxygen supersaturation protects coastal marine fauna from ocean warming. *Science Advances*, *5*(9), eaax1814. https://doi.org/10.1126/sciadv.aax1814

Glover, C. N., & Wood, C. M. (2005). The disruption of *Daphnia magna* sodium metabolism by humic substances: mechanism of action and effect of humic substance source. *Physiological and Biochemical Zoology*, *78*(6), 1005–1016. https://doi.org/10.1086/432858

Gomez Isaza, D. F., Cramp, R. L., Smullen, R., Glencross, B. D., & Franklin, C. E. (2019). Coping with climatic extremes: Dietary fat content decreased the thermal resilience of barramundi (*Lates calcarifer*). *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *230*, 64–70. https://doi.org/10.1016/j.cbpa.2019.01.004

Grinder, R. M., Bassar, R. D., & Auer, S. K. (2020). Upper thermal limits are repeatable in Trinidadian guppies. *Journal of Thermal Biology*, *90*, 102597. https://doi.org/10.1016/j.jtherbio.2020.102597

Gunderson, A. R. (2024). Disentangling physiological and physical explanations for body size-dependent thermal tolerance. *Journal of Experimental Biology*, *227*(5), jeb245645. https://doi.org/10.1242/jeb.245645

Haney, D. C., & Walsh, S. J. (2003). Influence of Salinity and Temperature on the Physiology of *Limia melanonotata* (Cyprinodontiformes: Poeciliidae): A Search for Abiotic Factors Limiting Insular Distribution in Hispaniola. *Caribbean Journal of Science*, *3*, 327–337.

Hasnain, S. S., Shuter, B. J., & Minns, C. K. (2013). Phylogeny influences the relationships linking key ecological thermal metrics for North American freshwater fish species. *Canadian Journal of Fisheries and Aquatic Sciences*, *70*(7), 964–972. https://doi.org/10.1139/cjfas-2012-0217

947 Healy, T. M., & Schulte, P. M. (2012). Factors affecting plasticity in whole-organism thermal
948 tolerance in common killifish (*Fundulus heteroclitus*). *Journal of Comparative Physiology*
949 *B*, *182*(1), 49–62. https://doi.org/10.1007/s00360-011-0595-x
950 Holland, O. J., Smythe, C., Clark, T. D., Ragg, N. L. C., Mondon, J., Corbett, P., & Miller, A. D.
951 (2024). Size-dependent thermal limits in Australian hybrid abalone: Implications for
952 productivity shifts with ocean warming. *Reviews in Fish Biology and Fisheries*, *34*(1),
953 271–291. https://doi.org/10.1007/s11160-023-09812-4
954 Holman, L., Head, M. L., Lanfear, R., & Jennions, M. D. (2015). Evidence of experimental bias in
955 the life sciences: why we need blind data recording. *PLOS Biology*, *13*(7), e1002190.
956 https://doi.org/10.1371/journal.pbio.1002190
957 Hymel, M. K., Baltz, D. M., Chesney, E. J., Tarr, M. A., & Kolok, A. S. (2002). Swimming
958 performance of juvenile Florida Pompano exposed to ethylene glycol. *Transactions of*
959 *the American Fisheries Society*, *131*(6), 1152–1163. https://doi.org/10.1577/1548-
960 8659(2002)131<1152:SPOJFP>2.0.CO;2
961 Jørgensen, L. B., Malte, H., Ørsted, M., Klahn, N. A., & Overgaard, J. (2021). A unifying model to
962 estimate thermal tolerance limits in ectotherms across static, dynamic and fluctuating
963 exposures to thermal stress. *Scientific Reports*, *11*(1), 12840.
964 https://doi.org/10.1038/s41598-021-92004-6
965 Jutfelt, F., Roche, D. G., Clark, T. D., Norin, T., Binning, S. A., Speers-Roesch, B., Amcoff, M.,
966 Morgan, R., Andreassen, A. H., & Sundin, J. (2019). Brain cooling marginally increases
967 acute upper thermal tolerance in Atlantic cod. *Journal of Experimental Biology*,
968 jeb.208249. https://doi.org/10.1242/jeb.208249
969 Karp, N. A., Pearl, E. J., Stringer, E. J., Barkus, C., Ulrichsen, J. C., & Percie Du Sert, N. (2022). A
970 qualitative study of the barriers to using blinding in in vivo experiments and suggestions
971 for improvement. *PLOS Biology*, *20*(11), e3001873.

https://doi.org/10.1371/journal.pbio.3001873

Khursigara, A. J., Ackerly, K. L., & Esbaugh, A. J. (2019). Oil toxicity and implications for environmental tolerance in fish. *Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology*, *220*, 52–61. https://doi.org/10.1016/j.cbpc.2019.03.003

Killen, S. S., Christensen, E. A. F., Cortese, D., Závorka, L., Norin, T., Cotgrove, L., Crespel, A., Munson, A., Nati, J. J. H., Papatheodoulou, M., & McKenzie, D. J. (2021). Guidelines for reporting methods to estimate metabolic rates by aquatic intermittent-flow respirometry. *Journal of Experimental Biology*, *224*(18), jeb242522. https://doi.org/10.1242/jeb.242522

Kitagawa, T., & Kimura, S. (2006). An alternative heat-budget model relevant to heat transfer in fishes and its practical use for detecting their physiological thermoregulation. *Zoological Science*, *23*(12), 1065–1071. https://doi.org/10.2108/zsj.23.1065

Kochhann, D., Sarmento, C. G., De Oliveira, J. C., Queiroz, H. L., Val, A. L., & Chapman, L. J. (2021). Take time to look at the fish: Behavioral response to acute thermal challenge in two Amazonian cichlids. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology*, *335*(9–10), 735–744. https://doi.org/10.1002/jez.2541

Kraskura, K., Anderson, C. E., & Eliason, E. J. (2024). Pairing lab and field studies to predict thermal performance of wild fish. *Journal of Thermal Biology*, *119*, 103780. https://doi.org/10.1016/j.jtherbio.2023.103780

LeBlanc, S., Middleton, S., Gilmour, K. M., & Currie, S. (2011). Chronic social stress impairs thermal tolerance in the rainbow trout (*Oncorhynchus mykiss*). *Journal of Experimental Biology*, *214*(10), 1721–1731. https://doi.org/10.1242/jeb.056135

Leclair, A. T. A., Drake, D. A. R., Pratt, T. C., & Mandrak, N. E. (2020). Seasonal variation in thermal tolerance of redside dace *Clinostomus elongatus*. *Conservation Physiology*, *8*(1), coaa081. https://doi.org/10.1093/conphys/coaa081

Lutterschmidt, W. I., & Hutchison, V. H. (1997). The critical thermal maximum: History and

997 critique. *Canadian Journal of Zoology*, *75*(10), 1561–1574. https://doi.org/10.1139/z97-
998 783
999 Lutterschmidt, W. I., Schaefer, J. F., & Fiorillo, R. A. (2007). The ecological significance of
1000 helminth endoparasites on the physiological performance of two sympatric fishes.
1001 *Comparative Parasitology*, *74*(2), 194–203. https://doi.org/10.1654/4248.1
1002 Lydy, M. J., & Wissing, T. E. (1988). Effect of sublethal concentrations of copper on the critical
1003 thermal maxima (CTmax) of the fantail (*Etheostoma flabellare*) and johnny (*E. nigrum*)
1004 darters. *Aquatic Toxicology*, *12*(4), 311–321. https://doi.org/10.1016/0166-
1005 445X(88)90059-8
1006 McArley, T. J., Morgenroth, D., Zena, L. A., Ekström, A. T., & Sandblom, E. (2022). Experimental
1007 hyperoxia (O_2 supersaturation) reveals a gill diffusion limitation of maximum aerobic
1008 performance in fish. *Biology Letters*, *18*(11), 20220401.
1009 https://doi.org/10.1098/rsbl.2022.0401
1010 McDonnell, L. H., & Chapman, L. J. (2015). At the edge of the thermal window: Effects of
1011 elevated temperature on the resting metabolism, hypoxia tolerance and upper critical
1012 thermal limit of a widespread African cichlid. *Conservation Physiology*, *3*(1), cov050.
1013 https://doi.org/10.1093/conphys/cov050
1014 McKenzie, D. J., Zhang, Y., Eliason, E. J., Schulte, P. M., Claireaux, G., Blasco, F. R., Nati, J. J. H.,
1015 & Farrell, A. P. (2021). Intraspecific variation in tolerance of warming in fishes. *Journal of*
1016 *Fish Biology*, *98*(6), 1536–1555. https://doi.org/10.1111/jfb.14620
1017 Messmer, V., Pratchett, M. S., Hoey, A. S., Tobin, A. J., Coker, D. J., Cooke, S. J., & Clark, T. D.
1018 (2017). Global warming may disproportionately affect larger adults in a predatory coral
1019 reef fish. *Global Change Biology*, *23*(6), 2230–2240. https://doi.org/10.1111/gcb.13552
1020 Mora, C., & Maya, M. F. (2006). Effect of the rate of temperature increase of the dynamic
1021 method on the heat tolerance of fishes. *Journal of Thermal Biology*, *31*(4), 337–341.

https://doi.org/10.1016/j.jtherbio.2006.01.005

Morgan, R., Finnøen, M. H., Jensen, H., Pélabon, C., & Jutfelt, F. (2020). Low potential for evolutionary rescue from climate change in a tropical fish. *Proceedings of the National Academy of Sciences*, *117*(52), 33365–33372. https://doi.org/10.1073/pnas.2011419117

Morgan, R., Finnøen, M. H., & Jutfelt, F. (2018). CTmax is repeatable and doesn't reduce growth in zebrafish. *Scientific Reports*, *8*(1), 7099. https://doi.org/10.1038/s41598-018-25593-4

Morgan, R., Sundin, J., Finnøen, M. H., Dresler, G., Vendrell, M. M., Dey, A., Sarkar, K., & Jutfelt, F. (2019). Are model organisms representative for climate change research? Testing thermal tolerance in wild and laboratory zebrafish populations. *Conservation Physiology*, *7*(1), coz036. https://doi.org/10.1093/conphys/coz036

Mottola, G., López, M. E., Vasemägi, A., Nikinmaa, M., & Anttila, K. (2022). Are you ready for the heat? Phenotypic plasticity versus adaptation of heat tolerance in three-spined stickleback. *Ecosphere*, *13*(4), e4015. https://doi.org/10.1002/ecs2.4015

Moyano, M., Candebat, C., Ruhbaum, Y., Álvarez-Fernández, S., Claireaux, G., Zambonino-Infante, J.-L., & Peck, M. A. (2017). Effects of warming rate, acclimation temperature and ontogeny on the critical thermal maximum of temperate marine fish larvae. *PLOS ONE*, *12*(7), e0179928. https://doi.org/10.1371/journal.pone.0179928

Nakamura, I., Matsumoto, R., & Sato, K. (2020). Body temperature stability observed in the whale sharks, the world's largest fish. *Journal of Experimental Biology*, jeb.210286. https://doi.org/10.1242/jeb.210286

O'Donnell, M. J., Regish, A. M., McCormick, S. D., & Letcher, B. H. (2024). Thermal transfer rate is slower in bigger fish: How does body size affect response time of small, implantable temperature recording tags? *Ecology of Freshwater Fish*, e12794. https://doi.org/10.1111/eff.12794

Ørsted, M., Jørgensen, L. B., & Overgaard, J. (2022). Finding the right thermal limit: A framework to reconcile ecological, physiological and methodological aspects of CT_{max} in ectotherms. *Journal of Experimental Biology*, *225*(19), jeb244514. https://doi.org/10.1242/jeb.244514

Patra, R. W., Chapman, J. C., Lim, R. P., & Gehrke, P. C. (2007). The effects of three organic chemicals on the upper thermal tolerances of four freshwater fishes. *Environmental Toxicology and Chemistry*, *26*(7), 1454–1459. https://doi.org/10.1897/06-156R1.1

Penman, R. J., Bugg, W., Rost-Komiya, B., Earhart, M. L., & Brauner, C. J. (2023). Slow heating rates increase thermal tolerance and alter mRNA HSP expression in juvenile white sturgeon (*Acipenser transmontanus*). *Journal of Thermal Biology*, *115*, 103599. https://doi.org/10.1016/j.jtherbio.2023.103599

Percie du Sert, N., Ahluwalia, A., Alam, S., Avey, M. T., Baker, M., Browne, W. J., Clark, A., Cuthill, I. C., Dirnagl, U., Emerson, M., Garner, P., Holgate, S. T., Howells, D. W., Hurst, V., Karp, N. A., Lazic, S. E., Lidster, K., MacCallum, C. J., Macleod, M., … Würbel, H. (2020). Reporting animal research: Explanation and elaboration for the ARRIVE guidelines 2.0. *PLOS Biology*, *18*(7), e3000411. https://doi.org/10.1371/journal.pbio.3000411

Pörtner, H. O., Lucassen, M., & Storch, D. (2005). Metabolic Biochemistry: Its Role in Thermal Tolerance and in the Capacities of Physiological and Ecological Function. In *Fish Physiology* (Vol. 22, pp. 79–154). Elsevier. https://doi.org/10.1016/S1546-5098(04)22003-9

Pörtner, H. O., & Peck, M. A. (2010). Climate change effects on fishes and fisheries: Towards a cause-and-effect understanding. *Journal of Fish Biology*, *77*(8), 1745–1779. https://doi.org/10.1111/j.1095-8649.2010.02783.x

Pottier, P., Burke, S., Drobniak, S. M., & Nakagawa, S. (2022). Methodological inconsistencies define thermal bottlenecks in fish life cycle: A comment on Dahlke et al. 2020.

Evolutionary Ecology, *36*(2), 287–292. https://doi.org/10.1007/s10682-022-10157-w

Recsetar, M. S., Zeigler, M. P., Ward, D. L., Bonar, S. A., & Caldwell, C. A. (2012). Relationship between fish size and upper thermal tolerance. *Transactions of the American Fisheries Society*, *141*(6), 1433–1438. https://doi.org/10.1080/00028487.2012.694830

Reemeyer, J. E., & Chapman, L. J. (2024). Effects of acute hypoxia exposure and acclimation on the thermal tolerance of an imperiled Canadian minnow. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology*, *341*(8), 937–949. https://doi.org/10.1002/jez.2847

Reemeyer, J. E., Rumball, D., Mandrak, N. E., & Chapman, L. J. (2024). Seasonal variation in thermal tolerance and hypoxia tolerance of a threatened minnow and a non-imperilled congener: A cautionary tale for surrogate species in conservation. *Conservation Physiology*, *12*(1), coae071. https://doi.org/10.1093/conphys/coae071

Reeve, C., Rowsey, L. E., & Speers-Roesch, B. (2022). Inactivity and the passive slowing effect of cold on resting metabolism as the primary drivers of energy savings in overwintering fishes. *Journal of Experimental Biology*, *225*(8), jeb243407. https://doi.org/10.1242/jeb.243407

Rodgers, E. M., & Gomez Isaza, D. F. (2022). Stress history affects heat tolerance in an aquatic ectotherm (Chinook salmon, *Oncorhynchus tshawytscha*). *Journal of Thermal Biology*, *106*, 103252. https://doi.org/10.1016/j.jtherbio.2022.103252

Ruthsatz, K., Dahlke, F., Alter, K., Wohlrab, S., Eterovick, P. C., Lyra, M. L., Gippner, S., Cooke, S. J., & Peck, M. A. (2024). Acclimation capacity to global warming of amphibians and freshwater fishes: Drivers, patterns, and data limitations. *Global Change Biology*, *30*(5), e17318. https://doi.org/10.1111/gcb.17318

Rutledge, C. J., & Beitinger, T. L. (1989). The effects of dissolved oxygen and aquatic surface respiration on the critical thermal maxima of three intermittent-stream fishes.

Environmental Biology of Fishes, *24*(2), 137–143. https://doi.org/10.1007/BF00001283

Sandblom, E., Clark, T. D., Gräns, A., Ekström, A., Brijs, J., Sundström, L. F., Odelström, A., Adill, A., Aho, T., & Jutfelt, F. (2016). Physiological constraints to climate warming in fish follow principles of plastic floors and concrete ceilings. *Nature Communications*, *7*(1), Article 1. https://doi.org/10.1038/ncomms11447

Smit, K., & Glassom, D. (2017). Large fluctuations but constant mean temperatures allow corals to persist in intertidal rock pools on the east coast of South Africa. *Helgoland Marine Research*, *71*(1), 3. https://doi.org/10.1186/s10152-017-0482-2

Stewart, E., Bowman, J. C., Wilson, C. C., & Raby, G. D. (2024). Local conditions drive interpopulation variation in field-based critical thermal maximum of brook trout. *Conservation Physiology*, *12*(1), coae086. https://doi.org/10.1093/conphys/coae086

Stewart, E. M. C., Frasca, V. R., Wilson, C. C., & Raby, G. D. (2023). Short-term acclimation dynamics in a coldwater fish. *Journal of Thermal Biology*, *112*, 103482. https://doi.org/10.1016/j.jtherbio.2023.103482

Sunday, J. M., Bates, A. E., & Dulvy, N. K. (2012). Thermal tolerance and the global redistribution of animals. *Nature Climate Change*, *2*(9), 686–690. https://doi.org/10.1038/nclimate1539

Tassone, S. J., Besterman, A. F., Buelo, C. D., Ha, D. T., Walter, J. A., & Pace, M. L. (2023). Increasing heatwave frequency in streams and rivers of the United States. *Limnology and Oceanography Letters*, *8*(2), 295–304. https://doi.org/10.1002/lol2.10284

Terblanche, J. S., Hoffmann, A. A., Mitchell, K. A., Rako, L., Le Roux, P. C., & Chown, S. L. (2011). Ecologically relevant measures of tolerance to potentially lethal temperatures. *Journal of Experimental Biology*, *214*(22), 3713–3725. https://doi.org/10.1242/jeb.061283

Tuyttens, F. A. M., De Graaf, S., Heerkens, J. L. T., Jacobs, L., Nalon, E., Ott, S., Stadig, L., Van Laer, E., & Ampe, B. (2014). Observer bias in animal behaviour research: Can we believe

what we score, if we score what we believe? *Animal Behaviour*, *90*, 273–280. https://doi.org/10.1016/j.anbehav.2014.02.007

Verberk, W. C. E. P., Overgaard, J., Ern, R., Bayley, M., Wang, T., Boardman, L., & Terblanche, J. S. (2016). Does oxygen limit thermal tolerance in arthropods? A critical review of current evidence. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *192*, 64–78. https://doi.org/10.1016/j.cbpa.2015.10.020

Verhille, C. E., Lee, S., Todgham, A. E., Cocherell, D. E., Hung, S. S. O., & Fangue, N. A. (2016). Effects of nutritional deprivation on juvenile green sturgeon growth and thermal tolerance. *Environmental Biology of Fishes*, *99*(1), 145–159. https://doi.org/10.1007/s10641-015-0463-8

Vernon, H. M. (1899). The death temperature of certain marine-organisms. *The Journal of Physiology*, *25*(2), 131. https://doi.org/10.1113/jphysiol.1899.sp000782

Vinagre, C., Leal, I., Mendonça, V., & Flores, A. A. V. (2015). Effect of warming rate on the critical thermal maxima of crabs, shrimp and fish. *Journal of Thermal Biology*, *47*, 19–25. https://doi.org/10.1016/j.jtherbio.2014.10.012

Webb, T. J., Lines, A., & Howarth, L. M. (2020). Occupancy-derived thermal affinities reflect known physiological thermal limits of marine species. *Ecology and Evolution*, *10*(14), 7050–7061. https://doi.org/10.1002/ece3.6407

Wells, Z. R. R., McDonnell, L. H., Chapman, L. J., & Fraser, D. J. (2016). Limited variability in upper thermal tolerance among pure and hybrid populations of a cold-water fish. *Conservation Physiology*, *4*(1), cow063. https://doi.org/10.1093/conphys/cow063

Woodin, S. A., Hilbish, T. J., Helmuth, B., Jones, S. J., & Wethey, D. S. (2013). Climate change, species distribution models, and physiological performance metrics: Predicting when biogeographic models are likely to fail. *Ecology and Evolution*, *3*(10), 3334–3346. https://doi.org/10.1002/ece3.680

Woolway, R. I., Anderson, E. J., & Albergel, C. (2021). Rapidly expanding lake heatwaves under climate change. *Environmental Research Letters*, *16*(9), 094013. https://doi.org/10.1088/1748-9326/ac1a3a

Yusishen, M. E., Yoon, G. R., Bugg, W., Jeffries, K. M., Currie, S., & Anderson, W. G. (2020). Love thy neighbor: Social buffering following exposure to an acute thermal stressor in a gregarious fish, the lake sturgeon (*Acipenser fulvescens*). *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, *243*, 110686. https://doi.org/10.1016/j.cbpa.2020.110686

Supplementary Information for

Measuring critical thermal limits in aquatic ectotherms: a practical guide

Fig. S1. Example of generator-powered field-based CT_{max} experiment (Stewart et al. 2024; CTmax arena shown in Fig. 3c). Note that a tent was used (at all field sites) to ensure the CT_{max} arena was shaded from direct sunlight to help control warming rates. In both cases, individual recovery containers (buckets) were prepared in advance, with water set to an intermediate temperature (above ambient but well below CT_{max}) for recovery. Once animals were given *ca*. 10-20 minutes to recover (and confirm survival), they could be individually weighed and measured to link mass and length data to CT_{max} at the individual animal level (the fish were then released back to the wild).

Table S1. Overall warming rates and R^2 achieved in CT_{max} trials in Fig. 6. The achieved rate (slope, *b*) and R^2 were calculated from the linear regression of each trial. Classifications were based on the consistency of the warming rate throughout the trial (in Fig. 6: blue = optimal, orange = non-optimal). The warming rates correspond to the same panel(s) in Fig. 6, where 'optimal' = the blue lines, 'non-optimal' = the orange lines.

Warming rate	**Classification**	**Achieved rate (*b*) (°C h^{-1})**	***R*²**
0.3°C min^{-1} (18°C h^{-1})	Non-optimal	18.339	0.994
0.3°C min^{-1} (18°C h^{-1})	Optimal	17.760	0.999
1°C h^{-1}	Non-optimal	0.993	0.999
1°C h^{-1}	Optimal	0.919	1.000
2°C h^{-1}	Non-optimal	1.540	0.975
2°C h^{-1}	Optimal	1.808	0.999
4°C h^{-1}	Non-optimal	3.874	0.998
4°C h^{-1}	Optimal	4.083	1.000

Table S2. Warming rates and R^2 achieved in the initial and final stages of the CT_{max} trials in Fig. 6. Final vs. initial sections were classified on a trial-by-trial basis by visual inspection of curves in Fig. 6 (Section). The achieved rate (slope, *b*) and R^2 were calculated from a linear regression of each section of each trial. Classifications were based on the consistency of the warming rate throughout the trial (in Fig. 6: blue = optimal, orange = non-optimal). The warming rates correspond to the same panel(s) in Fig. 6.

Targeted warming rate	Classification	Section	Achieved rate (*b*) (°C h^{-1})	R^2
2°C h^{-1}	Non-optimal	Initial (< 4 hours)	1.920	0.996
2°C h^{-1}	Non-optimal	Final (> 4 hours)	0.914	0.994
2°C h^{-1}	Optimal	Initial (< 3 hours)	1.846	0.998
2°C h^{-1}	Optimal	Final (> 3 hours)	1.669	0.998
0.3°C min^{-1} (18°C h^{-1})	Non-optimal	Initial (< 0.4 hours)	19.801	0.998
0.3°C min^{-1} (18°C h^{-1})	Non-optimal	Final (> 0.4 hours)	13.956	0.996
0.3°C min^{-1} (18°C h^{-1})	Optimal	Initial (< 0.5 hours)	18.442	1.000
0.3°C min^{-1} (18°C h^{-1})	Optimal	Final (> 0.5 hours)	16.110	0.999

Inter-trial variability analysis

Methods

The multi-species dataset consists of 12 species. We only used data from each species' control treatment, which was measured in four replicate trials. The number of individuals in a given trial replicate ranged from 5–10 (mean = 8.69 individuals per trial; overall N = 417). The zebrafish dataset consisted of eight treatments. The treatments represent the F2 generation of a selection experiment selecting on CT_{max} where H = selection for high CT_{max}, L = selection for low CT_{max}, R = a control, and AH = selection for high CT_{max} after acclimation to a higher temperature, with two replicates of each (1 and 2) (more details on treatment, holding, and trial conditions in Morgan et al., 2020). Zebrafish CT_{max} was measured at least 12 times in each treatment, so data was filtered to include only the first twelve trials per treatment. The number of individuals per trial replicate ranged from 9–18 (mean = 12.42 individuals per trial; overall N = 1192).

All statistical analysis was conducted using R for Mac OS X (R Core Team 2020; version 4.2.2). We estimated trial effects on CT_{max} using Bayesian linear regression (*brms* package; Bürkner 2017), which uses prior knowledge to build regression models, both from the observed data and our understanding of its distribution. The model is built by drawing thousands of random samples from the posterior predictive distribution to generate model parameters and indicate where parameters differ from the intercept. Individual trials were modelled as covariates of the initial trial to facilitate comparison of each trial's mean CT_{max} with the initial trial's mean CT_{max}. By doing so, we detect where it is highly likely that a subsequent trial's mean would differ from the initial trial, indicating a trial effect.

A normal prior distribution was used for each set of models, with a mean of 0 and a standard deviation of 1 for the multi-species dataset, and 0.5 for the zebrafish dataset, and both datasets were parameterized by population-level effects (i.e., class = b). This choice reflects moderate prior knowledge on the range of variation expected within a given species or treatment in the respective datasets. Each model ran using four Markov chains, each with 5000 burn-in iterations followed by 15,000 iterations. The 95% credible interval (CI) indicates a 95% probability that the true estimate would fall within the interval, given the evidence in the observed data, so trial effects were where the 95% CI did not include 0. R-hat and effective sample sizes (ESS; bulk and tail) were used as diagnostic criteria for model convergence and efficiency. To meet criteria, R-hat had to be equal to 1, and bulk and tail ESS had to each be greater than or equal to 100 * the number of Markov Chains (4). All models met criteria for convergence and efficiency (Table S3, S4).

Results

Of the 12 species in the multi-species dataset, there were seven species which had no trial effects detected, meaning there were no differences detected between trial replicates and the initial trial's mean CT_{max}. The remaining five species demonstrated evidence of trial effects, where at least one of the three trial replicates' mean CT_{max} differed from the initial trial. A trial effect was detected in one of the three subsequent trials in brown shrimp, clownfish, and humbug damselfish (Table S3). In lesser pipefish, two of the three replicate trials' means differed from the initial trial, and in bluntnose minnow, all three replicate trials' means differed from the initial trial (Table S3).

In the zebrafish dataset, seven of the eight treatments showed evidence of trial effects, where at least one of the 11 subsequent trial replicates' mean CT_{max} differed from the initial

trial. There were, on average, three trials per treatment where a trial effect was detected, but the number of replicate trials that differed ranged from 1–6 within a given treatment (mean = 3; Table S4).

Table S3. Multi-species CT_{max} Bayesian regression coefficients (Estimate) and 95% credible intervals (l-95% CI, u-95% CI) of all four trial replicates for each of twelve species. Trials were modelled as covariates to test for trial effects (i.e., differences in mean CT_{max} from the initial trial [Intercept] for each replicate trial). Where the credible interval does not include 0, evidence of a trial effect exists. R-hat, Bulk ESS (estimated sample size), and Tail ESS are diagnostic criteria for model convergence and efficiency where R-hat should = 1, and Bulk and Tail ESS should be ≥ 100 per Markov Chain.

Species	Parameter	Estimate	Est.Error	l-95% CI	u-95% CI	R-hat	Bulk ESS	Tail ESS
bluegill	Intercept	33.1661	0.4215	32.3358	33.9945	1.0001	60626.4335	42563.5107
bluegill	Trial 2	-0.3487	0.5812	-1.4853	0.7966	1.0001	64689.0364	46966.2571
bluegill	Trial 3	0.3625	0.5768	-0.7782	1.4896	1.0000	63961.1369	46287.1184
bluegill	Trial 4	0.3550	0.6125	-0.8566	1.5456	1.0001	65115.9652	44627.6026
bluntnose minnow	Intercept	32.8407	0.3211	32.2420	33.5023	1.0000	43899.6043	37557.1958
bluntnose minnow	Trial 2	1.0825	0.4412	0.1798	1.9132	1.0000	47398.8682	42256.9970
bluntnose minnow	Trial 3	1.7760	0.4736	0.7978	2.6533	1.0000	47146.6710	43799.0444
bluntnose minnow	Trial 4	2.0111	0.4934	0.9869	2.9310	1.0000	47200.8863	44502.4712
brook trout	Intercept	28.5563	0.0992	28.3606	28.7518	1.0000	44780.2092	39356.4640
brook trout	Trial 2	-0.1447	0.1402	-0.4203	0.1341	1.0001	49663.4795	43699.8190
brook trout	Trial 3	-0.0307	0.1474	-0.3212	0.2592	1.0000	50017.5921	45137.6072
brook trout	Trial 4	-0.1788	0.1479	-0.4723	0.1132	1.0000	49915.1019	45050.6376
brown shrimp	Intercept	30.4667	0.4147	29.6644	31.2924	1.0001	41505.6529	39078.8683

brown shrimp	Trial 2	-0.9494	0.5108	-1.9564	0.0491	1.0001	44990.1049	42864.3987
brown shrimp	Trial 3	0.4293	0.5145	-0.6013	1.4231	1.0001	46042.1611	44225.4521
brown shrimp	Trial 4	1.7886	0.5234	0.7339	2.7974	1.0001	44838.8835	44059.6151
clownfish	Intercept	38.0280	0.1761	37.6846	38.3782	1.0000	41120.4738	37903.7351
clownfish	Trial 2	0.1040	0.2414	-0.3782	0.5778	1.0001	45423.0126	43004.3129
clownfish	Trial 3	0.6554	0.2403	0.1784	1.1237	1.0001	45723.7287	44206.1009
clownfish	Trial 4	0.2251	0.2418	-0.2533	0.6995	1.0000	45690.5199	44324.2971
European flounder	Intercept	30.2104	0.1575	29.9021	30.5215	1.0000	42467.8552	40240.5934
European flounder	Trial 2	0.0774	0.2158	-0.3471	0.5047	1.0000	47705.5694	44577.1023
European flounder	Trial 3	-0.0224	0.2271	-0.4710	0.4232	1.0000	48186.7880	43737.6473
European flounder	Trial 4	0.3584	0.2209	-0.0786	0.7895	1.0001	46597.9947	43685.0464
humbug damselfish	Intercept	39.5010	0.1130	39.2784	39.7230	1.0000	40211.0911	38723.9657
humbug damselfish	Trial 2	-0.2352	0.1598	-0.5476	0.0820	1.0000	45650.9664	44065.3116
humbug damselfish	Trial 3	-0.3766	0.1517	-0.6750	-0.0755	1.0000	44498.3038	44467.3270
humbug damselfish	Trial 4	-0.2674	0.1516	-0.5637	0.0340	1.0001	44069.9896	43356.3438
lesser pipefish	Intercept	31.4417	0.1210	31.2049	31.6835	1.0002	42465.7737	39456.1583
lesser pipefish	Trial 2	0.5273	0.1707	0.1864	0.8597	1.0000	48028.4467	45614.6510
lesser pipefish	Trial 3	0.2218	0.1711	-0.1195	0.5572	1.0000	47390.1964	46058.6636
lesser pipefish	Trial 4	0.4620	0.1712	0.1212	0.7983	1.0001	47575.7777	45036.0627
rusty crayfish	Intercept	36.2290	0.3146	35.6170	36.8545	1.0000	53635.6325	42831.7327
rusty crayfish	Trial 2	-0.0423	0.4512	-0.9356	0.8388	1.0000	58540.4141	47182.0458
rusty crayfish	Trial 3	0.4024	0.4532	-0.5020	1.2825	1.0000	56774.5600	44540.4908
rusty crayfish	Trial 4	0.6529	0.4529	-0.2567	1.5253	1.0001	58160.1249	45945.1090

sand goby	Intercept	29.0790	0.2390	28.6122	29.5526	1.0000	48490.9467	42461.8673
sand goby	Trial 2	0.0207	0.3466	-0.6686	0.7011	1.0000	53783.7789	45992.7940
sand goby	Trial 3	0.4356	0.3366	-0.2316	1.0936	1.0000	55392.8532	45682.0705
sand goby	Trial 4	0.3928	0.3920	-0.3863	1.1580	1.0001	56322.1934	46904.6140
stickleback	Intercept	32.7518	0.1384	32.4785	33.0232	1.0001	48324.4733	39332.9552
stickleback	Trial 2	-0.0878	0.2067	-0.4939	0.3245	1.0001	53698.9324	46112.5676
stickleback	Trial 3	-0.0112	0.1945	-0.3942	0.3737	1.0001	52922.9133	46422.9029
stickleback	Trial 4	0.2551	0.2162	-0.1725	0.6811	1.0000	53240.6943	46207.6986
zebrafish	Intercept	41.8655	0.1415	41.5876	42.1483	1.0001	43855.8163	40478.2023
zebrafish	Trial 2	-0.1102	0.1986	-0.5056	0.2795	1.0001	47285.4425	44775.5639
zebrafish	Trial 3	0.2982	0.2075	-0.1131	0.7060	1.0000	45830.8670	43410.0271
zebrafish	Trial 4	0.2736	0.1999	-0.1243	0.6674	1.0001	48035.7973	44432.7396

Table S4. Zebrafish CT_{max} Bayesian regression coefficients (Estimate) and 95% credible intervals (l-95% CI, u-95% CI) for all 12 trials in each of eight treatments. Trials were modelled as covariates to test for trial effects (i.e., differences in mean CT_{max} from the initial trial [Intercept] for each replicate trial). Where the credible interval does not include 0, evidence of a trial effect exists. R-hat, Bulk ESS (estimated sample size), and Tail ESS are diagnostic criteria for model convergence and efficiency where R-hat should = 1, and Bulk and Tail ESS should be ≥ 100 per Markov Chain.

Treatment	Parameter	Estimate	Est.Error	l-95% CI	u-95% CI	R-hat	Bulk ESS	Tail ESS
AH1	Intercept	42.9023	0.0903	42.7244	43.0792	1.0005	20526.6216	32198.0101
AH1	Trial 2	0.0296	0.1394	-0.2455	0.3020	1.0002	35213.5323	42723.5527
AH1	Trial 3	-0.1924	0.1392	-0.4655	0.0838	1.0003	34892.7728	43079.4443
AH1	Trial 4	0.2038	0.1391	-0.0705	0.4754	1.0002	35717.4394	43166.0346
AH1	Trial 5	-0.0422	0.1388	-0.3150	0.2277	1.0002	35918.4964	44005.6714
AH1	Trial 6	0.0751	0.1427	-0.2051	0.3541	1.0002	35989.7108	45166.1165
AH1	Trial 7	0.1206	0.1457	-0.1678	0.4053	1.0002	37821.9835	43218.3819
AH1	Trial 8	0.2517	0.1389	-0.0207	0.5257	1.0001	35212.5739	41647.0312
AH1	Trial 9	-0.0775	0.1461	-0.3625	0.2074	1.0001	37584.0507	42660.7063
AH1	Trial 10	0.0667	0.1423	-0.2106	0.3456	1.0003	37431.7329	42490.5778
AH1	Trial 11	0.1964	0.1393	-0.0769	0.4703	1.0003	34971.6912	42244.5788
AH1	Trial 12	0.2278	0.1391	-0.0439	0.4992	1.0002	36026.6429	45066.9333
AH2	Intercept	43.0362	0.0609	42.9164	43.1553	1.0006	14025.2873	23304.5477
AH2	Trial 2	0.0539	0.0891	-0.1205	0.2278	1.0002	23377.3115	36575.2423
AH2	Trial 3	-0.0110	0.0867	-0.1816	0.1593	1.0003	22750.4596	36261.5412
AH2	Trial 4	0.1282	0.0864	-0.0424	0.2970	1.0003	22699.7965	35779.1993
AH2	Trial 5	-0.0518	0.0870	-0.2244	0.1177	1.0002	22390.6295	35834.8915
AH2	Trial 6	0.0221	0.0871	-0.1491	0.1934	1.0003	22830.8147	35137.3192

AH2	Trial 7	0.0301	0.0868	-0.1402	0.2007	1.0002	22692.9156	34821.2821
AH2	Trial 8	-0.0355	0.0940	-0.2205	0.1487	1.0002	25071.7305	37386.1302
AH2	Trial 9	0.0463	0.0874	-0.1266	0.2168	1.0004	22857.5017	35987.8412
AH2	Trial 10	-0.3542	0.0851	-0.5208	-0.1874	1.0003	21767.2978	34435.3617
AH2	Trial 11	0.0957	0.0868	-0.0753	0.2655	1.0004	22717.4150	35592.7606
AH2	Trial 12	0.1040	0.0870	-0.0674	0.2754	1.0003	22320.1617	34909.5957
H1	Intercept	41.8898	0.0688	41.7556	42.0245	1.0002	13986.5187	25690.3549
H1	Trial 2	-0.1365	0.1007	-0.3333	0.0625	1.0001	23882.8490	37109.0143
H1	Trial 3	-0.1041	0.1011	-0.3009	0.0937	1.0001	24045.7891	37531.5985
H1	Trial 4	-0.4115	0.0990	-0.6038	-0.2165	1.0002	23010.3255	35852.3206
H1	Trial 5	-0.1406	0.1043	-0.3459	0.0632	1.0001	25255.7085	37091.6292
H1	Trial 6	-0.2307	0.0995	-0.4276	-0.0354	1.0001	23466.0559	36405.2989
H1	Trial 7	-0.2339	0.1014	-0.4330	-0.0334	1.0001	24293.3179	37743.8316
H1	Trial 8	-0.3814	0.0987	-0.5740	-0.1890	1.0001	23594.1175	37200.9548
H1	Trial 9	0.0856	0.0992	-0.1090	0.2814	1.0002	23581.9633	36479.2823
H1	Trial 10	-0.2306	0.0995	-0.4262	-0.0357	1.0001	23852.8406	36748.8472
H1	Trial 11	-0.0952	0.0992	-0.2880	0.0999	1.0001	23212.4002	35067.6192
H1	Trial 12	-0.3137	0.0993	-0.5077	-0.1176	1.0000	22929.7131	36256.2331
H2	Intercept	41.6815	0.0666	41.5510	41.8141	1.0002	20197.1655	31266.7641
H2	Trial 2	-0.1200	0.1042	-0.3252	0.0840	1.0000	34371.6991	40575.4514
H2	Trial 3	-0.2368	0.1023	-0.4373	-0.0348	1.0001	33877.4637	41409.4825
H2	Trial 4	-0.1766	0.1070	-0.3865	0.0339	1.0002	35773.9333	42700.5112

H2	Trial 5	0.1477	0.1048	-0.0603	0.3532	1.0001	34941.3326	41726.8981
H2	Trial 6	-0.3614	0.1066	-0.5688	-0.1510	1.0000	35879.7624	42221.5554
H2	Trial 7	0.0507	0.1045	-0.1538	0.2557	1.0000	34721.1688	42365.7775
H2	Trial 8	-0.2557	0.1076	-0.4668	-0.0444	1.0000	35103.0681	40492.5512
H2	Trial 9	0.1857	0.1072	-0.0239	0.3972	1.0001	35976.0328	41175.6207
H2	Trial 10	-0.2820	0.1041	-0.4843	-0.0787	1.0001	34655.6879	43509.2752
H2	Trial 11	-0.0527	0.1069	-0.2631	0.1573	1.0001	35538.7292	41983.7691
H2	Trial 12	-0.3062	0.1045	-0.5125	-0.1005	1.0000	34434.4594	39678.3505
L1	Intercept	41.4607	0.0795	41.3053	41.6159	1.0002	15105.3924	26139.4244
L1	Trial 2	-0.2767	0.1172	-0.5044	-0.0462	1.0000	26365.1641	38730.2112
L1	Trial 3	0.1522	0.1202	-0.0833	0.3877	1.0000	26305.0869	38698.8003
L1	Trial 4	-0.2531	0.1154	-0.4783	-0.0267	1.0001	25710.2903	37833.7419
L1	Trial 5	0.0281	0.1235	-0.2148	0.2705	1.0000	27987.3836	39113.6659
L1	Trial 6	-0.2159	0.1146	-0.4403	0.0097	1.0001	25042.6669	37360.7017
L1	Trial 7	0.0380	0.1104	-0.1783	0.2548	1.0000	23919.6223	37008.3450
L1	Trial 8	-0.1315	0.1170	-0.3611	0.0993	1.0000	25706.4851	37691.5744
L1	Trial 9	-0.1391	0.1178	-0.3701	0.0931	1.0000	26196.5018	38271.1359
L1	Trial 10	-0.1484	0.1151	-0.3740	0.0800	1.0001	25139.6053	37199.9914
L1	Trial 11	0.0803	0.1126	-0.1421	0.2997	1.0000	24524.6261	38151.5914
L1	Trial 12	0.0308	0.1148	-0.1952	0.2551	1.0001	25103.0755	37093.4416
L2	Intercept	41.3223	0.0832	41.1603	41.4857	1.0003	16881.9546	29109.1249
L2	Trial 2	-0.1105	0.1254	-0.3600	0.1337	1.0001	29169.2356	38886.4485

L2	Trial 3	0.0548	0.1228	-0.1863	0.2951	1.0001	28015.0940	39590.6186
L2	Trial 4	-0.1102	0.1249	-0.3559	0.1341	1.0001	28443.4657	37816.7828
L2	Trial 5	0.0825	0.1254	-0.1623	0.3274	1.0001	28696.8594	40534.6876
L2	Trial 6	-0.1180	0.1245	-0.3619	0.1270	1.0001	29603.9325	41109.0041
L2	Trial 7	0.3489	0.1254	0.1019	0.5947	1.0001	29347.5442	40424.5846
L2	Trial 8	-0.0040	0.1311	-0.2622	0.2511	1.0001	31192.1053	40849.8455
L2	Trial 9	0.1567	0.1253	-0.0894	0.4001	1.0001	28584.3999	40207.5960
L2	Trial 10	0.0474	0.1231	-0.1948	0.2887	1.0000	29090.7030	41690.4945
L2	Trial 11	0.2905	0.1288	0.0348	0.5411	1.0002	30443.3584	39947.6945
L2	Trial 12	0.1488	0.1253	-0.0973	0.3948	1.0001	28967.5629	39324.1410
R1	Intercept	41.5221	0.0707	41.3840	41.6615	1.0002	14687.7946	27334.6604
R1	Trial 2	-0.2614	0.1046	-0.4652	-0.0562	1.0001	24587.9841	36964.7409
R1	Trial 3	0.0237	0.1039	-0.1791	0.2279	1.0001	24304.8816	38519.8900
R1	Trial 4	-0.4046	0.1040	-0.6091	-0.2016	1.0001	24122.6452	38573.5013
R1	Trial 5	0.0234	0.1041	-0.1824	0.2279	1.0001	24582.3064	38439.7514
R1	Trial 6	0.0309	0.1046	-0.1738	0.2358	1.0001	24946.0670	37626.2186
R1	Trial 7	-0.0354	0.1021	-0.2345	0.1656	1.0002	23793.9407	37476.7390
R1	Trial 8	-0.0440	0.1045	-0.2482	0.1612	1.0001	24276.9424	38155.4892
R1	Trial 9	0.1600	0.1022	-0.0400	0.3605	1.0000	23852.7274	38030.0135
R1	Trial 10	0.0537	0.1042	-0.1516	0.2572	1.0001	24142.1636	38147.4082
R1	Trial 11	0.0608	0.1039	-0.1426	0.2637	1.0002	25102.8438	37718.4950
R1	Trial 12	-0.2541	0.1044	-0.4585	-0.0484	1.0001	24685.6069	38695.1635

R2	Intercept	41.5030	0.0857	41.3336	41.6710	1.0002	17258.5516	27479.7174
R2	Trial 2	-0.1887	0.1279	-0.4400	0.0627	1.0001	29610.1562	41091.7025
R2	Trial 3	-0.1433	0.1300	-0.3986	0.1129	1.0001	30483.2141	39851.7888
R2	Trial 4	-0.2389	0.1302	-0.4941	0.0185	1.0001	30291.3656	40794.4971
R2	Trial 5	0.0630	0.1299	-0.1917	0.3171	1.0001	29738.4345	39574.7877
R2	Trial 6	-0.3275	0.1306	-0.5846	-0.0711	1.0001	30577.2449	40210.3498
R2	Trial 7	-0.0303	0.1198	-0.2655	0.2050	1.0001	26909.7462	39218.3057
R2	Trial 8	-0.1066	0.1336	-0.3685	0.1561	1.0001	31040.8455	40099.7578
R2	Trial 9	-0.0546	0.1302	-0.3081	0.2009	1.0002	30306.8543	41325.6925
R2	Trial 10	-0.1224	0.1324	-0.3803	0.1366	1.0001	31181.2529	40387.7562
R2	Trial 11	-0.3318	0.1355	-0.5975	-0.0638	1.0001	32274.8316	40754.8878
R2	Trial 12	-0.0853	0.1274	-0.3327	0.1643	1.0002	29538.5115	40348.0246