Preprint:
Winder, L.A., Simons, M.J.P. and Burke, T. (Submitted: 2023) The optimal clutch size revisited: separating individual quality from the costs of reproduction. [Preprint - eLife] (Submitted)

https://doi.org/10.7554/eLife.87018.1

Evolutionary Biology

The optimal clutch size revisited: separating individual quality from the costs of reproduction

Reviewed Preprint

Published from the original preprint after peer review and assessment by eLife.

About eLife's process

Reviewed preprint posted
August 22, 2023 (this version)

Posted to bioRxiv
May 30, 2023

Sent for peer review
March 21, 2023

Lucy A. Winder ✉, **Mirre J.P. Simons, Terry Burke**

Ecology & Evolutionary Biology, School of Biosciences, The University of Sheffield, Sheffield S10 2TN, UK

Abstract

Life-history theory, central to our understanding of diversity in morphology, behaviour and senescence, describes how traits evolve through the optimisation of trade-offs in investment. Despite considerable study, there is only minimal support for trade-offs within species between the two traits most closely linked to fitness – reproduction and survival – questioning the theory's general validity. We used a meta-analysis to separate quality effects from the costs of reproduction using studies of parental investment and survival in birds. Experimental enlargement of broods caused reduced parental survival. However, the effect of experimental enlargements was small and opposite to the effect of phenotypic quality, where individuals that naturally produced larger clutches also survived better. The opposite effects on survival in experimental and observational studies of parental care provides the first meta-analytic evidence for theory suggesting that quality differences mask trade-offs. Fitness projections using the overall effect size revealed that reproduction presented negligible costs, except when reproductive effort was forced beyond the level observed within species, to that seen between species. We conclude that there is little support for the most fundamental life-history trade-off, between reproduction and longevity, operating within a population. We suggest that within species, the fitness landscape of the reproduction– survival trade-off is flat until it reaches the boundaries of the between-species fast–slow life-history continuum. Our interpretation explains why the costs of reproduction are not apparent and why variation in reproductive output persists within species.

eLife assessment

In this potentially **useful** study, the authors attempt to use comparative meta-analysis to advance our understanding of life history evolution. Unfortunately, both the meta-analysis and the theoretical model is **inadequate** and proper statistical and mechanistic descriptions of the simulations are lacking. Specifically, the interpretation overlooks the effect of well-characterised complexities in the relationship between clutch size and fitness in birds.

Main

Across taxa, we see wide variation in life-history traits, such as the number of offspring produced and time spent raising young [1–3]. The central idea in life-history theory is that resources are finite, forcing trade-offs, meaning that investment in one aspect of life requires the sacrifice of another [4–7]. As reproduction is considered to be one of the most energetically demanding life stages, it is expected that within-species variation in offspring production will be driven by the cost of producing and raising young. It is thought that the fitness costs of reproduction are largely incurred as a decrement to survival, which would explain the fast– slow life-history continuum between reproduction and lifespan across species [6]. As reproduction and survival are the two components of life-history most closely related to fitness, this central trade-off has been the subject of much theoretical and empirical research.

Brood size manipulations of birds in natural conditions have provided arguably the best experimental paradigm in which to test the survival costs of reproduction. Experimental increases in brood size result in increased parental effort, suggesting that parents can typically cope with increased reproductive demands [8–12]. However, the expected increased costs of parental investment are not always detected and the current estimate across studies suggests only a small and inconsistent effect [13–16]. The absence of a cost of reproduction on survival means that costs must arise elsewhere or, alternatively, that individuals may differ in quality. Individuals may each be operating at their own maximum reproductive output, determined by their phenotypic condition, local or temporal genetic adaptation, and the surrounding environment [8–10,16]. The relative importance of the trade-off between reproduction and survival – central to life-history theory and the biology of ageing[12] – therefore remains unclear. In addition, the compelling theoretical explanation for the lack of an apparent trade-off due to the confounding effects of individual quality has not been investigated on a quantitative level [5,11,16].

Here, we present a meta-analysis that distinguishes between quality effects and the costs of reproduction. To do this, we tested how parental annual survival in birds is affected by the brood size they cared for in two different contexts: first in brood manipulation studies and, second, in observational studies of natural variation in clutch size. We expressed changes in survival on a per-egg basis, which, for the first time, allows for a quantitative comparison across studies. We find that quality is associated with higher survival chances, and that this effect is opposite but equal in magnitude to the costs of reproduction. The survival trade-off for offspring production within a population is therefore offset by differences in quality, potentially constraining the evolution of higher reproductive effort. Our analysis also uniquely allowed a quantitative comparison across species, as survival risk was expressed on a per-egg basis. We transformed the response variable, scaling for variance and mean, given that a per-egg increase in clutch size does not equate to the same proportional increase in parental effort for all species equally. Our findings suggest that species that generally lay smaller clutches are affected more severely by brood-size manipulations. This provides evidence that trade-offs are only detected when an individual is forced to perform far outside its optimum level.

To predict the evolutionary consequences of the effect sizes that we estimated using meta-analysis, we projected the fitness consequences for a change in brood-size life-history strategy. We found that the effects on parental survival translate into negligible fitness costs, with a relatively flat fitness landscape, suggesting that birds underproduce in terms of brood size, given the absence of fitness costs. This conclusion fits with our comparative finding that suggests that only brood sizes manipulated beyond the natural range incur substantial survival costs. Our results therefore suggest that, in wild populations, parental survival costs are, at most, a small component of the total fitness costs of investing parental effort. Our results do suggest that a cost of reproduction can be detected when an organism is pushed to the extreme of its reproductive capacity. We therefore infer that, though the survival– parental care trade-off does exist within species, it is too minimal to explain why variation in clutch size persists within a population. In addition, our work

shows that differences in individual quality counterbalance the trade-off between survival and reproduction, as previously theorised [5], and as such constrain reproductive effort and maintain clutch size variation in a population.

Results

The relationship between clutch size and survival was significantly different and opposite between observational and brood manipulation studies, irrespective of how brood size was scaled ($p < 0.01$, **Figure 1**, **Table 1**). Within observed natural variation, parents with larger clutches showed increased survival. In contrast, when broods were experimentally manipulated, the opposite relationship was found: increasing brood sizes decreased survival. Although the difference in overall effect size between experimental and natural variation in brood size was strongly significant in each comparison made, the individual overall effect sizes only became significant (from zero) when brood size was expressed as a proportional increase. Expressing brood size as a proportional increase corrects for the variation in average clutch size observed across the species included in this analysis, which ranged from 2 to 11. The parental effort required to raise two instead of one chick is potentially doubled, whereas one additional chick in a brood of 11 is likely to require only a marginal increase in effort. Indeed, also when using a between-species comparison, the effects of brood size manipulation and quality were strongest in the species that laid the smallest clutches, suggesting that costs to survival were only observed when a species was pushed beyond its natural limits (**Figure 2**, Supplementary Table 1).

Males and females did not differ in their survival response to changing clutch size (Supplementary Table 2, Supplementary Figure 1, contrary to Santos & Nakagawa 2012 [15]). The variance assigned to the random effects in the model was largely accounted for by study (Supplementary Table 3). Species accounted for more variation than the phylogeny, indicating that species vary in their survival for their brood size raised, irrespective of their shared evolutionary ancestry. However, our dataset included few closely related species, which reduces our ability to estimate phylogenetic effects (Supplementary Figure 2).

Projected fitness consequences of the costs of reproduction

From our meta-analysis we now have a quantifiable and comparable effect size for the survival costs of reproduction that we can use to predict its evolutionary consequences across a range of life histories. To this end, we projected the fitness consequences of increased reproductive effort, starting with the average effect size estimate per egg (**Table 1**) across a range of life histories, for a range of annual survival rates and clutch sizes (**Figure 3**). Overall, the effect size estimated in the meta-analysis (−0.05) resulted in a gain of fitness when reproductive output increased, especially in hypothetical species with low survival and small clutches. Conversely, the benefit of higher reproductive output was largely offset by the cost of survival when a species' survival rate and clutch size were high. When we increased the effect size upto five-fold, fitness costs of reproduction became more pronounced, but were still not present in species with small clutches and short lifespans.

Under long-term evolution these selection differentials should lead to individual hypothetical species moving towards the diagonal (bottom left to the right top corner). This diagonal represents the observed fast–slow pace of life continuum observed among species[17]. Exemplar species (i.e., with survival and average clutch size combinations observable in wild populations), for which we predicted the fitness consequences of the costs of reproduction, lie on this comparative diagonal in life history. In these exemplar species, the selection differential was observed to lie slightly above one, indicating that individuals having a higher clutch size than the species' average would gain a

	Parameter		Effect size	95% CI lower bounds	95% CI upper bounds	*p*	(individual)
Raw	Clutch size	Brood manipulation	-0.0522	-0.1406	0.0363	**0.0007**	(0.2477)
		Observational	0.0747	0.1571	-0.0288		(0.1571)
Standardised	Clutch size	Brood manipulation	-0.0651	-0.1478	0.0177	**0.0065**	(0.1232)
		Observational	0.1143	-0.0046	0.2333		(0.0595)
Proportional	Clutch size	Brood manipulation	-0.2703	-0.4984	-0.0423	**0.0005**	**(0.0202)**
		Observational	0.3850	0.0583	0.7116		**(0.0209)**

Table 1

Effect size estimates for the odds of survival with increasing clutch size (raw, standardised and proportional clutch size). The p-values indicate the difference between brood manipulations and observational data, with the individual effect *p*-values (from zero) in parentheses.

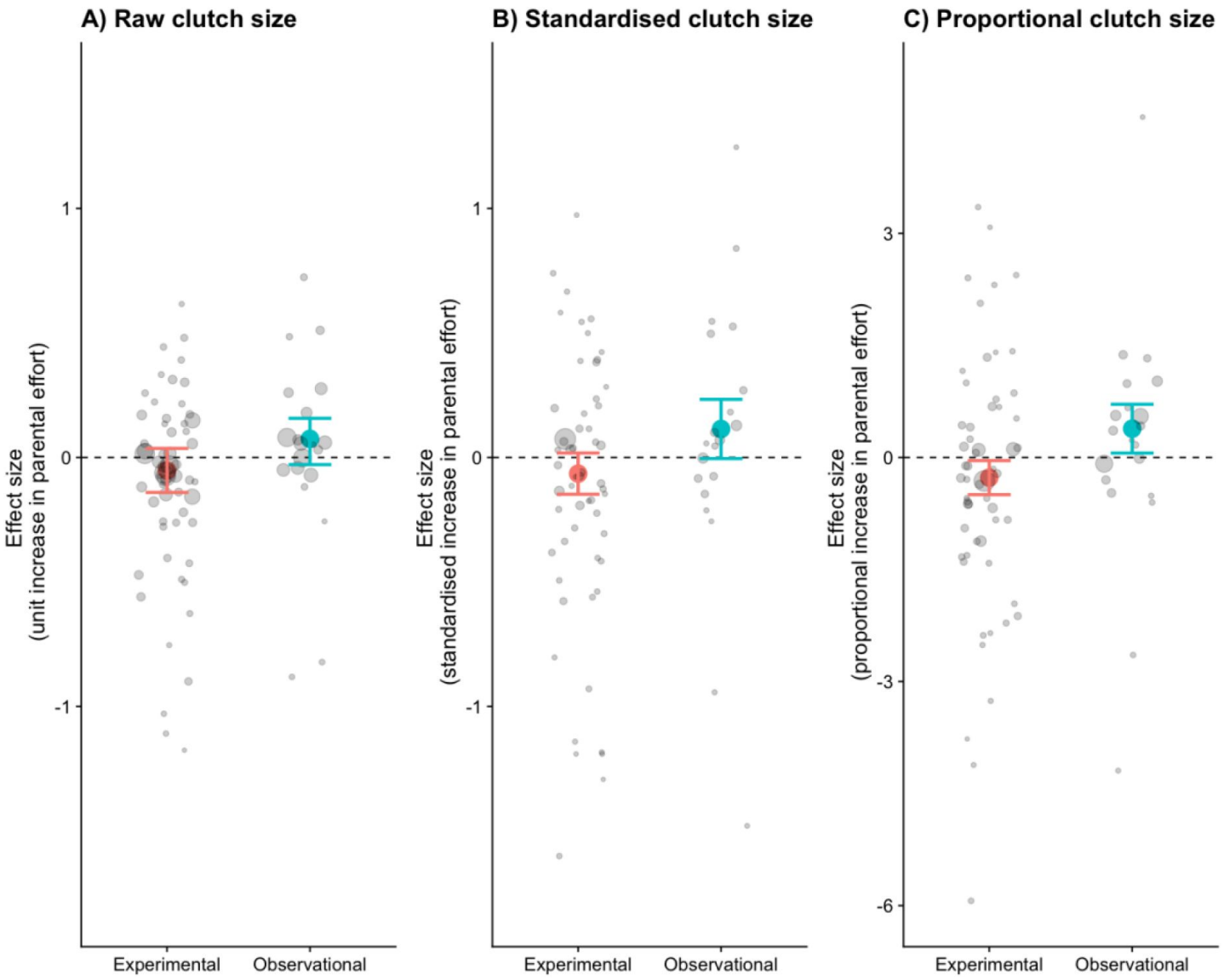

Figure 1

The effects size (log odds of survival) for three different measures of clutch size: (A) raw, (B) standardised, and (C) proportional clutch size. Coloured points are the combined effect sizes of the odds ratios with their 95% confidence intervals. Points are coloured by whether they represent brood manipulation experiments (costs of reproduction) or they are observational (quality). Grey points are the odds ratios of each study, with their sizes weighted by the variance.

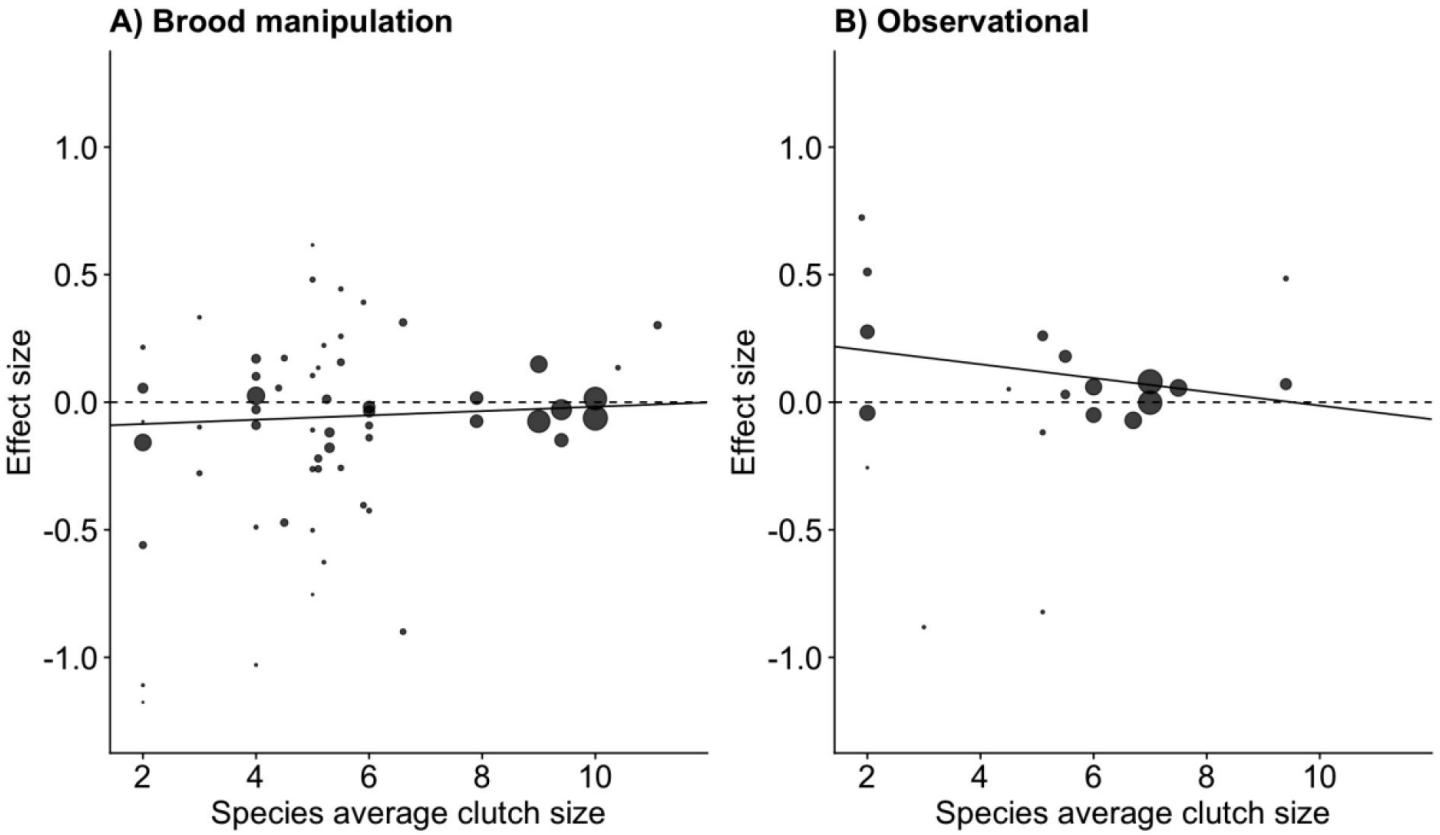

Figure 2

The meta-analytic linear regression (Table S1) of the effect of increasing clutch size (per egg) on parental survival, given the average clutch size for the species. Species with small clutch sizes showed stronger costs of reproduction and a stronger relationship with quality (p = 0.015). The points are the survival effect sizes (log odds ratio) per egg (as in **Figure 1A**) on parental survival in each study, with the point size reflecting the meta-analytic weight of that study.

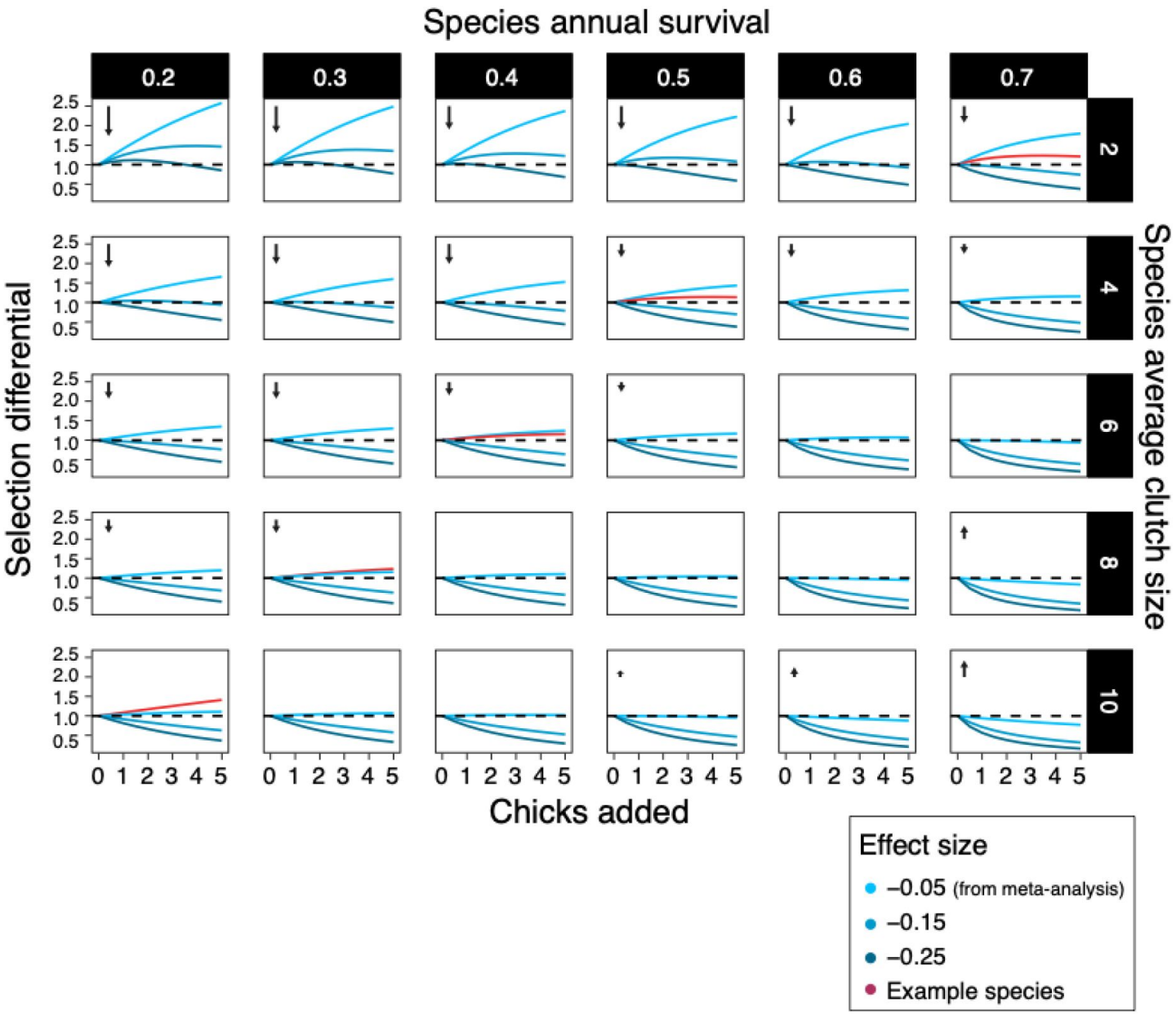

Figure 3

Isoclines of selection differentials among hypothetical control populations (in which individuals reproduce at the species' mean ratee) and hypothetical brood-manipulated populations (where individuals reproduce at an increased rate compared to control) for their whole lives. Selection differentials above 1 represents high lifetime fitness. Survival rates, clutch sizes, the magnitude of the manipulation (chicks added) and effect sizes represent the range of these variables present in the studies used in our meta-analysis. For each clutch size, we used a predicted survival rate and effect size to give isoclines that are biologically meaningful (exemplar birds shown in red). Arrows indicate the relative size and direction of selection in life-history space (on the reproduction axis). The costs of reproduction we estimated within species are predicted to result in a fast–slow life-history continuum across species, and the exemplar species we used as examples fit on this diagonal of survival rate/ clutch size combinations. We suggest that individual species show limited costs of reproduction, as they operate within relatively wide constraints imposed by the cost of reproduction that is responsible for the strong life-history trade-off observed across species.

slight fitness benefit. The fitness costs and benefits did, in general, not diverge substantially with the addition of chicks, but flattened, suggesting that the costs of survival counterbalance the benefits of reproduction across a range of reproductive outputs within a species.

The low costs of reproduction that we estimated could still be responsible for between-species life-history evolution, constraining species reproductive output and survival combination to fall along the diagonal of the fast–slow pace of life continuum. How selection pressures translate into short term and longer term evolutionary trajectories is uncertain. Often directional selection estimated in the wild does not translate to the inter-generational change on the population level[18]. Note, however, that only far away from the diagonal did our fitness projections reach a magnitude that would be predicted to lead to rapid evolutionary change [19] and see SI). The weak selection effects that lie on the diagonal are probably to be counterbalanced in the wild by factors such as environmental effects and genetic effects (e.g., gene flow from immigration or random mutation) [20]. We argue that within-species the minimal costs of reproduction, a flat fitness landscape and quality effects together explain why individuals appear to under-produce. Only when individuals are pushed beyond the observed between-species constraint do costs become apparent (**Figures 1**,**2**).

Our interpretation of the reproduction/lifespan life-history trade-off, based on our quantitative meta-analysis and subsequent fitness projections, explains several key observations and contradictions in the field. A strong trade-off is observed between species, but within species this trade-off is not apparent and variation in reproductive output is maintained within fitness boundaries similar to those that determine the between-species life history trade-off. The implication of this conclusion is that the costs of reproduction are likely to operate on a physiological level, but that the fitness consequences will remain largely flat over a species' observed variation in reproductive output. These effects are further obscured by the effects of quality, which are opposite in sign and magnitude to the cost of reproduction (**Figure 1**) and are likely to further flatten the fitness landscape.

Discussion

Our results provide the first meta-analytic evidence that differences in individual quality drive variation in clutch size. Here, we use the definition of quality as a combination of traits that give an individual higher fitness [5]. The reason selection is not acting on these individuals is currently unknown but it is likely that environmental variability leads to alternative phenotypes being selected for at different points in space and time (also discussed in [18]). It is possible that the quality effect could be representative of a terminal effect, where individuals have lower reproductive output in the year preceding their death, thereby driving the trend for naturally lower laying birds to have lower survival (e.g.,[21],[22]; also see [23] for age-related changes in reproductive output). The effect of birds having naturally-larger clutches was significantly opposite to the result of increasing clutch size through brood manipulation. For both costs of reproduction and quality effects, we found that species that laid the smallest clutches showed the largest effects. Brood manipulations that affect parental survival are thus likely to be the result of pushing parental effort beyond its natural limits. The classic trade-off between adult survival and the clutch size cared for is only apparent when an individual is forced to raise a clutch outside of its individual optimum, and these effects are confounded or even fully counterbalanced by differences in quality (as theorised in [5]).

Our fitness projections of the consequences of the costs of reproduction using the overall effect size we estimated suggest that, for current extant species, the within-species fitness landscape of the reproduction survival trade-off is flat. Species' life history decisions are constrained within a broader fast–slow life-history continuum, explaining why variation within species in reproductive effort, such as in clutch size, is large and near universal. Our interpretation assumes that other

fitness costs of reproduction are smaller or at least less relevant than survival costs. However, it is possible that such costs are important, such as effects on offspring quality (e.g.,[24,25]), parental condition other than survival (e.g., [26,27]) or future reproductive effort (e.g., [28]). Interestingly, the studies that have measured these different domains that contribute to fitness in brood-size manipulation studies concluded that only in combination do these costs result in balancing selection for the current most common brood size in the population [29,30]. Such classic trade-off explanations do, however, fail to explain why variation in reproductive effort is prevalent within species and why between-species life-history trade-offs appear so much stronger and conclusive. Our analysis and interpretation provide a novel explanation suggesting that, at its optimum, the within-species trade-off between survival and reproduction is relatively flat, and thus neutral to selection (supporting the theory presented in [16]). We suggest that the lack of evidence supporting trade-offs driving within-species variation does not necessarily mean that physiological costs of reproduction are non-existent (e.g., [24,31]), but rather that, within the wild and within the natural range of reproductive activities, such costs are not relevant to fitness. One key explanation for this effect supported by our meta-analysis and prior theory [5] is that individuals differ in quality.

Methods

Study sourcing & inclusion criteria

We used the following inclusion criteria (similar to [15]): the study must be on a wild population; must detail variation in the number of raised young (hereafter referred to as clutch size for simplicity) in relation to parental survival to the following year (including both experimental and observational studies) and must provide sample sizes. We did include studies where parental survival was reported for both parents combined. Excluded studies and the grounds for their removal are given in the supplementary information (Supplementary Table 4). We started by, first, extracting data from the included brood-manipulation studies and then searched the literature to include more recently-published studies (Supplementary Methods). In addition, we extracted data from studies that correlated variation in parental survival with natural variation in clutch size (observational studies). We aimed to pair each species in the brood manipulation studies with an observational study to ensure that the effects of quality were estimated across a similar range of species and so facilitate a more direct comparison. Where there was no equivalent study in the same species, we attempted to find a study of a congener. In most cases, observational data were obtained from either the same paper as the one describing brood manipulations, or via searching for other papers by the same authors. If this failed to produce observational data, a search was conducted following the same protocol as for the brood-manipulation experiments, but also specifying species, genus and/or common name in the search. Any additional brood manipulation studies or observational studies of different species found using this search were also included in the meta-analysis.

From the literature search, 78 individual effect sizes from 46 papers were used (20 observational and 58 experimental studies). While extracting these studies we also made note of the average clutch size of the species, the within-species standard deviation in clutch size and the longevity of the species. We extracted this information from the paper containing the study but if the information was missing, we searched other published literature with the aim to find the information from a similar population (i.e., at a similar latitude).

Extracting effect sizes

We used raw data to estimate an effect size by performing a logistic regression to obtain the log odds ratio for parental survival, given the clutch size (i.e., positive values indicate an increased chance of survival). Clutch size was averaged (mean) if a single estimate of survival was reported for multiple clutch sizes. 'Year' was included as an explanatory variable to correct for between-

year variation in adult survival, where data were presented for multiple years. We standardised the clutch size by the mean of the species and by the within-species standard deviation in clutch size. For species that have no within-species variance in clutch size, we used a value of 0.01 for the standard deviation in clutch size to prevent issues in calculations when using zero. We, therefore, expressed variation in clutch size in three ways: a raw increase in clutch size, a standardised clutch size and a proportional clutch size.

Meta-analysis

We ran a single model for each clutch size transformation to determine the cost of survival, given an increase in parental effort using the *metafor* package [32] in R 3.3.2 [33]. From these models we were also able to directly compare the effect size of brood manipulation studies and observational studies. We included phylogeny in these meta-analytic models to correct for shared ancestry. The phylogeny was obtained using BEAST to measure a distribution of 1,000 possible phylogenetic trees of the focal 30 species extracted from BirdTree [34]. We also included species and each studies' journal reference as random effects in the model. From these models, we calculated the proportion of variance explained by the phylogenetic effect [35].

We then tested the effect of the species' mean clutch size on the relationship between parental survival and clutch size. We ran a single model with the species' average clutch size in interaction with treatment (brood manipulation or observational). The clutch size was adjusted by the combined average clutch size of all the species used in the meta-analysis, subtracted from the species mean clutch size for each study. Species, phylogeny and reference were also included as random effects to correct for the similarity of effect sizes within species and studies.

The difference in survival for the different sexes was modelled for each clutch size measure. Brood manipulation studies and observational studies were analysed in separate models. Sex was modelled as a categorical moderator (41 female studies, 27 male studies and 10 mixed studies). Species, phylogeny and reference were included as random effects (Supplementary Table 2 and Supplementary Figure 1).

Publication bias

Much of the data used in this analysis were taken from studies where these data were not the main focus of the study. This reduces the risk that our results are heavily influenced by a publication bias for positive results. A funnel plot for the survival against raw clutch size model is presented in Supplementary Information (Supplementary Figure 3).

Fitness projections

We calculated various isoclines using the brood-manipulation overall effect size (based on raw clutch size) that we estimated. Here, an isocline is a trendline representing the change in fitness returns, given an increase in individual clutch size. An estimated lifetime reproductive fitness was calculated for hypothetical control populations, where all individuals consistently reproduce at the level of a species mean and have a consistent annual survival rate. We assumed species average clutch sizes to be 2, 4, 6, 8 and 10 and survival rates of 0.2, 0.3, 0.4, 0.5, 0.6 and 0.7, which reflected the range of clutch sizes and survival rates seen in the species in our meta-analysis. This lifetime reproductive fitness estimate was then repeated for a hypothetical population that reproduces at an increased level compared to control, i.e., brood size enlargement, throughout their lives. To obtain this, we added a range of 1–5 offspring to the clutch sizes of the control populations. Using a range of increased clutch sizes allowed us to investigate how increased reproductive effort would affect lifetime fitness. The survival costs were determined by the overall effect size found for brood manipulation studies (per egg). We modelled effect sizes of −0.05, −0.15 and −0.25, which represent, respectively, the meta-analytic overall effect size, its upper confidence interval and a further severe effect within the observed effect sizes (rounded to the closest 0.05 for

simplicity). We then calculated the selection differential ($LRS_{brood\ manipulation}$ / $LRS_{control}$) between the hypothetical control and "brood manipulation" populations for each combination of survival rate, clutch size and effect size, and plotted this as an isocline. We further plotted the fitness consequences for five exemplar species, where survival rate and clutch size combinations are observable in the wild. We used effect sizes from model predictions at these survival rate and clutch size combinations rather than the meta-analytic mean, thereby providing a biological context.

Acknowledgements

This work was supported by a Natural Environment Research Council (NERC) Adapting to the Challenges of a Changing Environment (ACCE) studentship to LAW, a Sir Henry Dale Fellowship (Wellcome and Royal Society; 216405/Z/19/Z) and an Academy of Medical Sciences Springboard Award (the Wellcome Trust, the Government Department of Business, Energy and Industrial Strategy (BEIS), the British Heart Foundation and Diabetes UK; SBF004\1085) to MJPS and a grant from NERC (NE/J024597/1) to TB.

References

1. Lack D. (1947) **The Significance of Clutch-size** *Ibis* **89**:302–352
2. Harvey P. H., Clutton-Brock T. H. (1985) **Life history variation in primates** *Evolution* **39**:559–559
3. Ricklefs R. E. (2000) **Density Dependence, Evolutionary Optimization, and the Diversification of Avian Life Histories** *Ornithol. Appl* **102**:9–22
4. Stearns S. C. (1976) **Life-History Tactics: A Review of the Ideas** *Q. Rev. Biol* **51**:3–47
5. van Noordwijk A. J., de Jong G. (1986) **Acquisition and Allocation of Resources: Their Influence on Variation in Life History Tactics** *Am. Nat* **128**:137–142
6. Kirkwood T. B. L., Rose M. R. (1991) **Evolution of senescence: late survival sacrificed for reproduction** *Philos. Trans. R. Soc. Lond. B. Biol. Sci* **332**:15–24
7. Lemaître J. F., Gaillard J. M. (2017) **Reproductive senescence: new perspectives in the wild** *Biol. Rev* **92**:2182–2199
8. Charnov E. L., Krebs J. R. (1974) **On Clutch-Size and Fitness** *Ibis* **116**:217–219
9. Pettifor R. A., Perrins C. M., McCleery R. H. (1988) **Individual optimization of clutch size in great tits** *Nature* **336**:160–162
10. Wilson A. J., Nussey D. H. (2010) **What is individual quality? An evolutionary perspective** *Trends Ecol. Evol* **25**:207–214
11. Descamps S., Gaillard J.-M., Hamel S., Yoccoz N. G. (2016) **When relative allocation depends on total resource acquisition: implication for the analysis of trade-offs** *J. Evol. Biol* **29**:1860–1866
12. Kirkwood T. B. L., Austad S. N. (2000) **Why do we age?** *Nature* **408**:233–238
13. Reznick D. (1985) **Costs of Reproduction : An Evaluation of the Empirical Evidence** *Oikos* **44**:257–267
14. Zera A. J., Harshman L. G. (2001) **The physiology of life history trade-offs in animals** *Annu. Rev. Ecology Syst* **32**:95–126
15. Santos E. S. A., Nakagawa S. (2012) **The costs of parental care: a meta-analysis of the trade-off between parental effort and survival in birds** *J. Evol. Biol* **25**:1911–1917
16. Cohen A. A., Coste C. F. D., Li X.-Y., Bourg S., Pavard S. (2020) **Are trade-offs really the key drivers of ageing and life span?** *Funct. Ecol* **34**:153–166
17. Healy K., Ezard T. H. G., Jones O. R., Salguero-Gómez R., Buckley Y. M. (2019) **Animal life history is shaped by the pace of life and the distribution of age-specific mortality and reproduction** *Nat. Ecol. Evol* **3**:1217–1224

18. Pujol B., et al. (2018) **The Missing Response to Selection in the Wild** *Trends Ecol. Evol* **33**:337–346

19. Gingerich P. D. (2009) **Rates of Evolution** *Annu. Rev. Ecol. Evol. Syst* **40**:657–675

20. Postma E., Visser J., Van Noordwijk A. J. (2007) **Strong artificial selection in the wild results in predicted small evolutionary change: Evolution is predictable, also in the wild** *J. Evol. Biol* **20**:1823–1832

21. Coulson J. C., Fairweather J. A. (2001) **Reduced reproductive performance prior to death in the Black-legged Kittiwake: senescence or terminal illness?** *J. Avian Biol* **32**:146–152

22. Rattiste K. (2004) **Reproductive success in presenescent common gulls (Larus canus): the importance of the last year of life** *Proc. R. Soc. Lond. B Biol. Sci* **271**:2059–2064

23. Hammers M., Richardson D. S., Burke T., Komdeur J. (2012) **Age-Dependent Terminal Declines in Reproductive Output in a Wild Bird** *PLoS ONE* **7**

24. Smith H. G., Kallander H., Nilsson J.-A. (1989) **The Trade-Off Between Offspring Number and Quality in the Great Tit Parus major** *J. Anim. Ecol* **58**:383–401

25. Conrad K. F., Robertson R. J. (1992) **Intraseasonal effects of clutch manipulation on parental provisioning and residual reproductive value of eastern phoebes (Sayornis phoebe)** *Oecologia* **89**:356–364

26. Reid W.V. (1987) **The cost of reproduction in the glaucous-winged gull** *Oecologia* **74**:458–467

27. Kalmbach E., Griffiths R., Crane J. E., Furness R. W. (2004) **Effects of experimentally increased egg production on female body condition and laying dates in the great skua Stercorarius skua** *J. Avian Biol* **35**:501–514

28. Järvistö P. E., Calhim S., Schuett W., Velmala W., Laaksonen T. (2016) **Sex-dependent responses to increased parental effort in the pied flycatcher** *Behav. Ecol. Sociobiol* **70**:157–169

29. Daan S., Dijkstra C., Tinbergen J. M. (1990) **Family planning in the kestrel (Falco Tinnunculus): The ultimate control of covariation of laying date and clutch size** *Behaviour* **114**:1–4

30. Verhulst S., Tinbergen J. M. (1991) **Experimental Evidence for a Causal Relationship between Timing and Success of Reproduction in the Great Tit Parus M. major** *J. Anim. Ecol* **60**:269–282

31. Lemaître J.-F., et al. (2015) **Early-late life trade-offs and the evolution of ageing in the wild** *Proc. Biol. Sci* **282**:20150209–20150209

32. Viechtbauer W. (2010) **Conducting meta-analyses in R with the metafor package** *J. Stat. Softw* **36**:1–48

33. R Development Core Team (2009) **R: A Language and Environment for Statistical Computing**

34. Rubolini D., Liker A., Garamszegi L. Z., Møller A. P., Saino N. (2015) **Using the birdtree.org website to obtain robust phylogenies for avian comparative studies: A primer** *Curr. Zool* **61**:959–965

35. Nakagawa S., Santos E. S. A. (2012) **Methodological issues and advances in biological meta-analysis** *Evol. Ecol* **26**:1253–1274

Author information

Lucy A. Winder
Ecology & Evolutionary Biology, School of Biosciences, The University of Sheffield, Sheffield S10 2TN, UK
For correspondence: lwinder1@sheffield.ac.uk
ORCID iD: 0000-0002-8100-0568

Mirre J.P. Simons
Ecology & Evolutionary Biology, School of Biosciences, The University of Sheffield, Sheffield S10 2TN, UK
ORCID iD: 0000-0001-7406-7708

Terry Burke
Ecology & Evolutionary Biology, School of Biosciences, The University of Sheffield, Sheffield S10 2TN, UK
ORCID iD: 0000-0003-3848-1244

Editors

Reviewing Editor
Ashleigh Griffin
University of Oxford, United Kingdom

Senior Editor
Christian Rutz
University of St Andrews, United Kingdom

Reviewer #1 (Public Review):

This paper falls in a long tradition of studies on the costs of reproduction in birds and its contribution to understanding individual variation in life histories. Unfortunately, the meta-analyses only confirm what we know already, and the simulations based on the outcome of the meta-analysis have shortcomings that prevent the inferences on optimal clutch size, in contrast to the claims made in the paper.

There was no information that I could find on the effect sizes used in the meta-analyses other than a figure listing the species included. In fact, there is more information on studies that were not included. This made it impossible to evaluate the data-set. This is a serious omission, because it is not uncommon for there to be serious errors in meta-analysis data sets. Moreover, in the long run the main contribution of a meta-analysis is to build a data set that can be included in further studies.

The main finding of the meta-analysis of the brood size manipulation studies is that the survival costs of enlarging brood size are modest, as previously reported by Santos & Nakagawa on what I suspect to be mostly the same data set. The paper does a very poor job of critically discussing whether we should take this at face value or whether instead there may be short-comings in the general experimental approach. A major reason why survival cost estimates are barely significantly different from zero may well be that parents do not fully

adjust their parental effort to the manipulated brood size, either because of time/energy constraints, because it is too costly and therefore not optimal, or because parents do not register increased offspring needs. Whatever the reason, as a consequence, there is usually a strong effect of brood size manipulation on offspring growth and thereby presumably their fitness prospects. In the simulations (Fig.4), the consequences of the survival costs of reproduction for optimal clutch size were investigated without considering brood size manipulation effects on the offspring. Effects on offspring are briefly acknowledged in the discussion, but otherwise ignored. Assuming that the survival costs of reproduction are indeed difficult to discern because the offspring bear the brunt of the increase in brood size, a simulation that ignores the latter effect is unlikely to yield any insight in optimal clutch size. It is not clear therefore what we learn from these calculations.

There are other reasons why brood size manipulations may not reveal the costs of reproduction animals would incur when opting for a larger brood size than they produced spontaneously themselves. Firstly, the manipulations do not affect the effort incurred in laying eggs (which also biases your comparison with natural variation in clutch size). Secondly, the studies by Boonekamp et al on Jackdaws found that while there was no effect of brood size manipulation on parental survival after one year of manipulation, there was a strong effect when the same individuals were manipulated in the same direction in multiple years. This could be taken to mean that costs are not immediate but delayed, explaining why single year manipulations generally show little effect on survival. It would also mean that most estimates of the fitness costs of manipulated brood size are not fit for purpose, because typically restricted to survival over a single year.

Details of how the analyses were carried out were opaque in places, but as I understood the analysis of the brood size manipulation studies, manipulation was coded as a covariate, with negative values for brood size reductions and positive values for brood size enlargements (and then variably scaled or not to control brood or clutch size). This approach implicitly assumes that the trade-off between current brood size (manipulation) and parental survival is linear, which contrasts with the general expectation that this trade-off is not linear. This assumption reduces the value of the analysis, and contrasts with the approach of Santos & Nakagawa.

The observational study selection is not complete and apparently no attempt was made to make it complete. This is a missed opportunity - it would be interesting to learn more about interspecific variation in the association between natural variation in clutch size and parental survival.

Reviewer #2 (Public Review):

I have read with great interest the manuscript entitled "The optimal clutch size revisited: separating individual quality from the costs of reproduction" by LA Winder and colleagues. The paper consists in a meta-analysis comparing survival rates from studies providing clutch sizes of species that are unmanipulated and from studies where the clutch sizes are manipulated, in order to better understand the effects of differences in individual quality and of the costs of reproduction. I find the idea of the manuscript very interesting. However, I am not sure the methodology used allows to reach the conclusions provided by the authors (mainly that there is no cost of reproduction, and that the entire variation in clutch size among individuals of a population is driven by "individual quality").

I write that I am not sure, because in its current form, the manuscript does not contain a single equation, making it impossible to assess. It would need at least a set of mathematical descriptions for the statistical analysis and for the mechanistic model that the authors infer from it.
The texts mixes concepts of individual vs population statistics, of within individual vs among-individuals measures, of allocation trade-offs and fitness trade-offs, etc ....which means it

would also require a glossary of the definitions the authors use for these various terms, in order to be evaluated.

This problem is emphasised by the following sentence to be found in the discussion "The effect of birds having naturally larger clutches was significantly opposite to the result of increasing clutch size through brood manipulation". The "effect" is defined as the survival rate (see Fig 1). While it is relatively easy to intuitively understand what the "effect" is for the unmanipulated studies: the sensitivity of survival to clutch size at the population level, this should be mentioned and detailed in a formula. Moreover, the concept of effect size is not at all obvious for the manipulated ones (effect of the manipulation? or survival rate whatever the manipulation (then how could it measure a trade-off ?)? at the population level? at the individual level ?) despite a whole appendix dedicated to it. This absolutely needs to be described properly in the manuscript.

Despite the lack of information about the underlying mechanistic model tested and the statistical model used, my impression is still that the interpretation in the introduction and discussion is not granted by the outputs of the figures and tables. Let's use a model similar to that of (van Noordwijk and de Jong, 1986): imagine that the mechanism at the population level is
a.c_(i,q)+b.s_(i,q)=E_q
Where c_(i,q) are s_(i,q) are respectively the clutch size for individual i which is of quality q, and E_q is the level of "energy" that an individual of quality q has available during the given time-step (and a and b are constants turning the clutch size and survival rate into energy cost of reproduction and energy cost of survival, and there are both quite "high" so that an extra egg (c_(i,q) is increased by 1) at the current time-step, decreases s_(i,q) markedly (E_q is independent of the number of eggs produced), that is, we have strong individual costs of reproduction). Imagine now that the variance of c_(i,q) (when the population is not manipulated) among individuals of the same quality group, is very small (and therefore the variance of s_(i,q) is very small also) and that the expectation of both are proportional to E_q. Then, in the unmanipulated population, the variance in clutch size is mainly due to the variance in quality. And therefore, the larger the clutch size c_(i,q) the higher E_q, and the higher the survival s_(i,q).
In the manipulated populations however, because of the large a and b, an artificial increase in clutch size, for a given E_q, will lead to a lower survival s_(i,q). And the "effect size" at the population level may vary according to a,b and the variances mentioned above. In other words, the costs of reproduction may be strong, but be hidden by the data, when there is variance in quality; however there are actually strong costs of reproduction (so strong actually that they are deterministic and that the probability to survive is a direct function of the number of eggs produced)

Moreover, it seems to me that the costs of reproduction are a concept closely related to generation time. Looking beyond the individual allocative (and other individual components of the trade-off) cost of reproduction and towards a populational negative relationship between survival and reproduction, we have to consider the intra-population slow fast continuum (some types of individuals survive more and reproduce less (are slower) than other (which are faster)). This continuum is associated with a metric: the generation time. Some individuals will produce more eggs and survive less in a given time-period because this time-period corresponds to a higher ratio of their generation time (Gaillard and Yoccoz, 2003; Gaillard et al., 2005). It seems therefore important to me, to control for generation time and in general to account for the time-step used for each population studied when analysing costs of reproduction. The data used in this manuscript is not just clutch size and survival rates, but clutch size per year (or another time step) and annual (or other) survival rates.

Finally, it is important to relate any study of the costs of reproduction in a context of individual heterogeneity (in quality for instance), to the general problem of the detection of

effects of individual differences on survival (see, e.g., Fay et al., 2021). Without an understanding of the very particular statistical behaviour of survival, associated to an event that by definition occurs only once per life history trajectory (by contrast to many other traits, even demographic, where the corresponding event (production of eggs for reproduction, for example) can be measured several times for a given individual during its life history trajectory).

References:
Fay, R. et al. (2021) 'Quantifying fixed individual heterogeneity in demographic parameters: Performance of correlated random effects for Bernoulli variables', Methods in Ecology and Evolution, 2021(August), pp. 1-14. doi: 10.1111/2041-210x.13728.
Gaillard, J.-M. et al. (2005) 'Generation time: a reliable metric to measure life-history variation among mammalian populations.', The American naturalist, 166(1), pp. 119-123; discussion 124-128. doi: 10.1086/430330.
Gaillard, J.-M. and Yoccoz, N. G. (2003) 'Temporal Variation in Survival of Mammals: a Case of Environmental Canalization?', Ecology, 84(12), pp. 3294-3306. doi: 10.1890/02-0409.
van Noordwijk, A. J. and de Jong, G. (1986) 'Acquisition and Allocation of Resources: Their Influence on Variation in Life History Tactics', American Naturalist, p. 137. doi: 10.1086/284547.

Reviewer #3 (Public Review):

The authors present here a comparative meta-analysis analysis designed to detect evidence for a reproduction/ survival trade-off, central to expectations from life history theory. They present variation in clutch size within species as an observation in conflict with expectations of optimisation of clutch size and suggest that this may be accounted for from weak selection on clutch size. The results of their analyses support this explanation - they found little evidence of a reproduction - survival trade-off across birds. They extrapolated from this result to show in a mathematical model that the fitness consequences of enlarged clutch sizes would only be expected to have a significant effect on fitness in extreme cases, outside of normal species' clutch size ranges. Given the centrality of the reproduction-survival trade-off, the authors suggest that this result should encourage us to take a more cautious approach to applying concepts the trade-off in life history theory and optimisation in behavioural ecology more generally. While many of the findings are interesting, I don't think the argument for a major re-think of life history theory and the role of trade-offs in fitness maximisation is justified.

The interest of the paper, for me, comes from highlighting the complexities of the link between clutch size and fitness, and the challenges facing biologists who want to detect evidence for life history trade-offs. Their results highlight apparently contradictory results from observational and experimental studies on the reproduction-survival trade-off and show that species with smaller clutch sizes are under stronger selection to limit clutch size.

Unfortunately, the authors interpret the failure to detect a life history trade-off as evidence that there isn't one. The construction of a mathematical model based on this interpretation serves to give this possible conclusion perhaps more weight than is merited on the basis of the results, of this necessarily quite simple, meta-analysis. There are several potential complicating factors that could explain the lack of detection of a trade-off in these studies, which are mentioned and dismissed as unimportant (lines 248-250) without any helpful, rigorous discussion. I list below just a selection of complexities which perhaps deserve more careful consideration by the authors to help readers understand the implications of their results:

• Reproductive output is optimised for lifetime reproductive success and so the consequences of being pushed off the optimum for one breeding attempt are not necessarily detectable in survival but in future reproductive success (and, therefore, lifetime reproductive success).

• The analyses include some species that hatch broods simultaneously and some that hatch sequentially (although this information is not explicitly provided (see below)). This is potentially relevant because species which have been favoured by selection to set up a size asymmetry among their broods often don't even try to raise their whole broods but only feed the biggest chicks until they are sated; any added chicks face a high probability of starvation. The first point this observation raises is that the expectation of more chicks= more cost, doesn't hold for all species. The second more general point is that the very existence of the sequential hatching strategy to produce size asymmetry in a brood is very difficult to explain if you reject the notion of a trade-off.
• For your standard, pair-breeding passerine, there is an expectation that costs of raising chicks will increase linearly with clutch size. Each chick requires X feeding visits to reach the required fledge weight. But this is not the case for species which lay precocious chicks which are relatively independent and able to feed themselves straight after hatching - so again the relationship of care and survival is unlikely to be detectable by looking at the effect of clutch size but again, it doesn't mean there isn't a trade-off between breeding and survival.
• The costs of raising a brood to adulthood for your standard pair-breeding passerine is bound to be extreme, simply by dint of the energy expenditure required. In fact, it was shown that the basal metabolic rate of breeding passerines was at the very edge of what is physiologically possible, the human equivalent being cycling the Tour de France (Nagy et al. 1990). If birds are at the very edge of what is physiologically possible, is it likely that clutch size is under weak selection?
• Variation in clutch size is presented by the authors as inconsistent with the assumption that birds are under selection to lay the Lack clutch. Of course, this is absurd and makes me think that I have misunderstood the authors' intended point here. At any rate, the paper would benefit from more clarity about how variable clutch size has to be before it becomes a problem for optimality in the authors' view (lines 84-85; line 246). See Perrins (1965) for an exquisite example of how beautifully great tits optimise clutch size on average, despite laying between 5-12 eggs.

Please note that the reviews of this paper were performed prior to data files supporting this manuscript being made publicly available. Supporting data can now be found here: https://doi.org/10.5061/dryad.q83bk3jnk

Author Response

Reviewer #1 (Public Review):

This paper falls in a long tradition of studies on the costs of reproduction in birds and its contribution to understanding individual variation in life histories. Unfortunately, the meta-analyses only confirm what we know already, and the simulations based on the outcome of the meta-analysis have shortcomings that prevent the inferences on optimal clutch size, in contrast to the claims made in the paper.

There was no information that I could find on the effect sizes used in the meta-analyses other than a figure listing the species included. In fact, there is more information on studies that were not included. This made it impossible to evaluate the data-set. This is a serious omission, because it is not uncommon for there to be serious errors in meta-analysis data sets. Moreover, in the long run the main contribution of a meta-analysis is to build a data set that can be included in further studies.

It is disappointing that two referees comment on data availability, as we supplied a link to our full dataset and the code we used in Dryad with our submitted manuscript. We were also

asked to supply our data during the review process and we again supplied a link to our dataset and code, along with a folder containing the data and code itself. We received confirmation that the reviewers had been given our data and code. We support open science and it was our intention that our dataset should be fully available to reviewers and readers. Our data and code are at https://doi.org/10.5061/dryad.q83bk3jnk.

> *The main finding of the meta-analysis of the brood size manipulation studies is that the survival costs of enlarging brood size are modest, as previously reported by Santos & Nakagawa on what I suspect to be mostly the same data set.*

We disagree that the main finding of our paper is the small survival cost of manipulated brood size. The major finding of the paper, in our opinion, is that the effect sizes for experimental and observational studies are in opposite directions, therefore providing the first quantitative evidence to support the influential theoretical framework put forward by van Noordwijk and de Jong (1986), that individuals differ in their optimal clutch size and are constrained to reproducing at this level due to a trade-off with survival. We show that while the manipulation experiments have been widely accepted to be informative, they are not in fact an effective test of whether within-species variation in clutch size is the result of a trade-off between reproduction and survival.

The comment that we are reporting the same finding as Santos & Nakagawa (2012) is a misrepresentation of both that study and our own. Santos & Nakagawa found an effect of parental effort on survival only in males who had their clutch size increased – but no effect for males who had their clutch size reduced and no survival effect on females for either increasing or reducing parental effort. However, we found an overall reduction in survival for birds who had brood sizes manipulated to make them larger (for both sexes and mixed sex studies combined). In our supplementary information, we demonstrate the overall survival effect of a change in reproductive effort to be close to zero for males, negative (though non-significant) for females and significantly negative for mixed sexes (which are not included in the Santos & Nakagawa study).

> *The paper does a very poor job of critically discussing whether we should take this at face value or whether instead there may be short-comings in the general experimental approach. A major reason why survival cost estimates are barely significantly different from zero may well be that parents do not fully adjust their parental effort to the manipulated brood size, either because of time/energy constraints, because it is too costly and therefore not optimal, or because parents do not register increased offspring needs. Whatever the reason, as a consequence, there is usually a strong effect of brood size manipulation on offspring growth and thereby presumably their fitness prospects. In the simulations (Fig.4), the consequences of the survival costs of reproduction for optimal clutch size were investigated without considering brood size manipulation effects on the offspring. Effects on offspring are briefly acknowledged in the discussion, but otherwise ignored. Assuming that the survival costs of reproduction are indeed difficult to discern because the offspring bear the brunt of the increase in brood size, a simulation that ignores the latter effect is unlikely to yield any insight in optimal clutch size. It is not clear therefore what we learn from these calculations.*

The reviewer's comment is somewhat of a paradox. We take the best studied example of the trade-off between reproductive effort and parental survival, a key theme in life-history and the biology of ageing, and subject this to a meta-analysis. The reviewer suggests we should interpret our finding as if there must be something wrong with the method or studies we included, rather than maybe considering the original hypothesis could be false or inflated in importance. The reviewer's inclination to question the premise of the data in favor of a held hypothesis we consider not necessarily the best scientific approach here. In many places in

our manuscript do we question and address issues in the underlying data and interpretation (L101-105, L149-150, 182-185 and L229-233). Moreover, we make it clear that we focus on the trade-off between current reproductive effort and subsequent parental survival and we are aware that other trade-offs could counter-balance or explain our findings, discussed on L189-191 & L246-253. Note that it is also problematic, when you do not find the expected response, to search for an alternative that has not been measured. In the case here, with trade-offs, there are endless possiblilities of where a trade-off might be incurred between traits. We purposfully focus on the one well-studied and theorised trade-off. We clearly acknowledge though that when all possible trade-offs are taken into account a trade-off on the fitness level can occur and cite two famous studies (Daan et al., 1990 and Verhulst & Tinbergen 1991) that have done just that (L250-253).

So whilst, we agree with the reviewer that the offspring may incur costs themselves, rather than costs being incurred by the parents, the aim of our study was to test for a generalised trend across species in the survival costs of reproductive effort. It is unrealistic to suggest that incorporating offspring growth into our simulations would add insight, as a change in offspring number rarely affects all offspring in the nest equally and there can even be quite stark differences; for example this will be most evident in species that produce sacrificial offspring. This effect will be further confounded by catch-up growth, for example, and so it is likely that increased sibling competition from added chicks alters offspring growth trajectories, rather than absolute growth as the reviewer suggests. There are mixed results in the literature on the effect of altering clutch size on offspring survival, with an increased clutch size through manipulation often increasing the number of recruits from a nest.

> *There are other reasons why brood size manipulations may not reveal the costs of reproduction animals would incur when opting for a larger brood size than they produced spontaneously themselves. Firstly, the manipulations do not affect the effort incurred in laying eggs (which also biases your comparison with natural variation in clutch size). Secondly, the studies by Boonekamp et al on Jackdaws found that while there was no effect of brood size manipulation on parental survival after one year of manipulation, there was a strong effect when the same individuals were manipulated in the same direction in multiple years. This could be taken to mean that costs are not immediate but delayed, explaining why single year manipulations generally show little effect on survival. It would also mean that most estimates of the fitness costs of manipulated brood size are not fit for purpose, because typically restricted to survival over a single year.*

First, our results did show a survival cost of reproduction for brood manipulations. We agree that there could be longer-term costs, and so our estimate of the survival cost for manipulated birds is likely to be an underestimate, meaning that our interpretation still holds – the cost to reproduce prevents individuals from laying beyond their optimal level. Note, however, that much theory is build on the immediate costs of reproduction and as such these costs are likely overinterpreted.

We agree with the reviewer that lifetime manipulations could be even more informative than single-year manipulations. Unfortunately, there are currently too few studies available to be able to draw generalisable conclusions across species for lifetime manipulations. This is, however, the reason we used lifetime change in clutch size in our fitness projections, which the reviewer seems to have missed – please see methods line 360-362, where we explicitly state that this is lifetime enlargement. Of course such interpretations do not include an accumulation of costs that is greater than the annual cost, but currently there is no clear evidence that such an assumption is valid. Such a conclusion can also not be drawn from the study on jackdaws by Boonekamp et al (2014) as the treatments were life-long and, therefore, cannot separate annual from accrued (multiplicative) costs that are more than the sum of annual costs incurred.

> *Details of how the analyses were carried out were opaque in places, but as I understood the analysis of the brood size manipulation studies, manipulation was coded as a covariate, with negative values for brood size reductions and positive values for brood size enlargements (and then variably scaled or not to control brood or clutch size). This approach implicitly assumes that the trade-off between current brood size (manipulation) and parental survival is linear, which contrasts with the general expectation that this trade-off is not linear. This assumption reduces the value of the analysis, and contrasts with the approach of Santos & Nakagawa.*

We thank the reviewer for highlighting a lack of clarity in places in our methods. We will add additional detail to this section in our revised manuscript.

For clarity in our response, each effect size was extracted by performing a logistic regression with survival as a binary response variable and clutch size was the absolute value of offspring in the nest (i.e., for a bird who laid a clutch size of 5 but was manipulated to have -1 egg, we used a clutch size value of 4). The clutch size was also standardised and, separately, expressed as a proportion of the species mean.

We disagree that our approach reduces the value of our analysis. First, our approach allows a direct comparison between experimental and observational studies, which is the novelty of our study. Our approach does differ from Santos & Nakagawa but we disagree that it contrasts. Our approach allows us to take into consideration the severity of the change in clutch size, which Santos & Nakagawa do not. Therefore, we do not agree that our approach is worse at accounting for non-linearity of trade-offs than the approach used by Santos & Nakagawa.

Our analysis, alongside a plethora of other ecological studies, does assume that the response to our predictor variable is linear. However, it is common knowledge that there are very few (if any) truly linear relationships. We use linear relationships because they serve a good approximation of the trend and provide a more rigorous test for an underlying relationship than would fitting nonlinear models. For many datasets there is not a range of chicks added for which a non-linear relationship could be estimated. The question also remains of what the shape of this non-linear relationship should be and is hard to determine a priori. We will address non-linear effects in our revised manuscript.

> *The observational study selection is not complete and apparently no attempt was made to make it complete. This is a missed opportunity - it would be interesting to learn more about interspecific variation in the association between natural variation in clutch size and parental survival.*

We clearly state in our manuscript that we deliberately made a tailored selection of studies that matched the manipulation studies (L279-282). We paired species extracted for observational studies with those extracted in experimental studies to facilitate a direct comparison between observational and experimental studies, and to ensure that the respective datasets were comparable. The reviewer's focus in this review seems to be solely on the experimental dataset. This comment dismisses the observational component of our analysis and thereby fails to acknowledge the question being addressed in this study.

Reviewer #2 (Public Review):

I have read with great interest the manuscript entitled "The optimal clutch size revisited: separating individual quality from the costs of reproduction" by LA Winder and colleagues. The paper consists in a meta-analysis comparing survival rates from studies providing clutch sizes of species that are unmanipulated and from studies where the

> *clutch sizes are manipulated, in order to better understand the effects of differences in individual quality and of the costs of reproduction. I find the idea of the manuscript very interesting. However, I am not sure the methodology used allows to reach the conclusions provided by the authors (mainly that there is no cost of reproduction, and that the entire variation in clutch size among individuals of a population is driven by "individual quality").*

We would like to highlight that we do not conclude that there is no cost of reproduction. Please see lines 258–260, where we state that our lack of evidence for trade-offs driving within-species variation in clutch size does not necessarily mean the costs of reproduction are non-existent. We conclude that individuals are constrained to their optima by the survival cost of reproduction. It is also an over-statement of our conclusion to say that we believe that variation in clutch size is only driven by quality. Our results show that unmanipulated birds who have larger clutch sizes also live longer, and we suggest this is evidence that some individuals are “better” than others, but we do not say, nor imply, that no other factors affect variation in clutch size.

> *I write that I am not sure, because in its current form, the manuscript does not contain a single equation, making it impossible to assess. It would need at least a set of mathematical descriptions for the statistical analysis and for the mechanistic model that the authors infer from it.*

We appreciate this comment, but this is the first time we have been asked to put equations in a manuscript rather than explain them in terms that are accessible to a wider audience. Note however that our meta-analysis is standard and based on logistic regression and standard meta-analytic practices. We do not think we need to repeat such equations and we cite the relevant data. For the simulation, we simply simulated the resulting effects and this is not something that we feel is captured more accurately in equations rather than in text and the associated graphs. We of course supplied our code for this along with our manuscript (https://doi.org/10.5061/dryad.q83bk3jnk), though as we mentioned above, we believe this was not shared with the reviewers despite us making this available for the review process. We therefore understand the reviewer feels the simulations were not explained thoroughly. We will revise our text to see if we can add additional explanation where relevant in our revision.

> *The texts mixes concepts of individual vs population statistics, of within individual vs among-individuals measures, of allocation trade-offs and fitness trade-offs, etc ….which means it would also require a glossary of the definitions the authors use for these various terms, in order to be evaluated.*

We would like to thank the reviewer for highlighting this lack of clarity in our text. We will simplify the terminology and define terms in our revised manuscript.

> *This problem is emphasised by the following sentence to be found in the discussion "The effect of birds having naturally larger clutches was significantly opposite to the result of increasing clutch size through brood manipulation". The "effect" is defined as the survival rate (see Fig 1). While it is relatively easy to intuitively understand what the "effect" is for the unmanipulated studies: the sensitivity of survival to clutch size at the population level, this should be mentioned and detailed in a formula. Moreover, the concept of effect size is not at all obvious for the manipulated ones (effect of the manipulation? or survival rate whatever the manipulation (then how could it measure a trade-off ?)? at the population level? at the individual level ?) despite a whole appendix dedicated to it. This absolutely needs to be described properly in the manuscript.*

We would like to thank the reviewer for bringing to our attention the lack of clarity on the details of our methodology. We will make this more clear in our revised manuscript.

For clarity, the effect size for both manipulated and unmanipulated nests was survival, given the brood size raised. We performed a logistic regression with survival as a binary response variable (i.e., number of individuals that survived and number of individuals that died after each breeding season), and clutch size was the absolute value of offspring in the nest (i.e., for a bird who laid a clutch size of 5 but was manipulated to have -1 egg, we used a clutch size value of 4). This allows for direct comparison of the effect size (survival given clutch size raised) between manipulated and unmanipulated birds.

> *Despite the lack of information about the underlying mechanistic model tested and the statistical model used, my impression is still that the interpretation in the introduction and discussion is not granted by the outputs of the figures and tables. Let's use a model similar to that of (van Noordwijk and de Jong, 1986): imagine that the mechanism at the population level is*
>
> *a.c_(i,q)+b.s_(i,q)=E_q*
>
> *Where c_(i,q) are s_(i,q) are respectively the clutch size for individual i which is of quality q, and E_q is the level of "energy" that an individual of quality q has available during the given time-step (and a and b are constants turning the clutch size and survival rate into energy cost of reproduction and energy cost of survival, and there are both quite "high" so that an extra egg (c_(i,q) is increased by 1) at the current time-step, decreases s_(i,q) markedly (E_q is independent of the number of eggs produced), that is, we have strong individual costs of reproduction). Imagine now that the variance of c_(i,q) (when the population is not manipulated) among individuals of the same quality group, is very small (and therefore the variance of s_(i,q) is very small also) and that the expectation of both are proportional to E_q. Then, in the unmanipulated population, the variance in clutch size is mainly due to the variance in quality. And therefore, the larger the clutch size c_(i,q) the higher E_q, and the higher the survival s_(i,q).*
>
> *In the manipulated populations however, because of the large a and b, an artificial increase in clutch size, for a given E_q, will lead to a lower survival s_(i,q). And the "effect size" at the population level may vary according to a,b and the variances mentioned above. In other words, the costs of reproduction may be strong, but be hidden by the data, when there is variance in quality; however there are actually strong costs of reproduction (so strong actually that they are deterministic and that the probability to survive is a direct function of the number of eggs produced)*

We would like to thank the reviewer for these comments. Please note that our simulations only take the experimental effect of brood size on parental survival into account. Our model does not incorporate quality effects. The reviewer is right that the relationship between quality and the effects exposed by manipulating brood size can take many forms and this is a very interesting topic, but not one we aimed to tackle in our manuscript. In terms of quality we make two points: 1) overall quality effects connecting reproduction and parental survival are present 2) these effects are opposite in direction to the effects when reproduction is manipulated and similar in magnitude. We do not go further than that in interpreting our results. The reviewer is right however that we do suggest and repeat suggestions by others that quality can also mask the trade-off in some individuals or circumstances (L63-65, L85-88 & L237-240), but we do not quantify this as this is dependent on the unknown relationships between quality and the response to the manipulation. A focussed set of experiments in that context would be interesting and there is some data that could get at this, i.e. the relationship between produced clutch size and the relative effect of the manipulation. Such information is however not available for all studies and although we explored also analyzing this, currently

this is not possible to do with sufficient confidence. We will include this rationale in our revision.

> *Moreover, it seems to me that the costs of reproduction are a concept closely related to generation time. Looking beyond the individual allocative (and other individual components of the trade-off) cost of reproduction and towards a populational negative relationship between survival and reproduction, we have to consider the intra-population slow fast continuum (some types of individuals survive more and reproduce less (are slower) than other (which are faster)). This continuum is associated with a metric: the generation time. Some individuals will produce more eggs and survive less in a given time-period because this time-period corresponds to a higher ratio of their generation time (Gaillard and Yoccoz, 2003; Gaillard et al., 2005). It seems therefore important to me, to control for generation time and in general to account for the time-step used for each population studied when analysing costs of reproduction. The data used in this manuscript is not just clutch size and survival rates, but clutch size per year (or another time step) and annual (or other) survival rates.*

The reviewer is right that this is interesting. There has been unexplained difference in temperate (seasonal) and tropical reproduction strategies. Most of our data come from seasonal breeders however. Although there is some variation in second brooding and such often these species only produce one brood. We do agree that a wider consideration here is relevant, but we are not trying to explain all of life-history in our paper. It is clearly the case that other factors will operate and the opportunity for trade-offs will vary among species according to their respective life histories. However, our study focuses on the two most fundamental components of fitness – longevity and reproduction – to test a major hypothesis in the field, and we uncover new relationships that contrast with previous influential studies, and cast doubt on previous conclusions. We question the assumed trade-off between reproduction and annual survival. We show quality is important and that the effect we find in experimental studies, is so small that it can only explain between-species patterns but is unlikely to be the selective force that constrains reproduction within-species. We do agree that there is a lot more work that can be done in this area. We hope we contribute to this, by questioning this central trade-off. We will try and incorporate some of these suggestions in the revision where possible.

> *Finally, it is important to relate any study of the costs of reproduction in a context of individual heterogeneity (in quality for instance), to the general problem of the detection of effects of individual differences on survival (see, e.g., Fay et al., 2021). Without an understanding of the very particular statistical behaviour of survival, associated to an event that by definition occurs only once per life history trajectory (by contrast to many other traits, even demographic, where the corresponding event (production of eggs for reproduction, for example) can be measured several times for a given individual during its life history trajectory).*

Thank you for raising this point. The reviewer is right that heterogeneity can dampen or augment selection. Note that by estimating the effect of quality here we give an example of how heterogeneity can possibly do exactly this. We thank the reviewer for raising that we should possibly link this to wider effects of heterogeneity and we aim to do so in the revision.

References:

Fay, R. et al. (2021) 'Quantifying fixed individual heterogeneity in demographic parameters: Performance of correlated random effects for Bernoulli variables', Methods in Ecology and Evolution, 2021(August), pp. 1-14. doi: 10.1111/2041-210x.13728.

Gaillard, J.-M. et al. (2005) 'Generation time: a reliable metric to measure life-history variation among mammalian populations.', The American naturalist, 166(1), pp. 119-123; discussion 124-128. doi: 10.1086/430330.

Gaillard, J.-M. and Yoccoz, N. G. (2003) 'Temporal Variation in Survival of Mammals: a Case of Environmental Canalization?', Ecology, 84(12), pp. 3294-3306. doi: 10.1890/02-0409.

van Noordwijk, A. J. and de Jong, G. (1986) 'Acquisition and Allocation of Resources: Their Influence on Variation in Life History Tactics', American Naturalist, p. 137. doi: 10.1086/284547.

Reviewer #3 (Public Review):

The authors present here a comparative meta-analysis analysis designed to detect evidence for a reproduction/ survival trade-off, central to expectations from life history theory. They present variation in clutch size within species as an observation in conflict with expectations of optimisation of clutch size and suggest that this may be accounted for from weak selection on clutch size. The results of their analyses support this explanation - they found little evidence of a reproduction - survival trade-off across birds. They extrapolated from this result to show in a mathematical model that the fitness consequences of enlarged clutch sizes would only be expected to have a significant effect on fitness in extreme cases, outside of normal species' clutch size ranges. Given the centrality of the reproduction-survival trade-off, the authors suggest that this result should encourage us to take a more cautious approach to applying concepts the trade-off in life history theory and optimisation in behavioural ecology more generally. While many of the findings are interesting, I don't think the argument for a major re-think of life history theory and the role of trade-offs in fitness maximisation is justified.

The interest of the paper, for me, comes from highlighting the complexities of the link between clutch size and fitness, and the challenges facing biologists who want to detect evidence for life history trade-offs. Their results highlight apparently contradictory results from observational and experimental studies on the reproduction-survival trade-off and show that species with smaller clutch sizes are under stronger selection to limit clutch size.

Unfortunately, the authors interpret the failure to detect a life history trade-off as evidence that there isn't one. The construction of a mathematical model based on this interpretation serves to give this possible conclusion perhaps more weight than is merited on the basis of the results, of this necessarily quite simple, meta-analysis. There are several potential complicating factors that could explain the lack of detection of a trade-off in these studies, which are mentioned and dismissed as unimportant (lines 248-250) without any helpful, rigorous discussion. I list below just a selection of complexities which perhaps deserve more careful consideration by the authors to help readers understand the implications of their results:

We would like to thank the reviewer for their thoughtful response and summary of the findings we also agree are central to our study. The reviewer also highlights areas where our manuscript could benefit from a deeper discussion and we will add detail to our discussion in our revised manuscript.

We would like to highlight that we do not interpret the failure to detect a trade-off as evidence that there isn't one. First, and importantly, we do find a trade-off but show this is only incurred when individuals lay beyond their optimal level. Secondly, we also state on lines 258-260 that the lack of evidence to support trade-offs being strong enough to drive variation in clutch size does not necessarily mean there are no costs of reproduction.

The statement that we have constructed a mathematical model based on the interpretation that we have not found a trade-off is, again, factually incorrect. We ran these simulations because the opposite is true – we did find a trade-off. There is a significant effect of clutch size when manipulated on annual parental survival. To appreciate whether this effect alone can explain why reproduction is constrained, we ran the simulations. From these simulations we find that this effect size is too small to explain the constraint so something else must be going on and we do spend a considerable amount of text discussing the possible explanations (L182-194). Note the possibly most parsimonious conclusion here is that costs of reproduction are not there so we also give that explanation some thought (L201-205 and L247-253).

We are disappointed by the suggestion that we have dismissed complicating factors which could prevent detection of a trade-off, as this was not our intention. We were aiming to highlight that what we have demonstrated to be an apparent trade-off can be explained through other mechanisms, and that the trade-off between clutch size and survival is not as strong in driving within-species variation in clutch size as previously assumed. We will add further discussion to our revised manuscript to make this clear and give readers a better understanding of the complexity of factors associated with life-history theory. Although we do feel we have addressed this (L248-255).

> *• Reproductive output is optimised for lifetime reproductive success and so the consequences of being pushed off the optimum for one breeding attempt are not necessarily detectable in survival but in future reproductive success (and, therefore, lifetime reproductive success).*

We agree this is a valid point, which is mentioned in our manuscript in terms of alternative stages where the costs of reproduction might be manifested (L248-250). We would also like to highlight that in our simulations, the change in clutch size (and subsequent survival cost) was assumed for the lifetime of the individual, for this very reason.

> *• The analyses include some species that hatch broods simultaneously and some that hatch sequentially (although this information is not explicitly provided (see below)). This is potentially relevant because species which have been favoured by selection to set up a size asymmetry among their broods often don't even try to raise their whole broods but only feed the biggest chicks until they are sated; any added chicks face a high probability of starvation. The first point this observation raises is that the expectation of more chicks= more cost, doesn't hold for all species. The second more general point is that the very existence of the sequential hatching strategy to produce size asymmetry in a brood is very difficult to explain if you reject the notion of a trade-off.*

We agree with the reviewer that the costs of reproduction can be absorbed by the offspring themselves, and may not be equal across offspring (we also highlight this at L249 in the manuscript). However, we disagree that for some species the addition of more chicks does not equate to an increase in cost, though we do accept this might be less for some species. This is, however, difficult to incorporate into a sensible model as the impacts will vary among species and some species do also exhibit catch-up growth. So without a priori knowledge on this we kept our model simple. To test whether the effect on parental survival (often assumed to be a strong cost) can explain the constraint on reproductive effort, and we conclude it does not.

We would also like to make clear that we are not rejecting the notion of a trade-off. Our study shows evidence that a trade-off between survival and reproductive effort likely does not drive within-species variation in clutch size. We do explicitly say this throughout our manuscript, and also provide suggestions of other areas where a trade-off may exist (L246-250). The point of our study is not whether trade-offs exist or not, it is whether there is a

generalisable across-species trend for a trade-off between reproductive effort and survival – the most fundamental trade-off in our field but for which there is a lack of conclusive evidence within species.

> *• For your standard, pair-breeding passerine, there is an expectation that costs of raising chicks will increase linearly with clutch size. Each chick requires X feeding visits to reach the required fledge weight. But this is not the case for species which lay precocious chicks which are relatively independent and able to feed themselves straight after hatching - so again the relationship of care and survival is unlikely to be detectable by looking at the effect of clutch size but again, it doesn't mean there isn't a trade-off between breeding and survival.*

Precocial birds still provide a level of parental care, such as protection from predators. Though we agree that the level of parental care in provisioning food (and in some cases in all parental care given) is lower in precocial than altricial birds, this would only make our reported effect size for manipulated birds to be an underestimate. Again, we would like to draw the reviewer's attention to the fact we did detect a trade-off in manipulated birds and we do not suggest that trade-offs do not exist. The argument the reviewer suggests here does not hold for unmanipulated birds, as we found that birds that naturally lay larger clutch sizes have higher survival.

> *• The costs of raising a brood to adulthood for your standard pair-breeding passerine is bound to be extreme, simply by dint of the energy expenditure required. In fact, it was shown that the basal metabolic rate of breeding passerines was at the very edge of what is physiologically possible, the human equivalent being cycling the Tour de France (Nagy et al. 1990). If birds are at the very edge of what is physiologically possible, is it likely that clutch size is under weak selection?*

If birds are at the very edge of what is physiologically possible, then indeed it would necessarily follow that if they increase the resource allocated in one area then expenditure in another area must be reduced. In many studies however, the overall brood mass is increased when chicks are added and cared for in an experimental setting, suggesting that birds are not operating at their limit all the time. Our simulations show that if individuals increase their clutch size, the survival cost of reproduction counterbalances the fitness gained by increasing clutch size and so there is no overall fitness gain to producing more offspring. Therefore, selection on clutch size is constrained to the within-species level. We do not say in our manuscript that clutch size is under weak selection – we only ask why variation in clutch size is maintained if selection always favours high-producing birds.

> *• Variation in clutch size is presented by the authors as inconsistent with the assumption that birds are under selection to lay the Lack clutch. Of course, this is absurd and makes me think that I have misunderstood the authors' intended point here. At any rate, the paper would benefit from more clarity about how variable clutch size has to be before it becomes a problem for optimality in the authors' view (lines 84-85; line 246). See Perrins (1965) for an exquisite example of how beautifully great tits optimise clutch size on average, despite laying between 5-12 eggs.*

We woud like to thank the reviewer for highlighting that our manuscript may be misleading in places, however, we are unsure which part of our conclusions the author is referring to here.The question we pose is "why all birds don't lay at the population optimum?", and is central to the decades-long field of life-history theory. Why is variation maintained at such a level? As the reviewer outlines it ranges massively with some birds laying half of what other birds lay.