

The role of theory in mitigating the ‘reproducibility crisis’

LUIS MARONE^{1,2}

¹ ECODES (Desert Community Ecology Research Team), IADIZA-CONICET, Mendoza, Argentina. ² Facultad de Ciencias Exactas y Naturales, UNCuyo, Mendoza, Argentina.

ABSTRACT. The lack of reproducibility of scientific results is jeopardizing the trust in science. An effort to inform the dynamic, but non-arbitrary, nature of scientific evidence is required along with strengthening the reliability of published results. Concerns and actions aimed at testing and increasing the reproducibility of scientific conclusions are usually directed at increasing data validation through careful management of the data-curation process, paying attention to the empirical aspects of hypotheses testing. But important as they are, these aspects may be vulnerable to the perils of radical empiricism and therefore should be combined with more theoretical and conceptual tools. Pattern consistency and an understanding of the physiological, behavioral, or ecological mechanisms that cause patterns make scientific assertions more robust. Testing a priori hypotheses (e.g., about functional biological traits) with fresh (and redundant) evidence offers theoretical as well as empirical support for ecological research and may help strengthen the reliability of published results.

[Keywords: a priori hypotheses, biological traits, computational reproducibility, empirical reproducibility, mechanisms, theory]

RESUMEN. La función de la teoría para mitigar la ‘crisis de reproducibilidad’. La falta de reproducibilidad de los resultados científicos desafía la confianza en ellos. Se requiere un esfuerzo tendiente a difundir la naturaleza dinámica, aunque no arbitraria, de la evidencia científica, además de instrumentar medidas para aumentar la confianza en los resultados publicados. La reflexión y las acciones dirigidas a evaluar e incrementar la reproducibilidad de las conclusiones científicas suelen estar dirigidas a aumentar la validez de los datos a través de un manejo cuidadoso de su proceso de curación, prestando atención a aspectos empíricos de la puesta a prueba de hipótesis. Sin embargo, estos importantes aspectos no son suficientes porque solo atienden elementos empíricos y técnicos de la prueba (i.e., diseño experimental, datos, análisis estadístico), por lo que deberían complementarse con herramientas teóricas y conceptuales que incrementen la confianza en las conclusiones. Tanto la robustez de los patrones ecológicos como la comprensión de los mecanismos fisiológicos, comportamentales y ecológicos que los causan contribuyen a dar solidez a las conclusiones. La puesta a prueba de hipótesis a priori (e.g., sobre rasgos biológicos funcionales) con evidencia novedosa (y redundante) ofrece a la investigación en ecología el apoyo simultáneo de datos y un entramado teórico actualizado y pertinente, contribuyendo a incrementar la confiabilidad en los resultados que se publican.

[Palabras clave: hipótesis a priori, mecanismos, rasgos biológicos, reproducibilidad computacional, reproducibilidad empírica, teoría]

The ability of science to understand the world is threatened by misinformation that jeopardizes trust in science (Amara 2022). In a ‘post-truth era’, there is much work to be done to explain the method of science to the public and policy makers, teaching the dynamic nature of research in which both failure to predict results and revision of earlier findings are not infrequent (González del Solar and Marone 2001; Roper 2022). If the provisional, but nonarbitrary, nature of scientific knowledge is not taught, shifts in understanding may impel people to believe that science cannot be trusted. However, this educational effort must be accompanied by a commitment from the research community to raise confidence in scientific

results by improving the quality of data and interpretations (Ioannidis 2014).

There are therefore serious concerns about the reproducibility and generalizability of scientific results, encouraging a cultural shift in transparency and data quality (Leonelli 2016; Parker et al. 2016; Berg 2018; McCord et al. 2021) in many disciplines (Marone et al. 2000; Baker and Penny 2016; Forstmeier et al. 2017; Ihle et al. 2017; Hutson 2018; Bishop 2019; Desjardins et al. 2021; Kaiser 2021; Sikorski 2022). An interesting example for ecology is the need to carefully assess the reliability and reproducibility of the evidence that can justify optimistic and pessimistic positions regarding the environmental crisis (Grau 2022, 2023). Hereinafter, I will review the main

causes of irreproducibility of research results in empirical sciences along with the typical actions the scientific community is carrying out to correct this rising problem (Oza 2023). Afterwards, I will point out and assess in detail the not sufficiently emphasized role that theory plays—in particular, evolutionary theory and natural-history hypotheses—in mitigating the reproducibility crisis in ecology and other disciplines.

To avoid confusion about the meaning of terms, empirical reproducibility (sometimes referred to as replicability) (Desjardins et al. 2021) considers the ability to obtain the same or a very similar result during the test of the same scientific hypothesis with the same or with a different, but equally pertinent, methodology in the lab or the field (Marone et al. 2019), whereas computational reproducibility focuses on the ability to produce equivalent analytical outcomes from the same data set using the same code and software as the original study (Peng 2011; Powers and Hampton 2019). Redundant results reached using different methodologies based on different assumptions (e.g., experiments, planned observations or simulations) are valuable because the inferences drawn from them often have higher external validity (i.e., they are robust and, in principle, can be generalized across a wider domain) (Marone et al. 2000; Munafó and Smith 2018; Desjardins et al. 2021). Reproducibility constitutes a requisite for reliable and predictive scientific knowledge, and reproducibility failures are among the most important concerns of the scientific community today (Baker and Penny 2016; Kaiser 2021).

First of all, irreproducibility in many disciplines may result from the variability of nature and historical contingency (Bissell 2013; Fidler et al. 2017; Pérez-Velázquez 2019; Desjardins et al. 2021). For example, ecological-evolutionary phenomena may be context-dependent (e.g., in space, time, species), and gradual changes in the ecosystem—as well as legacies of past events—may make such systems bear little resemblance to earlier states. They are ontological determinants of the lack of reproducibility, irregularities in the natural world that often disappoint the scientist engaged in field studies aiming at predicting natural phenomena. Aiming at solving this problem, Powers and Hampton (2019) suggested that ecologists can still achieve computational reproducibility even though field observations may never be

completely or perfectly reproduced. This claim is, however, arguable. Firstly, *in silico* ‘experiments’ are only plausible images of a real situation, not the actual real situation (Gunawardena 2014). Some unknown drivers will be absent in simulations because unrecognized parameters obviously cannot be consciously included in a model, yet they are present, exerting their effects, in nature. The existence of hidden assumptions makes empirical studies more realistic. Additionally, *in silico* irreproducibility also seems to be unavoidable if artificial intelligence algorithms are used to derive expertise from experience. In machine learning, for example, the way data are trained influences the performance of any algorithm that ‘learns’ by trial and error, because performance is sensitive not only to the exact code used, but also to the random numbers generated to initiate training and to settings that are not core to the algorithm, but that affect how quickly it ‘learns’ (‘hyperparameters’) (Hutson 2018). Despite the undeniable value of simulations in research and the stimulating advances in computer science, researchers dealing with the high variability of ecological phenomena in the real world will continue to need some conventional science methodology (Werner 1998; Marone et al. 2019).

Another source of irreproducibility is associated with the method of science, the variety of its associated techniques and various psychological and sociological aspects of scientific practice: they are mostly gnoseological determinants of irreproducibility. Baker and Penny (2016) reported the results of a survey of 1576 researchers who were asked about which factors they thought contributed most to irreproducible research. Psychosociological aspects that often interact with more methodological ones, such as selective reporting (70%), pressure to publish (>60%), insufficient mentoring (50%), raw data not available in the original lab (>40%), code unavailable (>40%), fraud (40%), and insufficient peer review (almost 40%) were highly rated, together with other purely methodological weaknesses, such as poor analysis (almost 60%), insufficient replication in the original lab (50%) and poor experimental design (>40%).

According to these figures, the combination of selective reporting and an excessive pressure to publish novel and ‘exciting’ results might be among the most important sources of the lack of confirmation of some published

studies. Selective reporting takes several forms depending on the type of study and consists of trying many analyses, but reporting only those that are 'statistically significant'. For example, repeating an experiment 18 times always reaching negative results and finally publishing the results of the 19th trial that rejects the null hypothesis (Bishop 2019) or reporting the significant correlation between one empirical indicator of an effect on a dependent variable after having carried out a multiple test that involved fifteen other unreported empirical indicators of the same effect, which did not correlate with the dependent variable (Forstmeier et al. 2017). Selective reporting is well known (Jennions and Moller 2002; Lehrer 2010; Frasser et al. 2018; Bishop 2019; Kaiser 2021), but its impact on the validity of scientific inferences is hardly taught to young scientists and may still be insufficiently debated (but see Ihle et al. 2017).

On the contrary, there is substantial concern about the need to pay attention to data curation and the quality of the techniques used to draw inferences from them (Editors 2014; Ioannidis 2014; McNutt 2014; Fidler et al. 2017; Ihle et al. 2017; Berg 2018). More precisely, there are many initiatives prescribing the improvement of study designs, the full description of experiments, the recording and sharing of raw data in repositories and more precise reporting of statistical analyses and their results. These initiatives are aimed at wisely managing the data-curation process to increase data validation.

A substantial part of hypothesis-driven research consists of testing hypotheses with fresh evidence. The data take the form of a number of observational statements that should reflect the natural facts (Figure 1, right side) (Hempel 1966; Bunge 1998). Therefore, well-validated data are those that correspond reliably to the facts, which depends on the careful consideration of several auxiliary hypotheses or assumptions that surround data collection (Figure 1, right side). Are the techniques used for data gathering sound? Are the theories underlying these techniques correct? What are the main sources of potential bias in the data or what are the main assumptions that, if proven false, would have biased the data? Also, on the right hand side of Figure 1, the confidence of results achieved by statistical inference matters (e.g., the statistical power of the analysis, whether the critical assumptions of statistical tests are met, the

number and types of replicates taken, the way variables are controlled and randomized, the criteria for excluding any data in the analysis, the level of significance established, the size of the effect detected). Studies that do not follow best practices on data curation and statistical analysis could reasonably fail in reproducibility and, conversely, there is evidence that 'methodologically optimal experiments' may be highly reproducible (Sikorski 2022). Then, the attention paid to data curation and sound analyses to increase data validation is widely justified (Ioannidis 2014).

But there is another way to mitigate the crisis. It is related to the left side of the scientific method (Figure 1) and is less considered in the irreproducibility literature than the psychosociological and empirical factors described above. Testing more or less formal theoretical constructs with the hypothetico-deductive method (e.g., Popper 1959; Bunge 1998; Marone and Galetto 2011) requires deducing a hypothesis (i.e., a description of how the reality under analysis would be if the construct is 'true') and predictions (i.e., the same description, but stated in a 'directly testable' way, under field or laboratory conditions) (Farji-Brener 2020). The premises of the deductive process will be the starting theory or theories along with other assumptions that may be called 'initial conditions'. If at least the most critical assumptions are considered and the theories are 'true' or approximately true, the hypothesis deduced is said to be plausible, not a certainty. While in formal logic a deduction from 'true' premises warrants the 'truth' of the conclusion, in the reasoning of the empirical sciences some factual uncertainty always remains in the premises and then the deduction can only make the conclusion plausible (Bunge 1998). At the same time that the hypothesis is deduced by proceeding through that logical sequence, the hypothesis is made plausible a priori (before data gathering) by going backward in the logical sequence to the support offered by the updated scientific knowledge (e.g., theories, initial conditions) included in the premises (Bunge 2012).

This situation is familiar to modelers, who consider their models (hypotheses) as mere descriptions of their assumptions about reality that could only be as plausible as the assumptions (Gunawardena 2014). According to this perspective, a model has the virtue that, as long as the math has been done correctly, if

the assumptions correspond to the real system, the predictions of the model will reasonably correspond to reality. In applied ecology, trustworthy assumptions make a model theoretically sound, and its predictions may occasionally be considered reliable enough without further assessment. For example, consider a minimum viable population value calculated by a wildlife management model. The manager will apply the value aiming at conserving an endangered population without making any further experiments. In the best case, the truth value of the prediction will be evaluated during a monitoring program following application. On the other hand, in basic ecology the model would be considered plausible a priori, which is a good enough reason to take the prediction seriously and, eventually, test it with fresh evidence (Kerr 1998; Bunge 2012; Marone et al. 2019).

By the way, the lack of theoretical foundation is typical of the assertions made by the pseudosciences or by post-truth activists. Debunking these unsubstantiated claims is more a matter of lack of theoretical support than of specific data (i.e., the data might eventually match some expectations of such allegations just by chance) (McNutt 2014) in the worthy task of avoiding the slipperiness of empiricism (Bunge 1998; Kerr 1998; Lehrer 2010) or the 'one shot game' in hypotheses testing (Marone et al. 2019).

This reasoning is the most important argument against hypothesizing after the results are known (HARKing) (Kerr 1998), which is typical of data-driven exploratory research (Ihle et al. 2017; Mitchell et al. 2018). HARKing often leads to generate a hypothesis that researchers would never have deduced from theory (Forstmeier et al. 2017), and whose plausibility was therefore never established (Bunge 1998). Together

with this hypothesis, there are, in principle, numerous other alternative hypotheses, which are usually ignored, but that could account for the same results (i.e., underdetermination, see Leonelli 2016). Further, the testing of such ad hoc and post hoc hypotheses is intrinsically circular because the same data are used to both generate and test them, and there is no way to reject them (Bunge 1998; Kerr 1998). Publishing the 'confirmation' of this kind of hypotheses is, consciously or not, a sort of 'selective reporting'. In Kerr's (1998) words: "... whereas a genuinely a priori hypothesis has some theoretical or empirical foundation that is independent of the current result ... an explicitly post hoc hypothesis implicitly acknowledges its dependence upon the result in hand as the cornerstone (or perhaps, the entirety) of its foundation...". All these arguments justify the use of a priori hypotheses as a plausible tool for establishing more reproducible research results and, consequently, more reliable inferences.

By way of example, the hypothesis that seed-eating birds maximize energy intake rates while feeding, and its prediction that seed-eating birds will consume large rather than small seeds in controlled experiments (Cueto et al. 2001), could be deduced from two major theoretical frameworks (i.e., the theory of adaptation by natural selection, and the associated optimal foraging theory), together with several critical assumptions (e.g., that larger seeds provide more energy reward per unit of time). Although theories are always open to revision, the former are usually assumed to be true due to their heuristic value and reasonable evidence supporting them. The critical assumptions, in turn, are more reliably assumed if they are corroborated in light of fresh, independent data. Thus, when such a priori hypothesis and their predictions are corroborated for some bird species (e.g.,

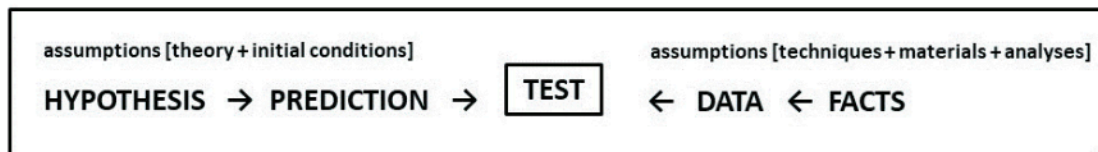


Figure 1. The hypothetico-deductive method outlined to test an a priori hypothesis. On the right side, the assumptions regarding data gathering and analyses (e.g., about the techniques and materials used in the field or the lab as well as the statistical analyses used to draw inferences) are emphasized. On the left side of the scheme, the assumptions highlighted are those corresponding to the starting theories used together with the initial conditions necessary to deduce the hypotheses and predictions to be tested.

Figura 1. Resumen del método hipotético deductivo empleado para poner a prueba una hipótesis concebida a priori. A la derecha de la prueba, los supuestos vinculados con la obtención y el análisis de datos (e.g., los que subyacen las técnicas y materiales empleados en el campo o el laboratorio y los asociados al análisis estadístico usado para hacer las inferencias). A la izquierda del esquema, las hipótesis y teorías de partida, que se suponen 'verdaderas', junto con algunas condiciones iniciales pertinentes para deducir las hipótesis de trabajo y sus predicciones.

Saltatricula multicolor, *Zonotrichia capensis*) (Marone et al. 2022), confidence in them comes from general evidence (i.e., adaptation, rational decision, information on seed energy reward) as well as more specific evidence (i.e., the result of laboratory trials carried out under the guidance of the hypothesis). Mechanistic a priori hypotheses (Bunge 1998; Marone and Bunge 1998) link theory (e.g., knowledge of natural history within an adaptive context) with ecological patterns (e.g., seed preferences or diet composition), so that the inference is supported simultaneously by a plausible theoretical framework and fresh specific evidence. The research program advised by Earl Werner (1998) is an example of this approach that emphasizes the integration of theoretical + experimental work to generate mechanistic hypotheses with more descriptive work to test predictions in the field. It is the iteration of theory, experiments and a proper field pattern that is so valuable (Werner 1998; Marone et al. 2019). A priori hypotheses are effectively at risk when faced with realistic field-ecological results. For example, and with respect to the studies of granivorous birds in central Argentina, lab experiments suggest that the abundance of the Yellow Cardinal (*Gubernatrix cristata*) should decrease in fields used for livestock due to the reduction of critical food resources there (i.e., medium and large grass seeds) (Marone and Camín 2022). If this pattern is not found after a reasonable research effort, the hypothesis must be discarded and other plausible reasons to account for the habitat selection of the cardinal should be proposed and tested.

The research approach based on the use of biological traits as a priori hypotheses on the natural history of the organisms to explain and predict ecological patterns successfully

foresaw, for example, many of the responses of herbaceous plants to grazing (Díaz et al. 2001), of interacting species to climate change (Schleuning et al. 2020), of birds invading new habitats (Sol et al. 2012), of bird composition in urbanization gradients (Crocí et al. 2008; Camín et al. 2022), of bird species to habitats subjected to bush thickening (Seymour and Dean 2010) and cattle grazing (Martin and Possingham 2005; Sagario et al. 2020). Besides the theoretical support (Figure 1, left side), the results of several empirical tests guided by the hypotheses are ideally needed (Figure 1, right side) to reduce the probability of false positives obtained by chance and of committing a Type I statistical error (i.e., 'empirical robustness' [Stegenga 2009; Nichols et al. 2019]). Obviously, sometimes the results do not meet predictions and the hypothesis should then be rejected. In any case, the combined feedback of theory with fresh and redundant evidence provides an instance of valuable reflection on the validity of results before they are published.

ACKNOWLEDGMENTS. Javier Lopez de Casenave, Rafael González del Solar, and Mario Bunge provided insightful discussions on these topics over the years. Fabián Jaksic, Martín Aguiar, Enrique Miranda, Marcelo Cabido, and Rosemary Scofield made useful comments on earlier drafts of the manuscript. I am indebted to all of them. The Fondo Nacional de Promoción Científica y Tecnológica (FONCYT-ANPCyT, most recently through PICT 2019 03217), and the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, most recently through PIP 2718), both from Argentina, helped me develop the empirical investigations that inspired this essay. Contribution number 120 of the Desert Community Ecology Research Team (Ecodes).

REFERENCES

- Amara, S. G. 2022. Empower with science. *Science* 375:699. <https://doi.org/10.1126/science.abo5963>.
- Baker, M., and D. Penny. 2016. Is there a reproducibility crisis? *Nature* 533:452-454. <https://doi.org/10.1038/533452a>.
- Berg, J. 2018. Progress on reproducibility. *Science* 359:9. <https://doi.org/10.1126/science.aar8654>.
- Bishop, D. 2019. Rein in the four horsemen of irreproducibility. *Nature* 568:435. <https://doi.org/10.1038/d41586-019-01307-2>.
- Bissell, M. 2013. The risks of the replication drive. *Nature* 503:333-334. <https://doi.org/10.1038/503333a>.
- Bunge, M. 1998. *Philosophy of science. Revised Edition.* Transaction, New York, New York, USA. <https://doi.org/10.4324/9781315126371>.
- Bunge, M. 2012. The correspondence theory of truth. *Semiotica* 188:65-75. <https://doi.org/10.1515/sem-2012-0004>.
- Camín, S. R., I. Figgini, and L. Marone. 2022. ¿Predicen los rasgos biológicos de las aves su ocupación del gradiente de urbanización en el área metropolitana de Mendoza? *Ecología Austral* 32:1078-1088. <https://doi.org/10.25260/EA.22.32.3.0.1862>.
- Crocí, S., E. Butet, and P. Clergeau. 2008. Does urbanization filter birds on the basis of their biological traits? *Condor* 110:223-240. <https://doi.org/10.1525/cond.2008.8409>.

- Cueto, V. R., L. Marone, and J. Lopez de Casenave. 2001. Seed preferences by birds: effects of the design of feeding-preference experiments. *Journal of Avian Biology* 32:275-278. <https://doi.org/10.1111/j.0908-8857.2001.320311.x>.
- Desjardins, E., J. Kurtz, N. Kranke, A. Lindeza, and S. H. Richter. 2021. Beyond standardization: improving external validity and reproducibility in experimental evolution. *BioScience* 71:543-552. <https://doi.org/10.1093/biosci/biab008>.
- Díaz, S. M., I. Noy-Meir, and M. Cabido. 2001. Can grazing response of herbaceous plants be predicted from simple vegetative traits? *Journal of Applied Ecology* 38:497-508. <https://doi.org/10.1046/j.1365-2664.2001.00635.x>.
- Editors. 2014. Journals unite for reproducibility. *Nature* 515:7. <https://doi.org/10.1038/515007a>.
- Farji-Brener, A. G. 2020. ¿18 años no es nada? Una nueva revisión del uso correcto, parcial e incorrecto de los términos 'hipótesis' y 'predicciones' en ecología. *Ecología Austral* 30:393-400. <https://doi.org/10.25260/EA.20.30.3.0.1129>.
- Fidler, F., Y. E. Chee, B. C. Wintle, M. A. Burgman, M. A. McCarthy, and A. Gordon. 2017. Metaresearch for evaluating reproducibility in ecology. *BioScience* 67:282-289. <https://doi.org/10.1046/10.1093/biosci/biw159>.
- Forstmeier, W., E. J. Wagenmakers, and T. H. Parker. 2017. Detecting and avoiding likely false-positive findings – a practical guide. *Biological Review* 92:1941-1968. <https://doi.org/10.1046/10.1111/brv.12315>.
- Fraser, H., T. Parker, S. Nakagawa, A. Barnett, and F. Fidler. 2018. Questionable research practices in ecology and evolution. *PLOS ONE* 13:e02000303. <https://doi.org/10.1371/journal.pone.0200303>.
- González del Solar, R., and L. Marone. 2001. The “freezing” of Science: consequences of the dogmatic teaching of ecology. *BioScience* 51:589-592. [https://doi.org/10.1641/0006-3568\(2001\)051\[0683:TFOSCO\]2.0.CO;2](https://doi.org/10.1641/0006-3568(2001)051[0683:TFOSCO]2.0.CO;2).
- Grau, H. R. 2022. Contra el consenso: Hallazgos que amenazan fundamentos del eco-alarmismo. *Ecología Austral* 32:33-44. <https://doi.org/10.25260/EA.22.32.1.0.1782>.
- Grau, H. R. 2023. Más hallazgos contra el consenso eco-alarmista. *Ecología Austral* 33:479-488. <https://doi.org/10.25260/EA.23.33.2.0.2158>.
- Gunawardena, J. 2014. Beware the tail that wags the dog: informal and formal models in biology. *Molecular Biology of the Cell* 25:3441-3444. <https://doi.org/10.1091/mbc.E14-02-0717>.
- Hempel, C. G. 1966. *Philosophy of natural science*. Prentice-Hall Inc., New York, New York, USA.
- Hutson, M. 2018. Artificial intelligence faces reproducibility crisis. *Science* 359:725-726. <https://doi.org/10.1126/science.359.6377.725>.
- Ihle, M., I. S. Winney, A. Krystalli, and M. Croucher. 2017. Striving for transparent and credible research: practical guidelines for behavioral ecologists. *Behavioral Ecology* 28:348-354. <https://doi.org/10.1093/beheco/axx003>.
- Ioannidis, J. P. A. 2014. How to make more published research true. *PLOS Medicine* 11:1001747. <https://doi.org/10.1371/journal.pmed.1001747>.
- Jennions, M. D. and A. P. Moller. 2002. Publication bias in ecology and evolution: An empirical assessment using the ‘trim and fill’ method. *Biological Reviews of the Cambridge Philosophical Society* 77:211-222. <https://doi.org/10.1017/S1464793101005875>.
- Kaiser, J. 2021. Key cancer results failed to be reproduced. *Science* 374:1311. <https://doi.org/10.1126/science.acx9788>.
- Kerr, N. L. 1998. HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review* 3: 197-217. https://doi.org/10.1207/s15327957pspr0203_4.
- Lehrer, J. 2010. The truth wears off. Is there something wrong with the scientific method? *The New Yorker* 13:52-57 [December].
- Leonelli, S. 2016. *Data-centric biology. A philosophical study*. University of Chicago Press, Chicago, Illinois, USA. <https://doi.org/10.1080/14636778.2017.1389263>.
- Marone, L., and M. Bunge. 1998. La explicación en Ecología. *Boletín de la Asociación Argentina de Ecología* 7:35-37.
- Marone, L., and L. Galetto. 2011. El doble papel de las hipótesis en la investigación ecológica y su relación con el método hipotético deductivo. *Ecología Austral* 21:201-216.
- Marone, L., and S. R. Camín. 2022. Seed preferences suggest a high vulnerability of the Yellow Cardinal (*Gubernatrix cristata*) to habitat degradation in Argentina. *Emu* 3-4:208-215. <https://doi.org/10.1080/01584197.2022.2112696>.
- Marone, L., J. Lopez de Casenave, and V. R. Cueto. 2000. Granivory in the southern South American deserts: conceptual issues and current evidence. *BioScience* 50:123-132. [https://doi.org/10.1641/0006-3568\(2000\)050\[0123:GISSAD\]2.3.CO;2](https://doi.org/10.1641/0006-3568(2000)050[0123:GISSAD]2.3.CO;2).
- Marone, L., J. Lopez de Casenave, and R. González del Solar R. 2019. The synthetic thesis of truth helps mitigate the reproducibility crisis and is an inspiration for predictive ecology. *Journal of Humanities of Valparaíso* 14:366-373. <https://doi.org/10.22370/rhv2019iss14pp363-376>.
- Marone, L., V. R. Cueto, J. Lopez de Casenave, A. Zarco, and S. R. Camín. 2022. Plausible causes of seed preferences and diet composition in seed-eating passerines. *Journal of Avian Biology* 2021:e02875. <https://doi.org/10.1111/jav.02875>.
- Martin, T. G., and H. P. Possingham. 2005. Predicting the impact of grazing on birds using foraging height data. *Journal of Applied Ecology* 42:400-408. <https://doi.org/10.1111/j.1365-2664.2005.01012.x>.
- McCord, S. E., N. P. Webb, J. W. Van Zee, S. H. Burnett, E. M. Christensen, E. M. Courtright, et al. 2021. Provoking a cultural shift in data quality. *BioScience* 71: 647-657. <https://doi.org/10.1093/biosci/biab020>.
- McNutt, M. 2014. Reproducibility. *Science* 343:229. <https://doi.org/10.1126/science.1250475>.
- Mitchell, M. S., S. N. Sells, K. J. Barker, S. B. Bassing, A. C. Keever, S. C. Forshee, and J. W. Goerz. 2018. Testing a priori hypotheses improves the reliability of wildlife research. *Journal of Wildlife Management* 82:1568-1571. <https://doi.org/10.1002/jwmg.21568>.

- Munafó, M. R. and G. D. Smith. 2018. Repeating experiments is not enough. *Nature* 553:399-401. <https://doi.org/10.1038/d41586-018-01023-3>.
- Nichols, J. D., W. L. Kendall, and G. S. Boomer. 2019. Accumulating evidence in ecology: Once is not enough. *Ecology and Evolution* 9:13991-14004. <https://doi.org/10.1002/ece3.5836>.
- Oza, A. 2023. Reproducibility goes on trial in ecology. *Nature* 622:677-678. <https://doi.org/10.1038/d41586-023-03177-1>.
- Parker, T. H., E. Main, S. Nakagawa, J. Gurevitch, F. Jarrad, and M. Burgman. 2016. Promoting transparency in conservation science. *Conservation Biology* 30:1149-1150. <https://doi.org/10.1111/cobi.12760>.
- Peng, R. D. 2011. Reproducible research in computational science. *Science* 334:1226-1227. <https://doi.org/10.1126/science.1213847>.
- Perez-Velázquez, J. L. 2019. The rise of the scientist-bureaucrat. Springer, New York, New York, USA. <https://doi.org/10.1007/978-3-030-12326-0>.
- Popper, K. 1959. The logic of scientific discovery. Basic Books, New York, New York, USA.
- Powers, S. M., and S. E. Hampton. 2019. Open science, reproducibility, and transparency in ecology. *Ecological Applications* 29:e01822. <https://doi.org/10.1002/eap.1822>.
- Roper, W. L. 2022. Science, health, and truth. *Science* 377:7. <https://doi.org/10.1126/science.add6477>.
- Sagario, M. C., V. R. Cueto, A. Zarco, R. G. Pol, and L. Marone. 2020. Predicting how seed-eating passerines respond to cattle grazing in a semi-arid grassland using seed preferences and diet. *Agriculture, Ecosystems and Environment* 289:e106736. <https://doi.org/10.1016/j.agee.2019.106736>.
- Seymour, C. I., and W. R. J. Dean. 2010. The influence of change of habitat structure on the species composition of bird assemblages in the southern Kalahari. *Austral Ecology* 35:581-592. <https://doi.org/10.1111/j.1442-9993.2009.02069.x>.
- Sikorski, M. 2022. Is forensic science in crisis? *Synthese* 200:188. <https://doi.org/10.1007/s11229-022-03685-z>.
- Sol, D., J. Maspons, M. Vall-llosera, I. Bartomeus, G. E. García-Peña, J. Piñol, and R. P. Freckleton. 2012. Unraveling the life history of successful invaders. *Science* 337:580-583. <https://doi.org/10.1126/science.1221523>.
- Werner, E. E. 1998. Ecological experiments and a research program in community ecology. Pp. 3-26 in W. J. Reserits and J. Bernardo (eds.). *Experimental ecology. Issues and perspectives*. Oxford University Press. New York.