Received: 23 March 2021 | Revised: 18 May 2021 | Accepted: 27 May 2021

DOI: 10.1111/jeb.13885

INVITED COMMENTARY

Journal of Evolutionary Biology eseb WILEY

A comment on *The adaptive value of gluttony: predators mediate the life history trade-offs of satiation threshold* by Pruitt & Krauel (2010)

Erik Postma[1] | **Alejandro Gonzalez-Voyer**[2] | **Luke Holman**[3]

[1]Centre for Ecology and Conservation, University of Exeter, Cornwall, UK

[2]Departamento de Ecología Evolutiva, Instituto de Ecología, Universidad Nacional Autónoma de México, Mexico City, Mexico

[3]School of Applied Sciences, Edinburgh Napier University, Sighthill Campus, Sighthill Court, Edinburgh, UK

Correspondence
Erik Postma, Centre for Ecology and Conservation, University of Exeter, Treliever Road, TR10 9FE, Cornwall, UK.
Email: e.postma@exeter.ac.uk

Abstract

Inspection of the data that accompany Pruitt and Krauel's study of individual variation in satiation threshold and a comparison of these data with the *Materials and Methods* and *Results* sections of the paper have revealed a number of issues that cast doubts on the reliability of the data and any results based on these data. In particular, we show that, following our analyses, the data are unlikely to have been obtained using the study design outlined in the publication and that statistical analyses of these data provide results that differ in important ways from those reported. These findings illustrate the importance of making raw data and analysis code available for the rigour and reproducibility of the scientific literature.

KEYWORDS

simulation, life-history evolution, quantitative genetics, trade-offs

1 | INTRODUCTION

The question of how individual variation in behavioural and other fitness-related traits is maintained is central to evolutionary biology. One popular explanation is that trade-offs between different fitness components allow individuals with different phenotypes to have equal fitness, which ultimately preserves phenotypic diversity in the face of selection (Roff & Fairbairn, 2007; Stearns, 1992). To test this, Pruitt and Krauel (2010b) (PK2010) measured how much an individual wolf spider (*Schizocosa ocreata*) eats during a foraging bout (termed its satiation threshold) and tested whether the consequences of this behaviour for its fecundity and survival were dependent on the intensity of predation. Here, we report a number of issues with the data and analyses in this study that raise serious concerns about its findings.

2 | AVAILABILITY OF RAW DATA

PK2010 presents data collected on wild-caught females from a single site, and their captive-born daughters. In short, a total of 514 immature spiders of unknown sex were collected and subsequently housed individually until maturation. After reaching maturity, wild-caught females were mated and their offspring were again housed individually until maturation. Both mothers and daughters had their satiation threshold measured within 3 days of maturation.

According to the *Materials and Methods*, satiation threshold was assessed 'by offering females size-matched prey items (25% of test female mass ±3%) at 10-min intervals until two consecutive prey items were rejected [...] Individuals were weighed 2 hr prior to their satiation threshold trials and were reweighed after

Accompanying articles for the 'Invited Commentary' type are 10.1111/j.1420-9101.2010.02070.x, 10.1111/jeb.13922, 10.1111/jeb.13915 and 10.1111/jeb.13972

their 24-hr feeding period. Satiation threshold was measured as an individual's percentage increase in mass following a gluttonous feeding trial'. Thus, calculating satiation threshold would have required measurement of each individual's mass before and after the feeding trial. However, the data file deposited on *Dryad* (Pruitt & Krauel, 2010a) only contains the pre-trial mass (measured in grams to 4 decimal places), but not the post-trial mass, and this for the captive-born daughters only. The data file does contain a third column with the change in mass in grams (from which one could infer the post-trial mass), but this column is calculated using an *Excel* formula from the pretrial mass and the (rounded) percentage increase (Figure S1).

According to the lead author (Pruitt, pers. comm.), the post-trial mass data are unavailable because the per cent change was calculated from the pre- and post-trial masses while collecting the post-trial mass data, and post-trial mass was then discarded. Pruitt further said that the back-calculated absolute change in mass was included as 'a service to other researchers' and that the satiation thresholds were rounded to the nearest per cent because that was considered 'the scale of the meaningful variation across individuals'.

In short, this means that most of the raw data for PK2010 are unavailable. Although daughter post-trial mass can in theory be deduced from the pretrial mass and satiation threshold, the latter variable was rounded, such that the post-trial masses can only be determined approximately. These data therefore cannot be assessed, which is especially unfortunate given the irregularities identified in Sections 3 and 4.

3 | PATTERNS IN DECIMALS

To infer the relative importance of additive genetic effects in shaping individual variation in satiation threshold, the heritability of satiation threshold was 'assessed using linear, dam-on-offspring (sic) regression' (see Section 6). Despite satiation thresholds having been rounded to the nearest percentage point prior to further analyses (see Section 2), 71% of the maternal values used in the mother–daughter regression end with a decimal, and 67% of these values end in .5 or .7 (Figure 1a,b). Similarly, as daughter satiation thresholds are the mean of two individuals, we would expect all means to end in either .0 or .5 if the individual data used in this analysis were integer-valued. However, 34% of the offspring means do not end in .0 or .5 (Figure 1c,d).

Strikingly, when *offspring* values are sorted in ascending order ignoring any decimal places—which with a few exceptions is how they are ordered in the data file on *Dryad* (Figures 1c and S2)—the decimal places of the *maternal* satiation thresholds show an inexplicable series of consecutive values that have identical decimals (Figures 1b and S2). There is no methodological or biological explanation for why the ordering of offspring mean sizes in the study as a whole should influence the mass of mothers at a particular decimal place.

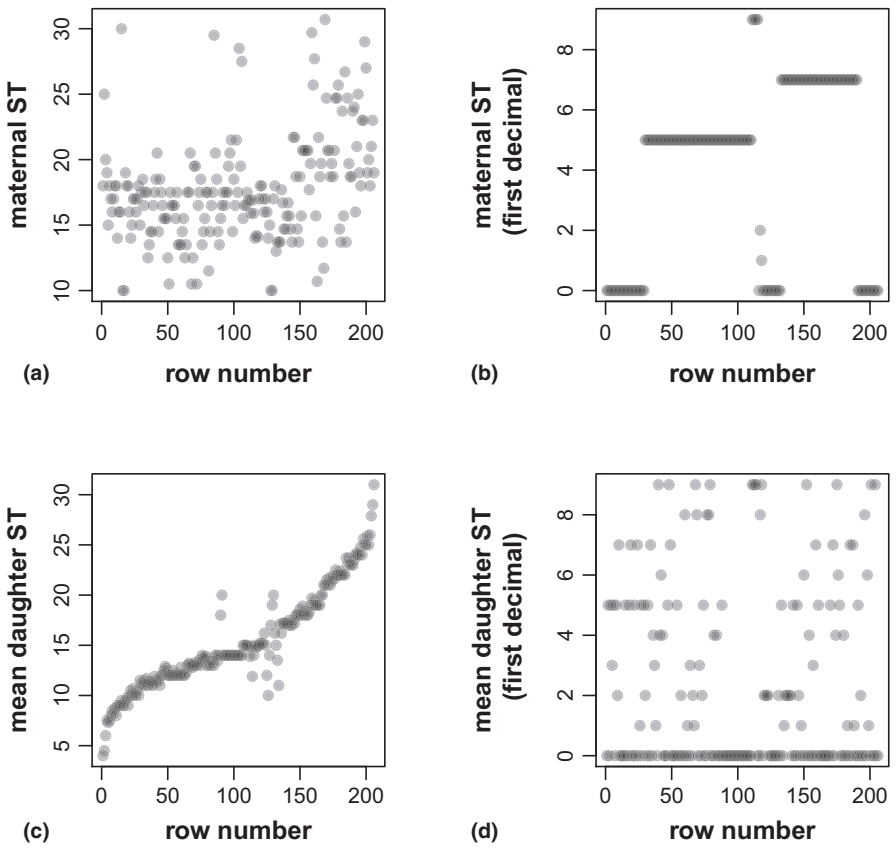

FIGURE 1 Visualization of mother and daughter satiation threshold (ST) data in the order in which they appear in the raw data file (a and c), and of the first decimal for these same ST values (b and d; for example, 5 means that there was a 5 immediately after the decimal point)

4 | DISTRIBUTION OF PRE-TRIAL BODY MASS

While neither post-trial mass nor the increase in mass is available (see Section 2), we found irregularities in the pre-trial body mass of captive-born daughters included in the mark–recapture study (Figure 2). As expected from the fact that females were size-matched before their introduction to one of six pairs of plots, either without (plots 1A–6A) or with predators (plots 1B–6B), there are large and systematic differences in mean mass *among* pairs of plots (Figure 2a). What is unexpected is the extremely low variance in body mass *within* each plot (±2% according to the *Materials and Methods*). For example, 18 of 32 females released in Plot 1A are exactly 0.1089 g. As a consequence, there are large gaps in the distribution of body size (Figure 2b,c), and, for example, not a single female in the study weighed between 0.0961 and 0.1023 g, despite many individuals being larger or smaller than this.

These features of the data are incompatible with the *Materials and Methods*, which state that these spiders represent almost all of the daughters of a collection of wild-caught females from a single population. We would expect the distribution of masses to be continuous (and probably Normal), and to contain few duplicated values when spiders are weighed to a tenth of a milligram. Although in theory a distribution similar to that observed could be generated by extreme subsetting of a larger population, this would require a pool of daughters to draw upon that is many times larger than could realistically be obtained following the protocol outlined.

5 | SAMPLE SIZES

After reaching maturity, wild-caught females were mated. These matings resulted in a total of 1,194 offspring from 217 females (i.e. females contributed 5.5 offspring on average). After their satiation threshold was assessed (see Section 2), two daughters per brood ($n = 412$ daughters) were randomly selected for a mark–recapture study, which implies that $412/2 = 206$ out of 217 broods (i.e. 95%) contained at least two daughters that survived until maturity.

This number seems unlikely, because we should expect more of the broods to contain fewer than 2 females, given that the average brood size is only 5.5. We calculated the probability that 206 (or more) of the 217 broods would contain at least two adult daughters, assuming a 50:50 sex ratio and no sex difference in survival, under the generous assumptions that all 1,194 offspring reached adulthood and that brood sizes were as even as possible. The probability of getting 206 broods with at least two females is extremely low and is lower still if we assume Poisson-distributed variation in brood size (<0.00001; Supplementary Material S4 and Figure S4.1). We thus conclude that it is highly unlikely that these data have been obtained following the study design outlined in the publication.

After mating, these 206 pairs of full-sibs were released into one of twelve equally sized experimental plots. According to the *Materials and Methods*, before their release, all resident *S. ocreata* were removed and the number of released captive-born individuals was equal to the number of residents removed, 'to mimic naturalistic densities' (PK2010). This implies that (1) the 12 plots combined contained 412 resident spiders and that (2) the number of individuals released in each plot mimics natural variation in density. The sample sizes in Table 1 imply that among-plot variation in spider density is strikingly low (mean ± standard deviation = 34.3 ± 2.5), much lower than expected under a simple Poisson model (Supplementary Material S4 and Figure S4.2). Even more surprisingly, the authors found six pairs of natural plots (treatment A vs. B) that contained exactly the same numbers of spiders (Table 1).

6 | STATISTICAL ANALYSES

According to the *Materials and Methods*, Pruitt and Krauel 'tested for differences in mean satiation threshold and starting mass between treatments using Student's *t*-tests. As outlined in the *Results*, they found that 'there were no significant differences in individuals' satiation threshold ($T = 1.43$, $df = 410$, $p = .15$) or starting mass ($T = 0.10$, $df = 410$, $p = .92$) among treatments'.

When we repeat these two *t*-tests (Supplementary Material S5), we indeed find that there is no statistically significant difference among the two treatment groups in starting mass. However, although the residual degrees of freedom is identical to that reported, the *t* value and *p*-value are different ($t = 0.263$, $p = .793$). Furthermore, in contrast to what is reported, there is a highly significant difference in satiation threshold between both treatment groups ($t = 3.883$, $p < .001$).

These discrepancies cannot be explained by differences among software packages (R vs. SAS) as suggested by Pruitt (pers. comm.). Instead, they suggest that the data used in these (and potentially other) statistical analyses are different from the deposited data.

Moving on to the heritability estimate and the accompanying statistical details as reported in PK2010 ($F_{1,204} = 26.31$, $p < .001$, $\beta = 0.28$, SE = 1.71, $h^2 = 0.56$), we see that the standard error reported is far too large given the *F* and *p* reported. Furthermore, the heritability reported in PK2010 comes from a regression of mother against daughter phenotype, which is the opposite way around to what one should use to calculate heritability (Supplementary Material S5).

When we instead regress daughter against mother phenotypes, we estimate h^2 as $2 \times 0.44550 = 0.89$ ($F_{1,204} = 30.2$, $p < .001$; Supplementary Material S5). This is exceptionally high for a behavioural trait (Postma, 2014). Furthermore, it is substantially higher than the within-individual repeatability of satiation threshold of 0.56 (Pruitt, 2010); repeatability typically sets an upper limit to the heritability (Dohm, 2002). On this note, none of the statistical analyses in PK2010 account for the similarity of data collected on sisters, and

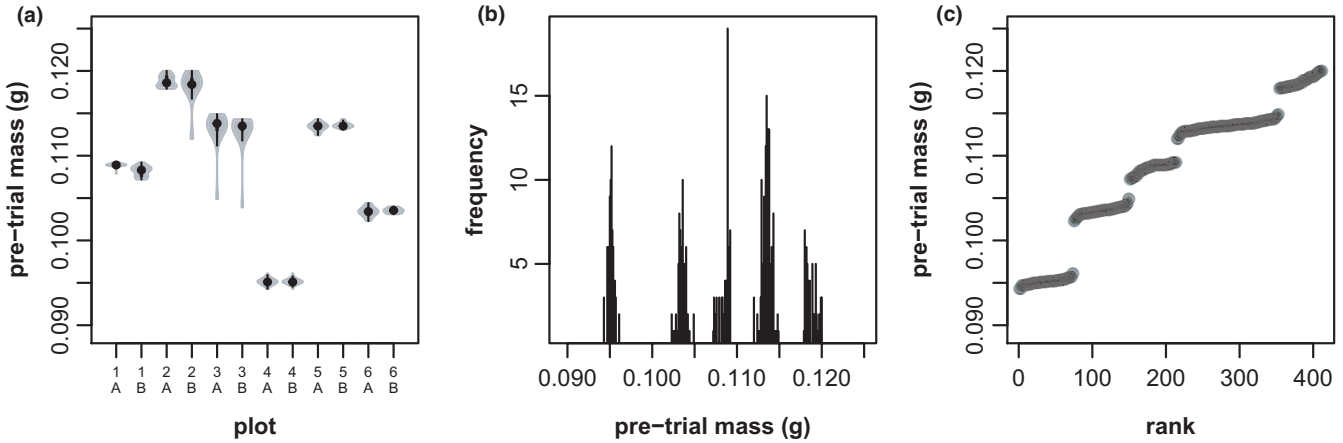

FIGURE 2 Distribution of pre-trial mass visualized in three different ways. See the text for more details

TABLE 1 Number of released individuals per plot and per treatment (A or B)

Plot	A	B
1	32	32
2	31	31
3	33	33
4	37	37
5	37	37
6	36	36

given the high reported heritability of satiation threshold, the analyses suffer from significant pseudoreplication.

7 | SUMMARY

PK2010 presents a series of analyses of the percentage increase in mass ('satiation threshold') in both wild-caught mothers and their captive-born daughters. Of the two measurements required to calculate this variable, only the pre-trial mass is available, and this only for the captive-born daughters. Furthermore, the sample sizes, distributions and patterns of decimal places in the raw data contain features that call into question whether the data were collected as described. Also, reanalysis of the archived data using simple statistical tests provides qualitatively and quantitatively different results from those reported.

This comment highlights the importance of data archiving for detecting errors or omissions in published research. We are glad that data archiving is becoming mandatory at a growing list of journals, and we encourage authors to publish their raw data and analysis code and provide it at the peer review stage (Culina et al., 2020; Roche et al., 2015). Making this the norm will help researchers, reviewers and readers to detect mistakes and clarify misunderstandings, with clear benefits to the rigour and reproducibility of research in all scientific disciplines, including evolutionary biology.

CONFLICT OF INTEREST

The authors have no conflict of interest to declare.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/jeb.13885.

OPEN RESEARCH BADGES

This article has been awarded Open Data Badge and Open Materials Badge. All materials and data are publicly accessible via the Open Science Framework at https://osf.io/gte2j/?view_only=db0913b1fc5042b88e7f2827f6d537a8

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in Dryad at https://doi.org/10.5061/dryad.1763. All R code is available as supplementary material.

ORCID

Erik Postma https://orcid.org/0000-0003-0856-1294
Alejandro Gonzalez-Voyer https://orcid.org/0000-0002-5072-1688
Luke Holman https://orcid.org/0000-0002-7268-2173

REFERENCES

Culina, A., van den Berg, I., Evans, S., & Sánchez-Tójar, A. (2020). Low availability of code in ecology: A call for urgent action. *PLOS Biology*, *18*, 1–9. Public Library of Science.

Dohm, M. R. (2002). Repeatability estimates do not always set an upper limit to heritability. *Functional Ecology*, *16*, 273–280.

Postma, E. (2014). Four decades of estimating heritabilities in wild vertebrate populations: Improved methods, more data, better estimates? In A. Charmantier, D. Garant, & L. E. B. Kruuk (Eds.), *Quantitative genetics in the wild*. Oxford University Press.

Pruitt, J. N. (2010). Differential selection on sprint speed and ad libitum feeding behaviour in active vs. sit-and-wait foraging spiders. *Functional Ecology*, *24*, 392–399.

Pruitt, J. N., & Krauel, J. J. (2010a). *Data from: The adaptive value of gluttony: Predators mediate the life history trade-offs of satiation threshold*. Dryad. https://doi.org/10.5061/dryad.1763

Pruitt, J. N., & Krauel, J. J. (2010b). The adaptive value of gluttony: Predators mediate the life history trade-offs of satiation threshold. *Journal of Evolutionary Biology*, *23*, 2104–2111.

Roche, D. G., Kruuk, L. E. B., Lanfear, R., & Binning, S. A. (2015). Public data archiving in ecology and evolution: How well are we doing? *PLOS Biology*, *13*, 1–12. Public Library of Science.

Roff, D. A., & Fairbairn, D. J. (2007). The evolution of trade-offs: Where are we? *Journal of Evolutionary Biology*, *20*, 433–447.

Stearns, S. C. (1992). *The evolution of life histories*. Oxford University Press.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Postma, E., Gonzalez-Voyer, A., & Holman, L. (2021). A comment on *The adaptive value of gluttony: predators mediate the life history trade-offs of satiation threshold* by Pruitt & Krauel (2010). *Journal of Evolutionary Biology*, 34, 1989–1993. https://doi.org/10.1111/jeb.13885