Journal of Negative Results
— Ecology & Evolutionary Biology —
vol. 1: 1–5 • Helsinki, November 2004
ISSN 1459-4625
http://www.jnr-eeb.org/

Editorial: The Journal of Negative Results in Ecology and Evolutionary Biology

D. Johan Kotze[1*], Christine A. Johnson[2], Robert B. O'Hara[3], Kari Vepsäläinen[1] & Mike S. Fowler[1]

[1)] *Department of Biological and Environmental Sciences, P.O. Box 65, FI-00014 University of Helsinki, Finland. * author for correspondence: johan.kotze@helsinki.fi*
[2)] *The Ohio State University, Evolution, Ecology & Organismal Biology, 314 Aronoff Laboratory, 318 West 12th Street, Columbus, Ohio 43210, USA*
[3)] *Department of Mathematics and Statistics, P.O. Box 68, FI-00014 University of Helsinki, Finland*

Over 15 years ago, Bruce Charlton (1987) suggested in his article 'Think Negative' that many disciplines would benefit if negative results were given public airing. He argued that science needs reports of negative results for the simple reason that similar investigations, which are often costly and time-consuming, are frequently duplicated and produce the same negative result. Since then, other scientists have expressed similar sentiments (e.g., Knight 2003). Across the range of biological fields, table talk among ecologists and evolutionary biologists at all levels of academia at some point produces a sigh followed by '*If I had only known, I would have done things differently*'. Recently, a number of the sciences have realised the gravity of the lack of published negative results, and have attempted to fill this publication void by producing journals that report negative results (e.g., the Journal of Negative Results in Biomedicine, the Journal of Negative Results in Speech and Audio Sciences, the Forum for Negative Results [computer science], the Journal of Negative Observations in Genetic Oncology, and the Journal of Articles in Support of the Null Hypothesis). We agree with these researchers that the awareness of negative results plays an important role in the advancement of science, and are realising the objective of making negative results in ecology and evolutionary biology more widely available. This is being done in part through the *Open Access System* of the *Public Knowledge Project* (http://www.pkp.ubc.ca/ojs/), an organisation whose primary goal is to make information available to all by providing journal software at no cost.

What is a negative result?

Defining a negative result may not be as intuitively easy as it first appears. Indeed, there are several types of negative results. The clearest type is when data are subjected to formal hypothesis tests. If the null hypothesis cannot be rejected at a pre-determined significance level (usually set at $p = 0.05$), the experimental treatment is considered to have no detectable effect and the result is considered to be negative.

There are other equally important ways negative results can arise. Analyses can give statistically significant results that are contrary to expectations. Such potentially awkward situations originate when there is an *a priori* directional expectation inherent in the hypothesis. Contradicting current scientific consensus, negative results often face an increased probability of rejection because they can appear to shatter the structure of an often large body of research.

A recent example comes from the behavioural ecology of mating in waterstriders (Vepsäläinen & Spence 2000, *see* below for further discussion of this and other cases).

Replication of previous studies also produces negative results. Replicate studies in time or space, although considered important, often reveal nothing new and are sometimes seen as producing negative results. They are, therefore, less likely to be considered publishable, and consequently less likely to be carried out. A recent return to the problems of replication in ecological studies highlights various conceptual differences, including terms specific to replication (Oksanen 2001, Hurlbert 2004, Oksanen 2004). Most biological systems are difficult to replicate under natural conditions, and large studies in particular suffer from limitations caused by a variety of factors (unfortunately including funding limitations). Papers reporting these studies are often rejected in review on the basis that the conclusions are not reliable owing to small sample sizes. Should we ignore biological systems that potentially contribute to a better understanding of the overall picture? We believe not. In fields such as taxonomy, information is often gathered and utilised from a single data point — descriptions of species are based on information gathered from a single specimen (e.g., *Gnamptogenys stellae* in Lattke 1995). Opinions about evolutionary history have been transformed on the basis of measurements of a single mandibular articulation (Engel & Grimaldi 2004). Hopefully such cases will be rare, with the documentation of new knowledge usually being based on a statistically representative sample. Whilst a conclusion based on a single ecological or biological event may be considered suspect, a collection of small data sets can lend itself to a larger meta-analytical study, which can reveal meaningful biological trends.

In summary, results that either fail to reject a null hypothesis or do not accord with current consensus are often not published, which may lead to a biased representation of natural processes (Knight 2003). Furthermore, problems that arise from a lack of replication only serve to compound our misery.

The necessity and difficulties of publishing negative results

There is no shortage of researchers who have spent years searching for the correct technique that allows them to complete their study and answer their question. Nor is there a shortage of researchers who have spent years searching for information that would allow them to generate a strong hypothesis (*see* Martin *et al.* in review). In a scientific community that is competing aggressively for funds, the investment of valuable resources and time into examinations that have already been carried out and deemed unworkable, can be all too costly (Charlton 1987, Knight 2003). In recent years, the pressures to publish only positive results (and hefty, often unrealistic publication lists) have led to scientific integrity being compromised ('The whole truth' [Editorial]: *New Scientist* 182 (2004): p. 3).

A study of Swedish and Finnish Ph.D. theses by Koricheva (2003) revealed that non-significant results are not necessarily less likely to be published than are significant results. However, non-significant results and results that contradict current expectations (Leimu & Koricheva 2004) are more likely to be published in journals with lower impact factors than studies with significant and expected results. Koricheva (2003) and Leimu and Koricheva (2004) concluded that this is not necessarily because journals reject non-significant studies at a higher rate than they do significant studies, but because authors send their non-significant and contradictory results to lower impact factor journals. Presumably, this is because authors assume that their "negative" studies will not be accepted by higher impact factor journals. The ramifications of this are that negative results may be less accessible to the wider scientific community, and therefore contribute to biases in meta-analytical studies. One important assumption behind Koricheva's (2003) findings is that "... all or at least most of the research conducted during the Ph.D. studies is included in the thesis": it may be that many negative results are not even written up in any form.

When meta-analytic studies are carried out, the publication bias that results from the under-reporting of non-significant results will lead to an over-estimation of the true effect size (Palmer 2000). Jennions *et al*. (2004) showed that in up to 21% of meta-analytical studies they examined, the null hypothesis was falsely rejected. A related problem is determining the extent to which hypotheses generated by scientists studying one system can be extrapolated to other systems (Palmer 2000, Vepsäläinen & Spence 2000); without the replication of studies in different systems, the extrapolations cannot be tested.

Perhaps as important as knowing which type of negative result has been encountered, is being able to explain why such results occur. Even with standard methodology unexpected results are possible. In a trivial case, this may simply be due to mistakes generated during the learning process (such as a new student working for the first time in a molecular biology laboratory), or to methodological errors, which can occur in any field of ecology or evolutionary biology. For instance, in a theoretical study by Bascompte and Solé (1994), the effects of dispersal on population dynamics in a single-species model were tested. They found that introducing a spatial dimension to the model destabilised the population dynamics, contrary to previous studies that concluded dispersal leads to a reduction in the tendency for populations to fluctuate (McCallum 1992, González-Andújar & Perry 1993a, 1993b, Stone 1993). Shortly thereafter, Hassell *et al*. (1995) pointed out that Bascompte and Solé's (1994) findings were due to a biologically impossible assumption contained within the dispersal function of the model, which failed to segregate survival and dispersal processes. Of course, this should more correctly be called a spurious result when interpreted biologically and compared with the correct results. The methods employed by the model led to a counter-intuitive (negative) result that was later shown to be entirely due to incorrect assumptions.

Alternatively, there may be genuine differences between results from the previous and current studies, not accounted for by differences (or errors) in methodology. Under these circumstances, it is important to be able to show why a certain method does not yield the previously published results, along with a clear explanation of why the results differ (Charlton, this issue). One example occurred in a study of waterstrider mating behaviour, which conflicted with earlier work on these behaviours. The established view of mating of many waterstrider species was that the copulatory phase was preceded by female struggle and followed by male guarding. This was ultimately broken by female struggling. The new work demonstrated that behavioural tactics are far more diverse and context-dependent than the earlier views suggested — notably, female-biased operational sex ratios experienced by males before mating experiments resulted in short matings and practically no post-copulatory guarding. The communication describing this work was rejected three times during the review process by the same referee, a proponent of an earlier, more generalised mating strategy (*see* footnote in Vepsäläinen & Spence 2000: pp. 234–235). This example highlights the problems that such results face when subjected to the review process. While it is hoped that individuals chosen to review work are at least objective, if not always independent of the subject matter, this is clearly not always the case. This has been true historically within biology, with current scientific fashion often dictating what is viewed as a negative result. This in turn will have consequences for the likelihood of important work being published. The long-running debate on matters such as "levels of selection" highlights how changing fashions often play as important a role in the publication of work as the quality of the work itself. In short, negative results arise because both nature and scientists change over time, because novel techniques are created and because individual perceptions and biases vary, even with attempts at standardisation.

Negative results in ecology: an example

One particular example that summarises many of the points made above stems from the ongoing discussion of the mechanism behind cycles in small mammal population sizes (Oli 2003a, 2003b). In short, there are two competing arguments that attempt to explain why cycles arise in vole population sizes: intrinsic (including maternal effects) and extrinsic (e.g., environmental and trophic effects, in particular the specialist predator hypothesis). Both camps provide evidence to support their chosen hypothesis. However, Graham and Lambin (2002) demonstrated that a reduction in numbers of specialist predators had no impact on the characteristic cycles in a British vole population. Could this be perceived as a negative result? Here, the control treatment was represented by populations under natural predation conditions, and were already known to cycle, whereas populations in the experimental treatment (areas with a reduction in numbers of specialist predators) had not been expected to cycle if the specialist predator hypothesis had been true. However, Graham and Lambin reported continued cycling of the populations following predator removal. Thus, there was no effect of the experimental treatment. Although there are undoubtedly limitations in large-scale studies such as this, problems arise when assessing the objectivity of criticisms of this work, as they often come from proponents of opposing theories (Korpimäki *et al*. 2003, Oli 2003a, 2003b).

This case also highlights one of the difficulties of assessing and reporting negative results mentioned above, namely replication. While laboratory-based studies may be controlled relatively easily, replicates in natural conditions will be buffeted by many uncontrollable factors, across a variety of scales. This should be countered by increasing the amount of replication in field studies, something which rarely is the case. The fact that the different populations in the example above are in different geographical regions (Britain, a large island with relatively sparse and fragmented forest remnants and a relatively high human population density as compared with Finland, a continental land area with a considerably lower human population density) means that results from one population may not be applicable across all populations of the same species. This is a common situation with ecological studies, where the units commonly under scrutiny (genes, individuals, populations, species, communities, ecosystems) do not always offer desirable characteristics for rigorous control or replication.

The establishment of JNR-EEB

In December of 2003 we established the *Journal of Negative Results — Ecology and Evolutionary Biology*, an online medium (http://www.jnr-eeb.org) for the publication of scientific work in ecology and evolutionary biology that may otherwise remain unknown. Our goal is not to provide a repository for bad science but to publish scientific studies that have been rigorously carried out, yet produce negative results. These will include, for example, studies in which no significant effect was obtained, studies in which a significant effect was obtained but in an unpredicted direction, studies that are replicates of previous studies (whether they yield expected or unexpected results), and studies in which the data are biologically interesting but lack statistical power. We will also encourage scientists to submit non-significant results based on small data sets, which can be published as short communications. Although these studies may be minor singly, their significance will accumulate as more studies are published and made available for meta-analyses from which regularities can be extracted.

JNR's focus will be ecology and evolutionary biology, and include areas such as behavioural, chemical, and community biology, conservation, evolutionary, population, or theoretical ecology and population genetics.

Will we succeed?

The *JNR-EEB* has enormous potential for success on a number of levels. We will succeed if scientists contribute their negative results to this journal, or if traditional journals begin publishing more negative results. Scientists have a great deal to benefit from the *JNR-EEB*. Firstly, they have an outlet to demonstrate their efforts and progress. Secondly, researchers will have information available about previous work that has produced negative results, on which they can build through either the modification of contemporary methodologies or conducting meta-analytical studies. *'Awareness of the problem [of lack of negative* *results] is growing'* (Knight 2003), and the *JNR-EEB* is a positive step in changing the future of published research in our field.

References

Bascompte, J. & Solé, R. 1994: Spatially induced bifurcations in single-species population dynamics. *J. Anim. Ecol*. 63: 256–264.

Charlton, B. 1987: Think negative. *New Scientist* 116: 72–72.

Charlton, B. G. 2004: Why a journal of negative results? *JNR-EEB* 1: 6–7.

Engel, M. S. & Grimaldi, D. A. 2004: New light shed on the oldest insect. *Nature* 427: 627–630.

González-Andújar, J. L. & Perry, J. N. 1993a: Chaos, metapopulations and dispersal. *Ecol. Model*. 65: 255–263.

González-Andújar, J. L. & Perry, J. N. 1993b: The effect of dispersal between chaotic and non-chaotic populations within a metapopulation. *Oikos* 66: 555–557.

Graham, I. M. & Lambin, X. 2002: The impact of weasel predation on cyclic field-vole survival: the specialist predator hypothesis contradicted. *J. Anim. Ecol*. 71: 946–956.

Hassell, M., Miramontes, O., Rohani, P. & May, R. M. 1995: Appropriate formulations for dispersal in spatially structured models; Reply to Bascompte and Solé. *J. Anim. Ecol*. 64: 662–664.

Hurlbert, S. H. 2004: On misinterpretations of pseudoreplication and related matters: a reply to Oksanen. *Oikos* 104: 591–597.

Jennions, M. D., Møller, A. P. & Hunt, J. 2004: Meta-analysis can "fail": reply to Kotiaho and Tomkins. *Oikos* 104: 191–193.

Knight, J. 2003: Negative results: Null and void. *Nature* 422: 554–555.

Koricheva, J. 2003: Non-significant results in ecology: a burden or a blessing in disguise? *Oikos* 102: 397–401.

Korpimäki, E., Klemola, T., Norrdahl, K., Oksanen, L., Oksanen, T., Banks, P. B., Batzli, G. O. & Henttonen, H. 2003: Vole cycles and predation. *TREE* 18: 494–495.

Lattke, J. E. 1995: Revision of the ant genus *Gnatmptogenys* in the New World (Hymenoptera: Formicidae). *Journal of Hymenoptera Research* 4: 137–193.

Leimu, R. & Koricheva, J. 2004. Cumulative meta-analysis: a new tool for detection of temporal trends and publication bias in ecology. *Proc. R. Soc. Lond. B* 271: 1961–1966.

Martin, S. J., Chaline, N., Jones, G. R. & Ratnieks, F. L. W. 2004: Searching for the egg-marking signal in honeybees. *JNR-EEB*, in review.

McCallum, H. I. 1992: Effects of Immigration on Chaotic Population Dynamics. *J. Theor. Biol*. 154: 277–284.

Oksanen, L. 2001: Logic of experiments in ecology: is pseudoreplication a pseudoissue? *Oikos* 94: 27–38.

Oksanen, L. 2004: The devil lies in details: reply to Stuart Hurlbert. *Oikos* 104: 598–605.

Oli, M. K. 2003a: Population cycles of small rodents are caused by specialist predators: or are they? *TREE* 18: 105–107.

Oli, M. K. 2003b: Response to Korpimäki et al.: vole cycles and predation. *TREE* 18: 495–496.

Palmer, A. R. 2000: Quasireplication and the contract of error: Lessons from sex ratios, heritabilities and fluctuating asymmetry. *Annu. Rev. Ecol. Syst*. 31: 441–480.

Stone, L. 1993: Period-doubling reversals and chaos in simple ecological models. *Nature* 365: 617–620.

Vepsäläinen, K. & Spence, J. R. 2000: Generalization in ecology and evolutionary biology: from hypothesis to paradigm. *Biol. Philos*. 15: 211–238.