

¿Ecólogos o Ególogos? Cuando las ideas someten a los datos

ALEJANDRO G FARJI-BRENER ✉

Laboratorio Ecotono-INIBIOMA, CRUB-Universidad Nacional del Comahue, Bariloche, Argentina

RESUMEN. Algunos ecólogos se apasionan tanto por sus ideas que terminan manipulando la información para que los resultados se ajusten a sus predicciones. En este trabajo describo ciertas características de estos investigadores (denominados ególogos porque viven el fracaso de sus hipótesis como una derrota de su ego) y detallo las formas mediante las cuales logran someter a los datos por las ideas. La manipulación de datos raros y la inquisición estadística (torturar los datos hasta que confiesen lo que se desea escuchar) son algunos de los procedimientos más comunes. Para comprender mejor el funcionamiento de la naturaleza, los ecólogos deberíamos ser esclavos de los datos y no de las hipótesis que guiaron su recolección.

[Palabras clave: diseño experimental, ego académico, manipulación estadística]

ABSTRACT. Ecologists or egologists? When the ideas enslave the data: Some ecologists feel in love with its ideas and, as a result, they tend to manipulate the information to artificially improve the fit between results and predictions. In this work I describe certain characteristics of this researchers (denominated egologist because they suffer the rejection of its ideas as a defeat of its ego), and detail how they enslave the data by their ideas. The manipulation of outliers' data and statistical inquisition (to torture the data until they confess) are some of the more frequent procedures. To better understand how nature works we, as ecologists, should be enslaved by the data and not by the hypotheses that guided their collection.

[Keywords: experimental design, academic ego, statistical manipulation]

Existen muchas ideas posibles que pueden explicar las causas de un patrón en la naturaleza. Sin embargo, el ecólogo elige probar solo algunas de ellas. Esta elección depende de muchos factores, incluyendo aquellos de índole histórica, sociológica, logística y hasta fortuita (Kuhn 1983). Sin embargo, es innegable que la elección de las ideas que un investigador pone a prueba involucra algo personal. Las hipótesis elegidas para guiar una investigación pueden representar indirectamente el grado de creatividad, nivel de información y capacidad deductiva de un ecólogo (de alguna forma, su ego). En otras palabras, elegir una idea y descartar otras para poner a prueba puede entenderse como una

apuesta intelectual en la mesa de juegos de la comunidad académica.

Los investigadores pueden vivir de diferentes maneras esta apuesta académica. Algunos la consideran como un medio para conocer los misterios de la naturaleza, y no temen que sus ideas sean incorrectas. Esta clase de investigadores acepta descubrir que su razonamiento estaba equivocado, porque entienden que la ciencia también avanza gracias al rechazo de hipótesis erróneas (Graham & Dayton 2002). Comprenden que la derrota parcial de su ego académico representa una victoria en la batalla por comprender cómo funciona la naturaleza. Contrariamente, otros investigadores se

✉ Laboratorio Ecotono-INIBIOMA, CRUB-Universidad Nacional del Comahue, (8400) Bariloche, Argentina.
alefarji@crub.uncoma.edu.ar, alefarji@yahoo.com

Recibido: 5 de enero de 2009; Fin de arbitraje: 3 de mayo de 2009; Revisión recibida: 12 de mayo de 2009; Segunda revisión recibida: 22 de mayo de 2009; Aceptado: 9 de junio de 2009

apasionan tanto de sus ideas que sufren su inexactitud como una derrota personal. Estos ecólogos se enamoran de las hipótesis que proponen (i.e., como una manifestación de su ego), en vez de enamorarse de los datos que estas ideas generan (i.e., como una manifestación de lo que sucede en la naturaleza). Por ende, este apasionamiento los lleva a manipular la información para que sus ideas triunfen, aun cuando no representan la realidad. A esta clase de investigadores los denominaré ególogos, porque es el ego -y no la curiosidad- lo que dirige su forma de trabajar.

SEIS MANERAS DE SOMETER LA INFORMACIÓN AL PODER DE LAS IDEAS (Y ALGUNAS SUGERENCIAS PARA EVITARLO)

El desarrollo de una investigación implica informarnos, ir al campo, diseñar un muestreo, recolectar datos, analizarlos e interpretar dicha información. En todas estas instancias existen formas para que los ególogos sometan los datos al poder de sus ideas. Este sometimiento puede comenzar en el origen mismo del proyecto, continuar a través de la recolección de datos, y finalizar con el análisis estadístico, mediante formas que van creciendo en complejidad. A continuación detallo algunas de las formas por las cuales, directa o indirectamente, un investigador puede imponer sus ideas independientemente de lo que sugiera la información obtenida.

1) *Apostar a lo seguro*

En este caso, el investigador elige poner a prueba ideas que, por información previa o sentido común, sabe de antemano que tienen una enorme probabilidad de verificarse. Esta forma encierra una paradoja para el ególogo. Por un lado, asegura la victoria de las hipótesis en su imaginaria batalla contra los datos. Pero por otra parte, esa victoria posee el sabor amargo de quien enfrenta y gana a un enemigo mucho más débil. Por lo general, los científicos principiantes eligen esta forma, dado que -posiblemente por inseguridad o falta de experiencia- plantean preguntas muy sencillas

cuyas respuestas son obvias, o interrogantes sesgados hacia problemas planteados (y resueltos) por otros investigadores con mayor experiencia. Ambas problemáticas podrían superarse apostando a la creatividad y a la capacidad crítica, postulando preguntas originales y relevantes cuyas respuestas sean una incógnita hasta el final de la investigación.

2) *Indiferencia académica*

En esta forma, los investigadores perciben sólo los datos que ajustan a sus predicciones. Aquellos resultados, observaciones de campo o información bibliográfica contrarios a lo que esperan son ignorados y excluidos de la investigación. Esta ceguera académica supone erróneamente que lo que se descarta, no existe. En este caso, la victoria de las ideas sobre los datos se da por un acto voluntario de omisión. Para superar este sesgo, toda información pertinente a la pregunta que guía la investigación debería incluirse sin especular si apoya o rechaza la idea previa que se tenía sobre el problema en estudio.

3) *Azar dirigido*

Durante el muestreo se puede torcer fácilmente el rumbo del azar hacia el puerto de las ideas. De forma más o menos inconsciente, un investigador puede elegir árboles más grandes, hojas con menos herbivoría, flores más coloridas, hormigas más rápidas o parcelas localizadas en determinados sitios según cómo estos sesgos faciliten que los datos apoyen sus hipótesis. En términos metafóricos, es como construir un escenario de dimensiones tan reducidas como para que pueda actuar un solo razonamiento. Y luego aplaudir el monólogo. Un muestreo planificado de antemano y ejecutado limpiamente puede minimizar este sesgo de manipular la información.

4) *Muestreo (im)perfecto*

El muestreo es la parte más débil de la saga, y es lo que primero se suele cortar cuando los resultados no sustentan el razonamiento que se suponía correcto. Pero, de manera paradójica, si los resultados apoyan las ideas que guiaron la investigación, pocas veces se considera

como posibles factores de sesgo al diseño de muestreo o a la recolección de los datos. En otras palabras: el mismo muestreo puede caer en el olvido o ser víctima de una autocritica feroz dependiendo de si los resultados apoyan o no las hipótesis propuestas, respectivamente. Todo diseño de muestreo puede ser motivo de autocritica, pero esta crítica siempre debe ser independiente de los resultados que se hayan obtenido a través de ese muestreo.

5) Manipulación numérica

Este procedimiento, lamentablemente común, se basa sobre el descarte o la inclusión de datos que se escapan del patrón general (datos "raros" o "outliers" en ingles) con el objetivo

de confirmar las ideas que se proponen. La decisión de eliminar o dejar los datos "raros" dependerá del efecto que ejerza su presencia o ausencia sobre los análisis estadísticos. Por ejemplo, si se desea probar que el peso de la semilla acarreada depende del peso de la hormiga que la transporta, se puede mantener (Figura 1a y b) o eliminar (Figura 1c y d) los datos "raros" para lograr la significancia estadística que apoye el preconcepto sobre el sistema estudiado. La pasión por las ideas ha promovido el secuestro selectivo de los datos. Ante este tipo de casos, plantear ambas situaciones (i.e., el análisis con y sin datos "raros") y dejar a los lectores que acompañen (o no) la interpretación del autor sería una forma

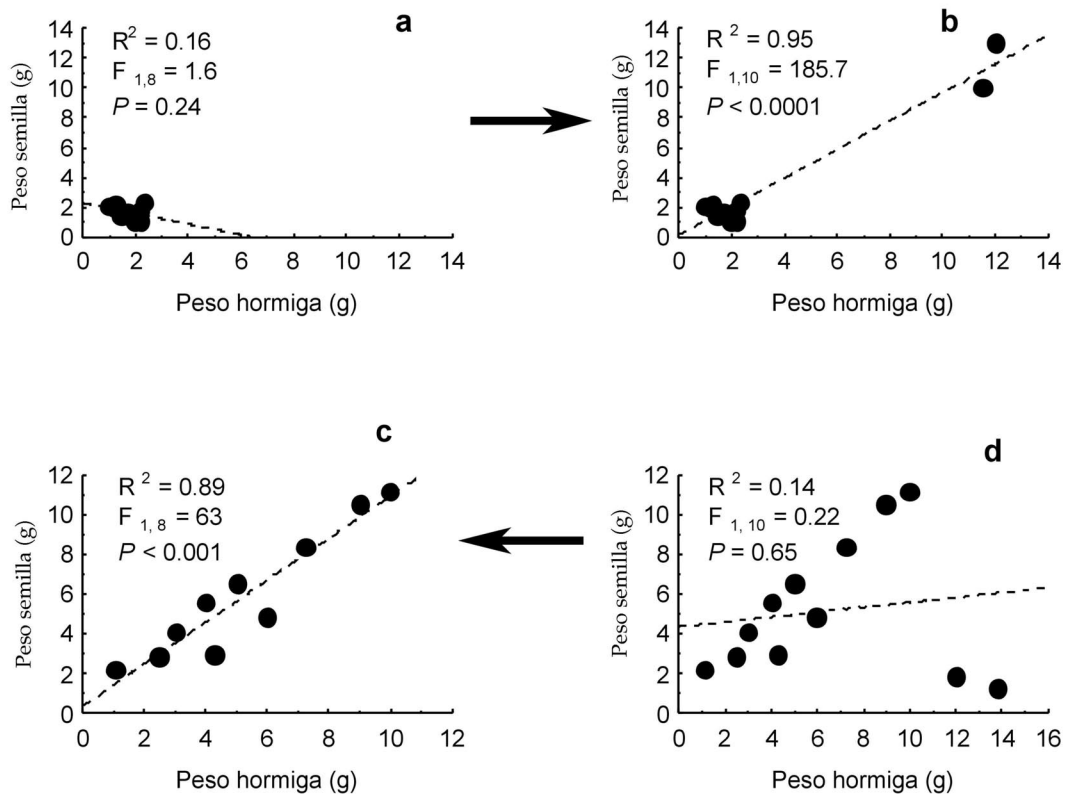


Figura 1. Ejemplo sencillo de manipulación numérica. Los datos "raros" son selectivamente incluidos (a→b) o descartados (c←d) sólo en función de lograr la significancia estadística que apoye la hipótesis propuesta.

Figure 1. Simple example of numerical manipulation. Outliers data are selectively incorporated (a→b) or discarded (c←d) with the merely purpose to obtain the statistical significance that support the proposed hypothesis.

mas honesta de poner a prueba las ideas que guiaron la investigación.

6) *Inquisición estadística (torturar los datos hasta que confiesen)*

En este proceso, el investigador utiliza diferentes pruebas estadísticas para lograr que los resultados de sus análisis se ajusten a sus predicciones. Las pruebas no son seleccionadas por ser las más adecuadas, sino por reunir la "virtud" de encontrar diferencias estadísticamente significativas (pero a veces biológicamente insignificantes). Los diferentes análisis se prueban con la tenacidad de quien cava buscando un tesoro escondido. En otras palabras, los mismos datos son sometidos a diversos instrumentos de tortura (e.g., las diferentes pruebas estadísticas) hasta que confiesan lo que el investigador desea escuchar. Supongamos que se pretende probar que las avispas prefieren las proteínas a los hidratos de carbono para alimentarse. Para ello se realiza un experimento distribuyendo al azar 20 potes con cebos (10 de cada tipo), y se registra la cantidad de avispas que recolectan alimento en cada pote en un tiempo dado (Tabla 1). Dependiendo de cómo se analicen estos datos, la idea que se quería poner a prueba será corroborada o no. Si no se considera la variación entre repeticiones (i.e., entre cada pote) y se usan los totales para realizar un análisis de X^2 , los resultados estadísticos apoyarán la idea previa de que las avispas prefieren proteínas. Si se analizan los mismos datos mediante una prueba de t , el resultado no será acorde a la predicción propuesta y la idea de una preferencia por proteínas debería ser rechazada (Tabla 1). Independientemente de que la última opción es la más correcta porque incorpora la variación entre unidades experimentales y evita el problema de pseudo-repetición sacrificante (Hulbert 1984), el investigador guiado por su ego elegirá la primera opción porque ese resultado se ajusta más a su pensamiento. Existen variantes mucho más sofisticadas que este sencillo ejemplo. Algunas incluyen utilizar pruebas estadísticas complejas o modelos incomprensibles para un lector promedio cuando los mismos datos podrían (y a veces hasta deberían) analizarse de forma más sencilla (Murtaugh 2007), o transformar

Tabla 1. Ejemplo sencillo de inquisición estadística (o cómo torturar los datos hasta que confiesen). Resultados de un experimento imaginario para verificar si las avispas prefieren recolectar cebos de proteínas (P) sobre los de hidratos de carbono (HC). Cada unidad experimental (i.e., cada repetición) consiste en un pote (con P ó HC) ubicado al azar (10 de cada tipo). Los números son la cantidad de avispas por pote que recolectaron cebo en un tiempo dado. La idea propuesta (la preferencia de las avispas por proteínas) puede ser apoyada o rechazada dependiendo de la prueba estadística utilizada (X^2 o prueba de t , respectivamente).

Table 1. Simple example of statistical inquisition (or how torture the data until they confess). Results of an imaginary experiment to test whether wasps prefer protein (P) than carbohydrate baits (CH). Each experimental unit (i.e., the replicate) was a recipient (with P or CH) randomly located (10 per type). The values are the number of wasp recollecting baits in a determinate lapse. The proposed prediction (the preference of wasp for protein baits) can be supported or rejected depending the statistical test employed (X^2 or t -test, respectively).

Experimento	H. Carbono	Proteínas
1	2	32
2	3	2
3	1	2
4	0	1
5	3	3
6	2	3
7	3	2
8	1	2
9	2	1
10	4	4
Totales	21	52
Media \pm SD	2.1 \pm 1.2	5.2 \pm 9.5

$$t = 1.03$$

$$P = 0.32$$

$$g.l. = 18$$

$$X^2 = 13.16$$

$$P = 0.003$$

$$g.l. = 1$$

los datos de formas tan intrincadas hasta lograr la significancia estadística, aunque luego su interpretación biológica sea prácticamente imposible. Esta forma de sometimiento es la más sofisticada y requiere cierto conocimiento de estadística, por lo cual la emplean, en general, los investigadores más experimentados.

Una solución a esta mala praxis académica es estudiar, consultar y elegir el método estadístico más adecuado a la pregunta que se realiza, al diseño experimental y a la naturaleza de nuestros datos "antes" de aplicarlo (e inclusive antes de la recolección de los datos). Y no cambiar de análisis porque los resultados no se ajustan a las predicciones.

Estos mecanismos de manipulación de la información también pueden tener lugar en los trabajos de investigadores que no se guían sólo por su ego. Por ejemplo, pueden ser producto de un diseño o análisis inadecuado. En otras palabras, un muestreo sesgado o el uso incorrecto de las herramientas estadísticas pueden deberse simplemente a la falta de experiencia y/o de conocimiento. En este caso, estamos ante la presencia de un investigador que comete errores. Por otra parte, los mecanismos descritos también podrían utilizarse para mejorar la productividad científica de un investigador, ya que los resultados que apoyan hipótesis son, por lo general, más fáciles de publicar que aquellos que las rechazan (Rosenthal 1979; Csada et al. 1996; Palmer 1999; Farji-Brener 2006). En este caso, los sesgos son cometidos de manera conciente, pero el objetivo final del autor no necesariamente implicaría fortalecer su amor propio. Sus metas podrían ser la obtención de alguna beca o un ascenso en la carrera académica a través de una tasa de publicación mayor. Estos investigadores presentan similitudes con los ególogos en cuanto a mecanismos, pero no necesariamente en cuanto a metas. Los investigadores guiados por su ego desean "algo más" que publicar: desean ser reconocidos intelectualmente. En este sentido, sus publicaciones no son sólo un medio para progresar académicamente, sino una manera para demostrarle al mundo lo apropiado de sus ideas.

¿BUDISMO ECOLÓGICO?

Casi todos, en menor o mayor medida, tenemos ego. Plantear que los investigadores debemos carecer de ego, tal como lo proponen ciertas disciplinas orientales, parece un poco ingenuo e irreal. Una alternativa es reconocer al ego como un motor de la actividad científica,

e intentar disciplinarlo y encausarlo hacia la creación honesta. Esto implica que deberíamos aceptar el fracaso de una idea si las evidencias así lo sugieren, discriminar si los resultados son tan potentes como para descartar la hipótesis o simplemente no son capaces de ponerla a prueba efectivamente, y desafiar a nuestro ego para generar nuevas ideas (en el primer caso) o para lograr un mejor diseño de muestreo (en el segundo). De esta forma, nuestro ego podría dejar de funcionar como un cristal que sólo refleja la información útil a nuestros preconcepciones, y ser aprovechado como un motor superador y creativo.

En nuestra búsqueda del conocimiento existen medios y fines. En este artículo intenté describir, a través de descripciones y metáforas extremas, cómo los medios y los fines pueden invertirse cuando el apasionamiento por una idea oscurece nuestra objetividad como científicos. En el universo del ecólogo común, las ideas deberían ser un medio para descubrir lo que nos sugieren los datos sobre el funcionamiento de la naturaleza. Cuando el orden se invierte y los datos se transforman en un simple medio para corroborar nuestro pensamiento, deberíamos detenernos y reflexionar sobre nuestra conducta como científicos. Y recordar que para comprender mejor el entorno que nos rodea debemos ser esclavos de los datos, no de las hipótesis que guiaron su recolección.

AGRADECIMIENTOS

Los comentarios de W. Eberhard, E. Jobbágy y María Semmartin colaboraron mucho a mejorar una versión preliminar de este manuscrito.

BIBLIOGRAFÍA

- CSADA, R; P JAMES & R ESPIE. 1996. The "file drawer problem" of non-significant results: does it apply to biological research? *Oikos* 76:591-593.
- FARJI-BRENER, AG. 2006. La (significativa) importancia biológica de la no-significancia estadística. *Ecología Austral* 16:79-84.
- GRAHAM, M. & P DAYTON. 2002. On the evolution

- of Ecological ideas: paradigms and scientific progress. *Ecology* **83**:1481-1489.
- HULBERT, S. 1984. Pseudoreplication and the design of ecological field experiments. *Ecol. Monogr.* **54**:187-211.
- KUHN, TS. 1983. *La estructura de las revoluciones científicas*. Fondo de Cultura Económica, México, DF.
- MURTAUGH, P. 2007. Simplicity and complexity in ecological data analysis. *Ecology* **88**:56-62.
- PALMER, AR. 1999. Detecting publication bias in meta-analyses: a case study of fluctuating asymmetry and sexual selection. *Am. Nat.* **154**:220-233.
- ROSENTHAL, R. 1979. The "file drawer problem" and tolerance for null results. *Psychol. Bull.* **86**:638-641.