**Methodological inconsistencies define thermal bottlenecks in fish life cycle: a comment on Dahlke *et al.* 2020**

Patrice Pottier1\*, Samantha Burke1, Szymon M. Drobniak1,2, and Shinichi Nakagawa1

1 Ecology & Evolution Research Centre, School of Biological, Earth and Environmental Sciences, The University of New South Wales, Sydney, New South Wales, Australia

2 Institute of Environmental Sciences, Jagiellonian University, Kraków, Poland

\*Corresponding author
E-mail address: p.pottier@unsw.edu.au (PP)

**Keywords:** climate vulnerability, thermal tolerance, comparability, comparative physiology, climate change

**Total word count:** 1496 excluding title page, references, figures, and declarations.

**Number of figures:** 1

**Number of tables:** 0

**Acknowledgements**: We thank Matthew Symonds for encouraging the submission of this manuscript.

**This preprint has not undergone peer review or any post-submission improvements or corrections. The Version of Record of this article is published in Evolutionary Ecology, and is available online at https://doi.org/10.1007/s10682-022-10157-w**

**Abstract**

Comparative analyses require researchers to not only ensure data quality, but also to make prudent and justifiable assumptions about data comparability. A failure to do so can lead to unreliable conclusions. As a case in point, we comment on a study that estimated the vulnerability of the world’s fish species to climate change using comparison between life stages (Dahlke et al. 2020, Science 369: 65-70). We highlight concerns with the data quality and argue that the metrics used to investigate ontogenetic differences in thermal tolerance were incomparable and confounded. Therefore, we provide caution when interpreting their results in the light of climate vulnerability. We suggest potential remedies and recommend thermal tolerance metrics that may be comparable across life stages. We also encourage the creation of guidelines to design, report, and assess comparative analyses to increase their reliability and reproducibility.

**Introduction**

Comparative analyses are essential for highlighting common and distinct patterns in the ecology and physiology of organisms across the tree of life. However, this is not a trivial exercise. The diversity of research questions, methods, and species investigated in ecological studies creates significant sources of data heterogeneity (Senior et al. 2016). Therefore, it is the researcher’s responsibility to ensure that the assumptions made regarding the quality and comparability of the compiled data are reasonable and justifiable.

Here, we use a case study on the vulnerability of the world’s fishes to temperatures (Dahlke et al. 2020) to illustrate how questionable assumptions may impact the reliability of research findings. To predict fish vulnerability to climate change, Dahlke and colleagues compiled published thermal tolerance limits of nearly 700 fish species at different life stages. Then, the authors inferred life stage-specific differences in thermal tolerance limits and assessed the vulnerability of fish species to climate change. The authors found that the mean thermal tolerance range of embryos and spawners consistently differ from larvae and non-spawning adults by over 20°C. While the implications of this study are topical, we contend that Dahlke et al. (2020) made questionable assumptions that compromise the robustness of their research findings. We provide a list of concerns, suggest potential remedies, and offer recommendations for future comparative studies in evolutionary ecology.

***Data comparability and confounding factors***

Thermal tolerance limits commonly define the minimum and maximum temperatures an organism can tolerate over a given period. Conventionally, two main metrics are used to quantify these traits: critical thermal minima/maxima (CTM; Fig. 1a,b; Cowles and Bogert 1944), and median lethal temperatures (LT50; Fig. 1c,d; Fry 1947). While CTM are easy to quantify in juveniles and adults by recording motor function, monitoring behavioural responses of embryos is challenging, and published evidence in spawners is sparse. Therefore, Dahlke et al. (2020) extended their study to include additional metrics.

Dahlke et al.’s data for spawners reflected “*the temperature at which ripe males and females* *(advanced stage of gonad maturation) have been observed in the field*”, which represent field occurrence data. While thermal tolerance limits of fish species may correlate with the temperature of their spawning sites, habitat use cannot be used as a comparable metric with CTM or LT50. In essence, field occurrence data underestimate thermal tolerance limits, and comparing these metrics is analogous to comparing fundamental and realized ecological niches. Their embryo thermal limits, on the other hand, were the “*High or low temperatures reported to cause a statistically significant change in survival relative to the optimum value”* (Fig. 1e,f)*.* Such a metric largely differs from the extreme temperatures an embryo can tolerate, which correspond to the far-ends of the thermal performance curve (Fig. 1e).

While these uncommon metrics provide *proxies* for thermal tolerance, they are unlikely to be directly comparable to CTM and LT50 without statistical validation. Dahlke et al. (2020) validated spawners *in situ* observations with laboratory estimates of temperatures “*that exert negative effects on gonad maturation and/or subsequent offspring survival*” (Fig 1e). While it is sensible to validate the compiled metrics, this method compares *in situ* observations with deviations from optimal reproductive performance which, as previously discussed for embryos, is, again, unlikely to be directly comparable with CTM and LT50.

Although comparative analyses can use different metrics, principles of experimental biology require the compared data not being confounded with other factors. Strikingly, the alternative metrics of Dahlke et al. (2020) were used solely for embryos and spawners, whereas CTM and LT50 data were used only in juveniles and adults. The separation into stage-specific metrics essentially confounds life stage and thermal tolerance metrics. Because *all things are not equal*, distinguishing whether differences in thermal tolerance limits are the result of varying metrics or varying life stages, is impossible. Therefore, we believe that the large stage-specific differences in thermal limits were, in fact, defined by underestimated metrics in embryos and spawners.

***Additional sources of concerns***

The piece from Dahlke et al. (2020) is subject to additional sources of procedural concerns. Thermal tolerance limits are plastic and highly dependent on prior thermal history, with higher temperatures and longer acclimations commonly inflating thermal tolerance. The authors made the effort to account for differences in acclimation “*by considering the value determined under the highest/longest acclimation treatment*”. However, experimental biologists often use durations and temperatures of acclimation outside the range naturally experienced by animals. Because this correction was used only in juveniles and adults, it likely inflated the estimates linked to these life stages relative to embryos and spawners.

In comparative physiology, it is also paramount to ensure data quality prior to its inclusion in analyses. Unfortunately, we identified data inconsistencies during a non-systematic and brief examination of the original data. For example, despite being used for collecting embryo data for 96 species, studies published by Mito (1961a,b; 1962a,b,c; 1963a,b,c,d) do not report thermal tolerance data. Indeed, while Dahlke et al. (2020) reported the thermal tolerance of *Oplegnathus fasciatus* as 23.4°C, the original reference (Mito 1963a) denotes that “*At water temperatures of 20.6-23.4°C, the embryos reach the seventh stromal stage at 16.5 hours post-fertilization*” (translated from Japanese), which does not indicate that thermal limits, nor deviations from optimum survival, were quantified.

***Suggested remedies***

We believe that the concerns we highlighted are mostly caused by dubious assumptions about the metrics’ comparability. Here, we suggest a homogenization of metrics among life stages to remove sources of confounding factors.

For instance, deviations from the optimum survival of juveniles, adults, and spawners may appear as comparable estimates to the ones currently used for embryos. However, the dependency of this metric on statistical significance makes it susceptible to differences in sample sizes between studies. Instead, we recommend using available CTM and LT50 in embryos and spawners as comparable metrics with the published evidence on juveniles and adults.

The CTM of embryos can be estimated by recording the temperature at which the heart stops beating (Del Rio et al. 2019; Zebral et al. 2018). In some instances, such metric provides evidence for higher CTM in embryos relative to juveniles (Del Rio et al. 2019). Using CTM in spawners is also interesting, and some evidence points pre-reproductive and reproductive Trinidadian guppies to differ by less than 1°C when using this metric (Auer et al. 2021). We also recommend employing the temperature resulting in a 50% decrease in embryo survival relative to the optimal temperature as a proxy for LT50. Interestingly, the concept of thermal tolerance landscapes (Rezende et al. 2014) also offers possibilities to unify thermal tolerance estimates to a single comparable metric (Jørgensen et al. 2019, 2021). We believe this framework opens fruitful avenues for reducing heterogeneity between methodologies and life stages.

Comparable estimates may not be available in sufficient amounts for all life stages, however. If true, we argue that the time is perhaps not ripe for a synthesis of this scale. We urge evolutionary ecologists to standardize thermal tolerance metrics across life stages. Until a more standardized system is in place can a synthesis reliably explore ontogenetic differences in thermal tolerance.

***Implications and future directions***

The implications of the study from Dahlke et al. (2020) are outstanding. If thermal bottlenecks in the life cycle were to occur and impact the vulnerability of fishes to climate change, Dahlke et al.’s conclusion could drastically influence research and conservation efforts. However, we are unsure that the data compiled by Dahlke et al. (2020) can confidently address this question. Therefore, we caution against generalizing their results too promptly.

Unfortunately, the concerns highlighted in Dahlke et al. (2020) are not an isolated case. Synthesizing a large body of evidence exposes researchers to a diversity of experimental designs that may impact the quality of the data. Therefore, ensuring that data is robust, representative, and comparable is critical for any comparative analysis.Although providing comprehensive guidelines is not the main aim of this piece, we encourage evolutionary ecologists to create guidelines for designing, reporting, and assessing comparative studies. In the meantime, we refer the reader to *Phylotocol* (DeBiasse and Ryan 2018), *PRISMA-EcoEvo* (O’Dea et al. 2021) and Nakagawa et al. (2017), which provide useful resources, albeit not specifically tailored for comparative studies in evolutionary ecology.

**Declarations:**

Funding: Financial support was provided by a UNSW Scientia Doctoral Scholarship to PP and SB. SN was supported by the Australian Research Council (ARC) Discovery Project (DP200100367). SMD was supported by the ARC Discovery Early Career Award (DE180100202) and the OPUS from Polish National Science Centre (UMO-2020/39/B/NZ8/01274).

Conflicts of interest/Competing interests: The authors declare no conflicts of interest or competing interests.

Ethics approval: Not applicable.

Consent to participate: All authors gave consent to participate in this work.

Consent for publication: All authors gave consent for publication.

Availability of data and material: Not applicable.

Code availability: Not applicable.

Authors’ contributions: PP conceived and coordinated the study. All authors participated in the conceptualization and critically revised the manuscript.

**References:**

Auer SK, Agreda E, Chen AH, et al (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

Cowles RB (Raymond B, Bogert CM (Charles M (1944) A preliminary study of the thermal requirements of desert reptiles. Bulletin of the American Museum of Natural History; v. 83, article 5.

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

DeBiasse MB, Ryan JF (2019) Phylotocol: Promoting Transparency and Overcoming Bias in Phylogenetics. Systematic Biology 68:672–678. https://doi.org/10.1093/sysbio/syy090

Del Rio AM, Davis BE, Fangue NA, Todgham AE (2020) Combined effects of warming and hypoxia on early life stage Chinook salmon physiology and development. Conservation Physiology 7(1). https://doi.org/10.1093/conphys/coy078

Fry FEJ (1947) Effects of the environment on animal activity. The University of Toronto Press, Toronto, 62 pp.

Jørgensen LB, Malte H, Ørsted M, et al (2021) A unifying model to estimate thermal tolerance limits in ectotherms across static, dynamic and fluctuating exposures to thermal stress. Sci Rep 11:12840. https://doi.org/10.1038/s41598-021-92004-6

Jørgensen LB, Malte H, Overgaard J (2019) How to assess *Drosophila* heat tolerance: Unifying static and dynamic tolerance assays to predict heat distribution limits. Functional Ecology 33:629–642. https://doi.org/10.1111/1365-2435.13279

Mito S (1961a) Pelagic fish eggs from Japanese waters-I. Clupeina, Chanina, Stomiatina, Myctophida, Anguillida, Belonida and Syngnathida. Science Bulletin of the Faculty of Agriculture, Kyushu University 18:285–310

Mito S (1961b) Pelagic fish eggs from Japanese waters-II. Lamprida, Zeida, Mugilina, Scombrina, Carangina and Stromateina. Science Bulletin of the Faculty of Agriculture, Kyushu University 18:451–466

Mito S (1962a) Pelagic fish eggs from Japanese waters-V. Calliomina and Ophidiina. Science Bulletin of the Faculty of Agriculture, Kyushu University 19:377–380

Mito S (1962b) Pelagic fish eggs from Japanese waters-VII. Chaetodontina, Balistina and Ostraciontina. Science Bulletin of the Faculty of Agriculture, Kyushu University 19:503–506

Mito S (1962c) Pelagic fish eggs from Japanese waters—VI. Labrina. Science Bulletin of the Faculty of Agriculture, Kyushu University 19:493–502

Mito S (1963a) Pelagic fish eggs from Japanese waters-III Percina. Japanese Journal of Ichthyology 11:39-64. https://doi.org/10.11369/jji1950.11.39

Mito S (1963b) Pelagic fish eggs from Japanese waters-IX. Echenieida and Pleuronectida. Japanese Journal of Ichthyology 11:81-102. https://doi.org/10.11369/jji1950.11.81

Mito S (1963c) Pelagic fish eggs from Japanese waters—IV. Trachinina and Uranoscopina. Science Bulletin of the Faculty of Agriculture, Kyushu University 19:369–376

Mito S (1963d) Pelagic fish eggs from Japanese waters-VIII. Cottina. Japanese Journal of Ichthyology 11:65-79

Nakagawa S, Noble DWA, Senior AM, Lagisz M (2017) Meta-evaluation of meta-analysis: ten appraisal questions for biologists. BMC Biology 15:18. https://doi.org/10.1186/s12915-017-0357-7

O’Dea RE, Lagisz M, Jennions MD, et al (2021). Preferred reporting items for systematic reviews and meta-analyses in ecology and evolutionary biology: a PRISMA extension. Biological Reviews. https://doi.org/10.1111/brv.12721

Rezende EL, Castañeda LE, Santos M (2014) Tolerance landscapes in thermal ecology. Functional Ecology 28:799–809. https://doi.org/10.1111/1365-2435.12268

Senior AM, Grueber CE, Kamiya T, et al (2016) Heterogeneity in ecological and evolutionary meta-analyses: its magnitude and implications. Ecology 97:3293–3299. https://doi.org/10.1002/ecy.1591

Zebral YD, Lansini LR, Costa PG, et al (2018) A glyphosate-based herbicide reduces fertility, embryonic upper thermal tolerance and alters embryonic diapause of the threatened annual fish *Austrolebias nigrofasciatus*. Chemosphere 196:260–269. https://doi.org/10.1016/j.chemosphere.2017.12.196

**Figures:**

Figure 1: Graphic overview of conventional thermal tolerance metrics, and the metrics used by Dahlke et al. (2020). **a.** Representation of the experimental design used to assess critical thermal limits. Here, the temperature is gradually increased (critical thermal maximum, CTmax), or decreased (critical thermal minimum, CTmin) and the temperature at which individual animals reach a pre-defined endpoint (e.g., loss of motor function) is recorded. The CTmax or CTmin is then defined as the arithmetic mean of the individual endpoints. **b.** Alternative representation of CTmin and CTmax. Depicted is a thermal performance curve that is bounded by CTmin and CTmax, the temperatures at which the animals lose their motor function. The temperature for optimal performance is defined as Topt­. **c.** Representation of median lethal temperature assessment. In these experiments, cohorts of animals are traditionally exposed to a range of temperatures and their survival is assessed at the end of the experiment. The survival rate at each temperature is then used to generate a logistic or probit survival curve and the LT50 is interpolated as the temperature lethal for 50% of the animals (**d.**). In **e.** and **f.** are depicted some metrics used by Dahlke et al. (2020). **e.** Metric used for the thermal limits of embryos and the validation of the metric used for spawners. Here, the grey-shaded area represents the hypothetical range of temperatures that are not statistically different from the temperature for optimal performance. With this metric, any temperature outside this zone can be defined as the thermal tolerance limit. Therefore, this metric strongly depends on the temperatures used in the experiment, which may be largely different from CTmax and CTmin. **f.** Difference between the metric used in Dahlke et al. 2020 considering the same experimental design as presented in **c.** Here, the first temperature tested causing a deviation from optimal survival is considered as the thermal tolerance limit, which is different from LT50. In all sections, red and blue stars represent true thermal limits, whereas grey stars denote thermal limits misidentified through fallacious design.