nature ecology & evolution

ARTICLES

https://doi.org/10.1038/s41559-018-0610-7

Logarithmic scales in ecological data presentation may cause misinterpretation

Duncan N. L. Menge[1*], Anna C. MacPherson[2], Thomas A. Bytnerowicz[1], Andrew W. Quebbeman[1], Naomi B. Schwartz[1], Benton N. Taylor[1] and Amelia A. Wolf[1,3]

Scientific communication relies on clear presentation of data. Logarithmic scales are used frequently for data presentation in many scientific disciplines, including ecology, but the degree to which they are correctly interpreted by readers is unclear. Analysing the extent of log scales in the literature, we show that 22% of papers published in the journal *Ecology* in 2015 included at least one log-scaled axis, of which 21% were log–log displays. We conducted a survey that asked members of the Ecological Society of America (988 responses, and 623 completed surveys) to interpret graphs that were randomly displayed with linear–linear or log–log axes. Many more respondents interpreted graphs correctly when the graphs had linear–linear axes than when they had log–log axes: 93% versus 56% for our all-around metric, although some of the individual item comparisons were even more skewed (for example, 86% versus 9% and 88% versus 12%). These results suggest that misconceptions about log-scaled data are rampant. We recommend that ecology curricula include explicit instruction on how to interpret log-scaled axes and equations, and we also recommend that authors take the potential for misconceptions into account when deciding how to visualize data.

Clear presentation of results is essential for scientific progress. One aspect of data display that poses challenges for clarity is the use of logarithms. Logarithms convert multiplicative relationships to additive ones, providing an elegant way to span many orders of magnitude[1], to show elasticities and other proportional changes[2], and to linearize power laws[3]. They are used in canonical scales of acidity[4], earthquake magnitude[5] and star brightness[6], and are frequently used for presenting income[7], time[1] and other quantities. They arise from the fundamental mathematics of population growth[8], radioactive decay[9] and other processes. They also have practical purposes, easing the computation of small numbers such as likelihoods[10] and transforming data to fit statistical assumptions[11]. Despite being useful, however, logarithms are fraught with misconceptions[12–15]. These misconceptions have been documented in high school[12,14,15] and college[13] students, but we hypothesized that logarithms would also be misinterpreted by many practising scientists. We focused on ecologists, who have substantial statistical training and frequently use and encounter logarithms. Species–area curves[16], allometry[17] and other scaling relationships[18], population growth[8], metrics of biodiversity[19] and biostatistics[11] are among the many standard topics in ecology that involve logarithms.

Here, we asked two broad questions. How often do ecologists encounter log-scaled data? How well do ecologists understand log-scaled data? To address these questions, we conducted a bibliometric analysis of all papers published in the journal *Ecology* in 2015 and a survey of ecologists (Methods, Supplementary Information). The survey asked respondents to evaluate a series of graphs and equations, to describe their level of comfort with logarithms, to express their preferences about data presentation and to explain when and why they presented data on logarithmic scales.

Results

Our bibliometric analysis showed that 10% of numerical (that is, not categorical) axes in graphs were log-scale, 22% of papers contained at least one log-scaled axis, and 44% of papers reported using logarithms in some way (Fig 1). When log scales were used in graphs, the corresponding figure caption or text usually (85%) mentioned the log scale (Fig 1). Bivariate graphs with at least one log-scaled axis were typically (74%) bivariate numerical (*x*–*y*) layouts rather than categorical on one axis (Fig. 1). Among bivariate numerical graphs, those with a linear-scale horizontal axis and a log-scale vertical axis (linear–log) were most common (32% of all bivariate graphs), followed by log–log (21%) and log–linear (16%) axes (Fig. 1). It was common (39% of panels) for more than one line or curve to be plotted on bivariate graphs with one or two log-scaled axes (Fig. 1). Although there are prominent examples of three-dimensional log–log–log graphs[20], we encountered only ten log-scaled *z* axes in our bibliometric analysis (Fig. 1).

When presenting log-scaled data, it is possible to use either untransformed values (for example, values of 1, 10 and 100 are equally spaced along the axis) or log-transformed values (for example, 0, 1, and 2). According to our bibliometric analysis, both untransformed (27%) and log-transformed values (73%) were relatively common on axis labels (Fig. 1). Overall, our results suggest that ecologists will encounter log-scaled axes regularly, often on bivariate numerical axes, and with both transformed and untransformed values. Therefore, any misconceptions about such logarithmic displays that are common among ecologists probably have a major impact on the extent to which ecologists understand ecological literature. Our survey was designed to evaluate the extent and type of such misconceptions.

The main part of our survey presented respondents with graphs that were randomly displayed on linear–linear scales (Fig. 2a,d and g), log–log scales with untransformed values (Fig. 2b,e and h), or log–log scales with log-transformed values (Fig. 2c,f and i). Two relationships were shown on each graph, with distance from the edge of a habitat on the horizontal axis and population size (of rabbits and chipmunks) on the vertical axis. We asked four

[1]Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, NY, USA. [2]American Museum of Natural History, New York, NY, USA. [3]Department of Plant Sciences, University of California, Davis, CA, USA. *e-mail: dm2972@columbia.edu

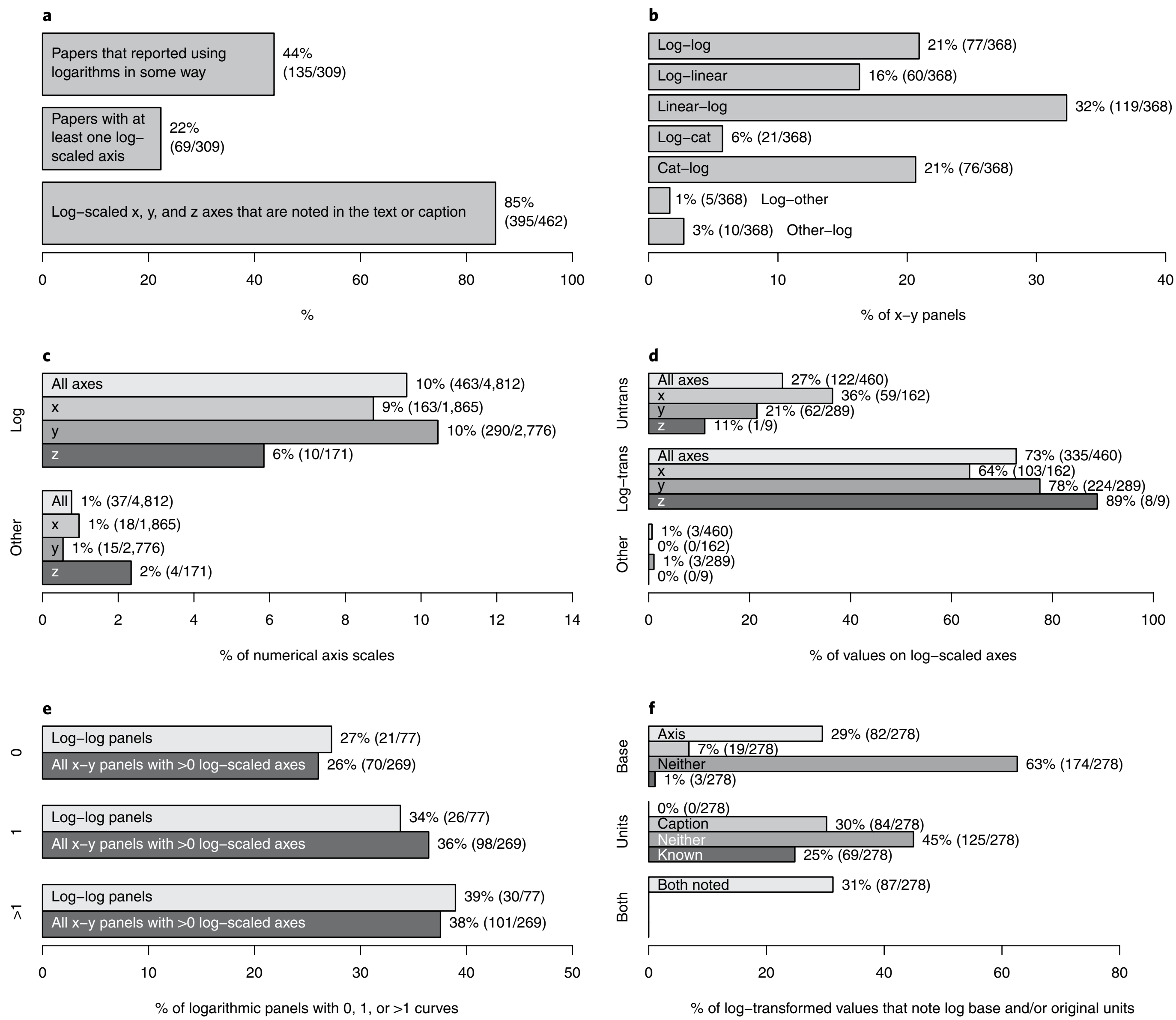

Fig. 1 | Logarithm usage in papers published in *Ecology* in 2015. All horizontal axes are percentages, with the group noted on the bars. Numbers listed to the right of bars are the percentages shown along with exact values in parentheses. **a**, Paper-wide general logarithm usage. **b**, Plot layouts, shown as the % of bivariate (x–y) panels in different forms. 'Cat' indicates 'categorical'. **c**, Breakdown of numerical axis scales. Logarithmic and 'other' (non-linear, non-logarithmic scales such as square-root, reciprocal sine and others) axes are shown, for all axes, x axes, y axes and z axes. Linear-scale axes comprise the remainder: 90% of all axes (4,312/4,812), 90% of x axes (1,684/1,865), 89% of y axes (2,471/2,776) and 92% of z axes (157/171). **d**, Breakdown of values for log-scaled axes into untransformed, log-transformed and 'other' values. **e**, Numbers of curves shown on log–log panels (light bars) or panels with at least one log-scaled axis (dark bars). **f**, Breakdown of whether log bases and original units are shown for axes with log-transformed values. In the top and middle groups of bars, light and light-medium grey bars indicate that the log base (top group) or the original units (middle group) were noted on the axis or in the caption, respectively, whereas the medium-dark bar indicates that they were not noted on either. 'Known' indicates that the log base or units were not indicated, but are easily assumed. For the log base, 'known' includes pH axes, for which we assume that base 10 is common knowledge. For units, 'known' includes values that are generally known to be unitless (such as log response ratios) or are otherwise generally known (such as pH). The bottom group shows only a single value, the % of log-transformed axes for which both the log base and the original units were either noted in the axis or caption or were known.

questions about the graphs. First, we asked respondents whether each population was increasing or decreasing with distance. Nearly all (99%) respondents answered correctly for all datasets, regardless of how data were displayed (Supplementary Fig. 1).

Second, we asked respondents how steeply the rabbit population changed with distance compared to the chipmunk population. In our first dataset, the rabbit population always increased less steeply than the chipmunk population (Fig. 2a–c), whereas in the second dataset the rabbit population changed more steeply near the edge of the habitat but less steeply further from the edge (Fig. 2d–f). When the data were displayed on linear–linear axes, 86% of the respondents correctly identified these relative changes (Fig. 3a and d), compared to only 9% when the data were presented on log–log axes (Fig. 3b,c,e and f). In our third dataset, where the rabbit population always increased less steeply (Fig. 2g–i) but both populations started at a non-zero value (and thus look different on log–log scales than

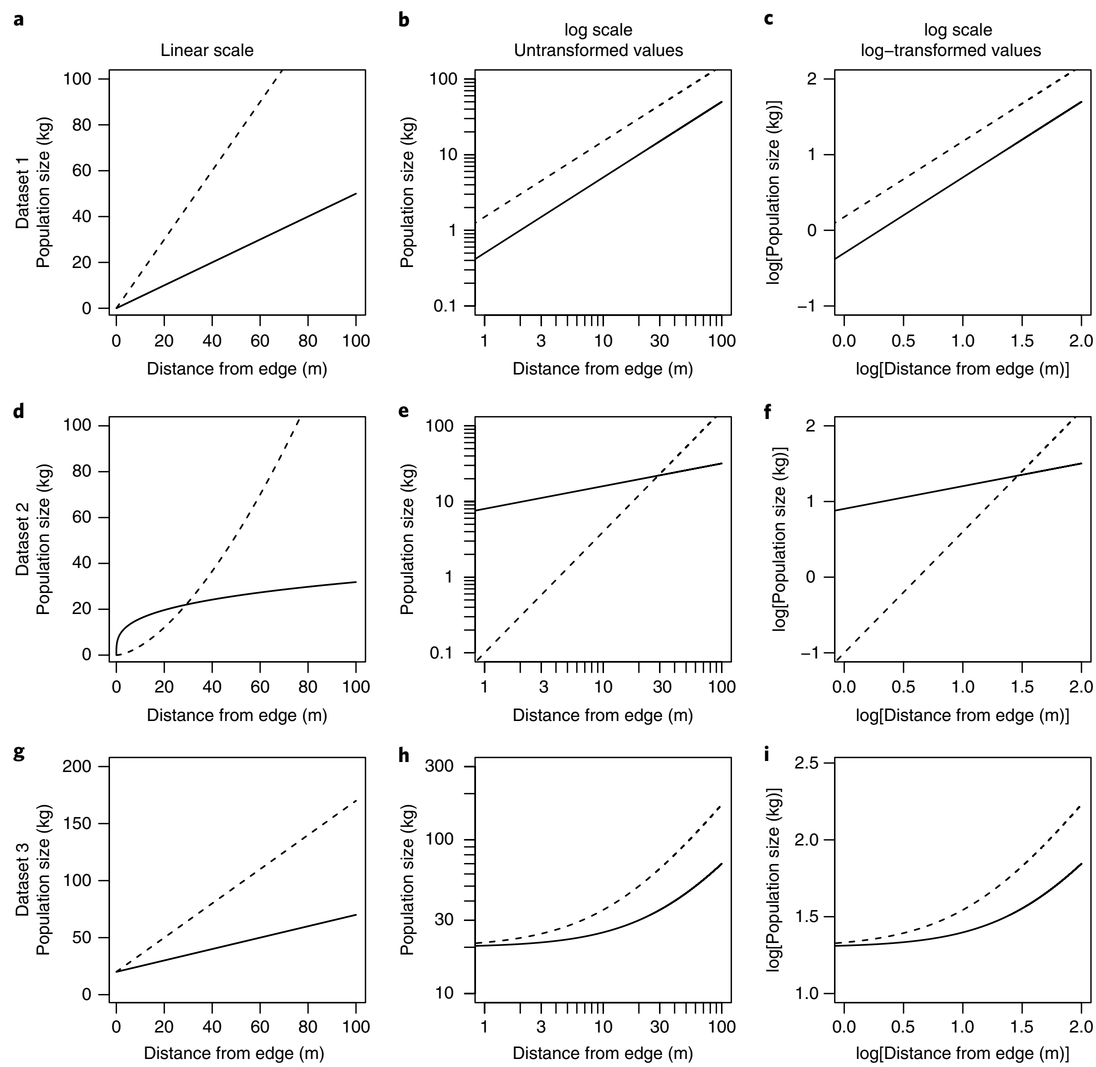

Fig. 2 | Graphs viewed in our survey. a–i, Population sizes of rabbits (solid) and chipmunks (dashed) are shown on linear–linear (**a**,**d**,**g**) and log–log (**b**,**c**,**e**,**f**,**h**,**i**) axes. log–log axes are shown with both untransformed (**b**,**e**,**h**) and log-transformed (**c**,**f**,**i**) values. All populations increase with distance from the habitat edge. In the first dataset (**a–c**), both populations are 0 at the habitat edge and increase linearly, and the rabbit population increases less steeply than the chipmunk population. In the second dataset (**d–f**), both populations are 0 at the habitat edge. Rabbits increase more steeply than chipmunks near the edge, but less steeply further from the edge. Rabbits increase in a decelerating fashion, whereas chipmunks increase in an accelerating fashion. In the third dataset (**g–i**), both populations are positive at the habitat edge and increase linearly, but chipmunks increase with distance faster than rabbits. Note that increasing lines in log–log space must intersect {0,0} in linear space, and can have accelerating, constant or decelerating trends (**a–f**). Note also that a non-zero, finite intercept in linear–linear space makes a curve in log–log space, even if the linear space relationship is a line (**g–i**). See also Supplementary Table 2 and Methods.

in the first dataset), responses were similar across all graph display types (Fig. 3g–i).

The large percentage of incorrect responses to our second question when the first two datasets were viewed on log–log axes (Fig. 3b,c,e and f) suggests a misconception—incorrect knowledge—not merely a lack of knowledge. No more than 6% of respondents chose 'wouldn't notice' (Supplementary Table 1), implying that respondents had some degree of confidence in their incorrect responses. The incorrect responses to log–log displays were also consistent with an obvious possible misinterpretation: that the slope of the log–log relationship indicates the slope of the linear–linear relationship (in truth it shows elasticity—the percentage change in y with a percentage change in x (ref. [2])—but not the steepness of the linear–linear relationship). When the log–log relationships were parallel lines, 80% of respondents said (incorrectly) that the two populations changed similarly steeply (Fig. 3b,c); when the log–log relationships converged and crossed, 84% said (incorrectly) that the rabbit population always changed less steeply (Fig. 3e,f); and when the log–log relationships diverged, 75% said (correctly, in this case) that the rabbit population always changed less steeply (Fig. 3h,i).

This suggests a misconception we call the 'hand-hold fallacy' (Table 1). Climbing something steep is harder than climbing something gradual, but climbing something smooth (such as a window) is harder than climbing something with hand-holds (such as a ladder), regardless of slope. The slope of a climb matters, but hand-holds also matter. The same is true for log–log graphs. The slope matters, but it is not the only thing that does; the intercept and the location on the horizontal axis also matter. A line in log–log space, $\ln(y) = \ln(c) + a\ln(x)$, represents a power law, $y = cx^a$. It might look like the log–log slope (a) indicates how steeply y changes with x, but c and x also matter, just like hand-holds. This can be seen by the appearance of c and x along with a in the derivative $\frac{dy}{dx} = acx^{a-1}$.

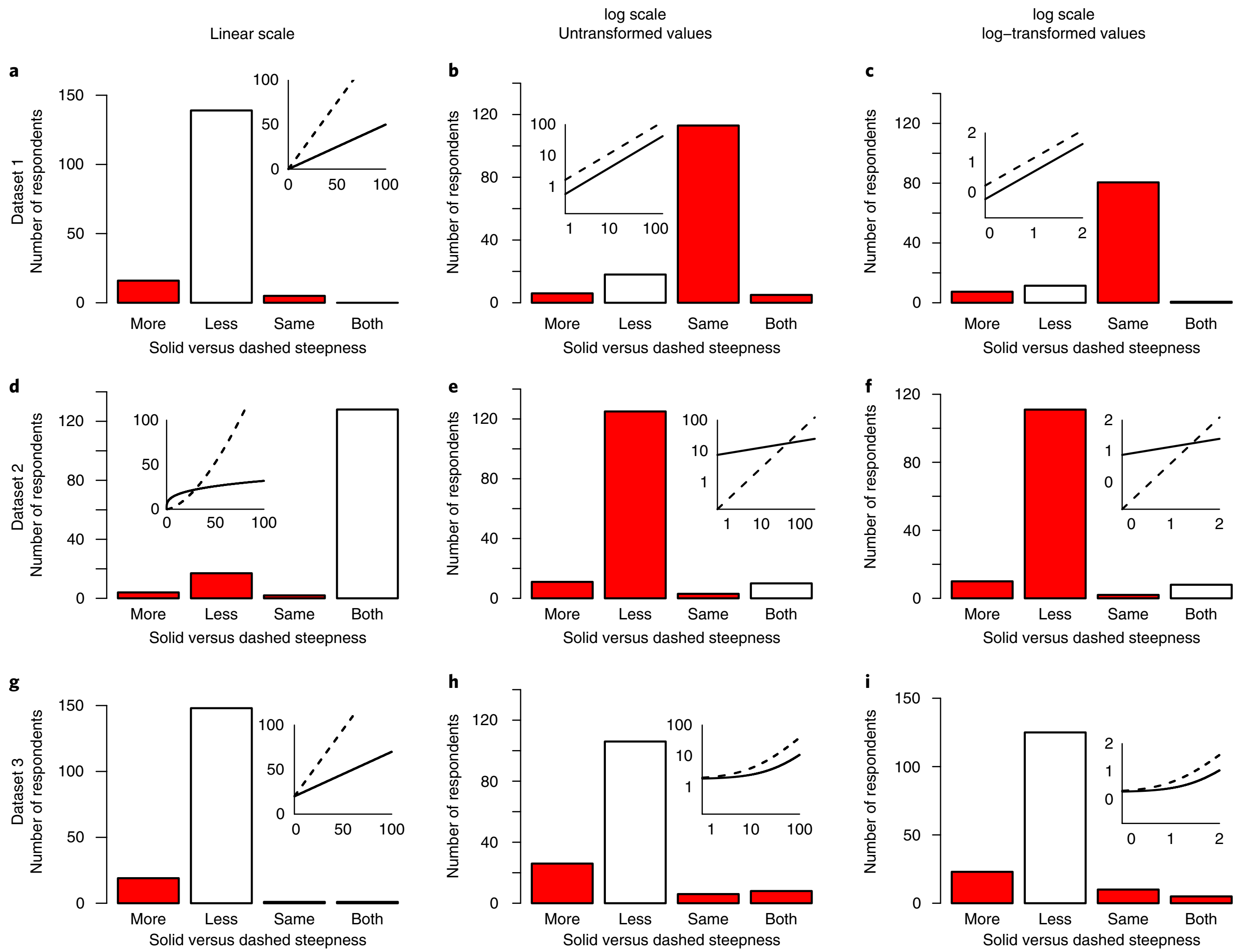

Fig. 3 | Survey responses about which population changes more steeply with distance. a–i, Each bar shows the number of responses in each category; 'wouldn't notice' responses are shown in Supplementary Table 1. These were responses to viewing three hypothetical datasets with linear–linear (**a**,**d**,**g**) and log–log scales (**b**,**c**,**e**,**f**,**h**,**i**). log–log scale graphs were viewed with untransformed (**b**,**e**,**h**) and log-transformed (**c**,**f**,**i**) values. Simplified forms of the graphs displayed to respondents are shown in the insets; the exact graphs are in Fig. 2. 'Both' is short for 'The solid population, compared to the dashed population, sometimes changes more steeply but sometimes changes less steeply with distance.' The full survey flow and items are in the Supplementary Information. Incorrect responses are red.

Third, we asked respondents about the population level at the edge of the habitat. All populations in the first two datasets were zero at the edge of the habitat, which 88% of respondents identified correctly when the data were displayed on linear–linear axes (Fig. 4a and d). In contrast, only 12% identified the correct population sizes on log–log axes (18% with untransformed values; Fig. 4b and e; 7% with log-transformed values; Fig. 4c and f). Both populations in the third dataset were positive at the edge of the habitat, which nearly all respondents identified correctly, regardless of scale (Fig. 4g–i). The breakdown of responses for the first two datasets with untransformed values, along with the small number of 'wouldn't notice' responses (Supplementary Table 1), suggests another misconception. In the datasets where the log–log lines crossed the left side of the graph, 94% said the edge population sizes were positive (both curves in Fig. 4b, solid curve in Fig. 4e), but in the dataset where the log–log line crossed at the lower-left corner of the graph, 61% correctly replied 'zero' when values were untransformed (dashed curve in Fig. 4e). When viewing log-transformed values, a substantial fraction (37%) of respondents said that the population sizes were negative for curves that crossed the vertical axis below the log-transformed value of 0 (solid curve in Fig. 4c, dashed curve in Fig. 4f). This result suggests a misunderstanding of the meaning of negative log-transformed values, which imply untransformed values between 0 and 1 (for example, 0.5 kg biomass), not less than 0 (negative biomass, which would be nonsensical).

The responses to our third question when the first two datasets were viewed on log–log axes (Fig. 4b,c,e and f) imply a misconception we call the 'Zeno's zero fallacy' (Table 1). Paraphrased, Zeno's famous paradox states that a distance can never be reduced to zero because it must be halved infinitely many times and it is impossible to perform an infinite number of tasks. The paradox is clearly wrong (try walking a metre), but it illustrates how multiplicative processes (halving; logarithms) can lead to misconceptions. In our examples, it might look like the populations will never reach zero if they are descending slowly (Fig. 4b,c,e and f), but we know mathematically that they will. Even though the logarithm of 0 is undefined, a positively sloped line in log–log space unambiguously implies a zero value of y when x is zero (see $y=cx^a$ when $a>0$). Conversely, functions that are lines with a non-zero intercept in linear–linear space ($y=b+cx$) approach a horizontal (if $b>0$; Fig. 2g–i) or vertical

Table 1 | Logarithm fallacies and related misconceptions

Fallacies	Observed misconceptions	Possible underlying misconception
Hand-hold fallacy	Steeper slopes in log–log relationships are steeper slopes in linear space. (Not always true; depends on the log–log slope, intercept and x range. Figs. 2 and 3.)	Linear extrapolation error[12]: logarithms cancel out like variables in algebraic expressions, which would imply that a log–log slope can be interpreted as a linear–linear slope. (Not true; $\log(y)=\log(c)+z\log(x)$ transforms to $y=cx^z$, not to $y=c+zx$.)
Zeno's zero fallacy	Positively sloped lines in log–log space can imply a non-zero value of y when x is zero. (Never true; positively sloped lines in log–log space unambiguously imply that $y=0$ when $x=0$. Figs. 2 and 4.)	Linear extrapolation error[12], which would also imply that a log–log intercept can be interpreted as a linear–linear intercept. (Not true; $\log(y)=\log(c)+z\log(x)$ transforms to $y=cx^z$, not to $y=c+zx$.)
Watch out for curves fallacies	(1) Lines in log–log space are lines in linear–linear space. (Only true if the log–log slope is 1. Figs. 2 and 5a–f.)(2) Lines in log–log space curve upward in linear–linear space. (Only true if the log–log slope is greater than 1. Figs. 2 and 5a–f)(3) Curves in log–log space have the same curvature in linear–linear space. (Not always true. Figs. 2 and 5g–i.)	(1) and (3) Linear extrapolation error[12], as explained for the hand-hold fallacy above.(2) log–log lines represent power laws (which are exponential relationships), and all exponential relationships curve upward. (Not true; power laws only curve upward if the log–log slope is greater than 1.)

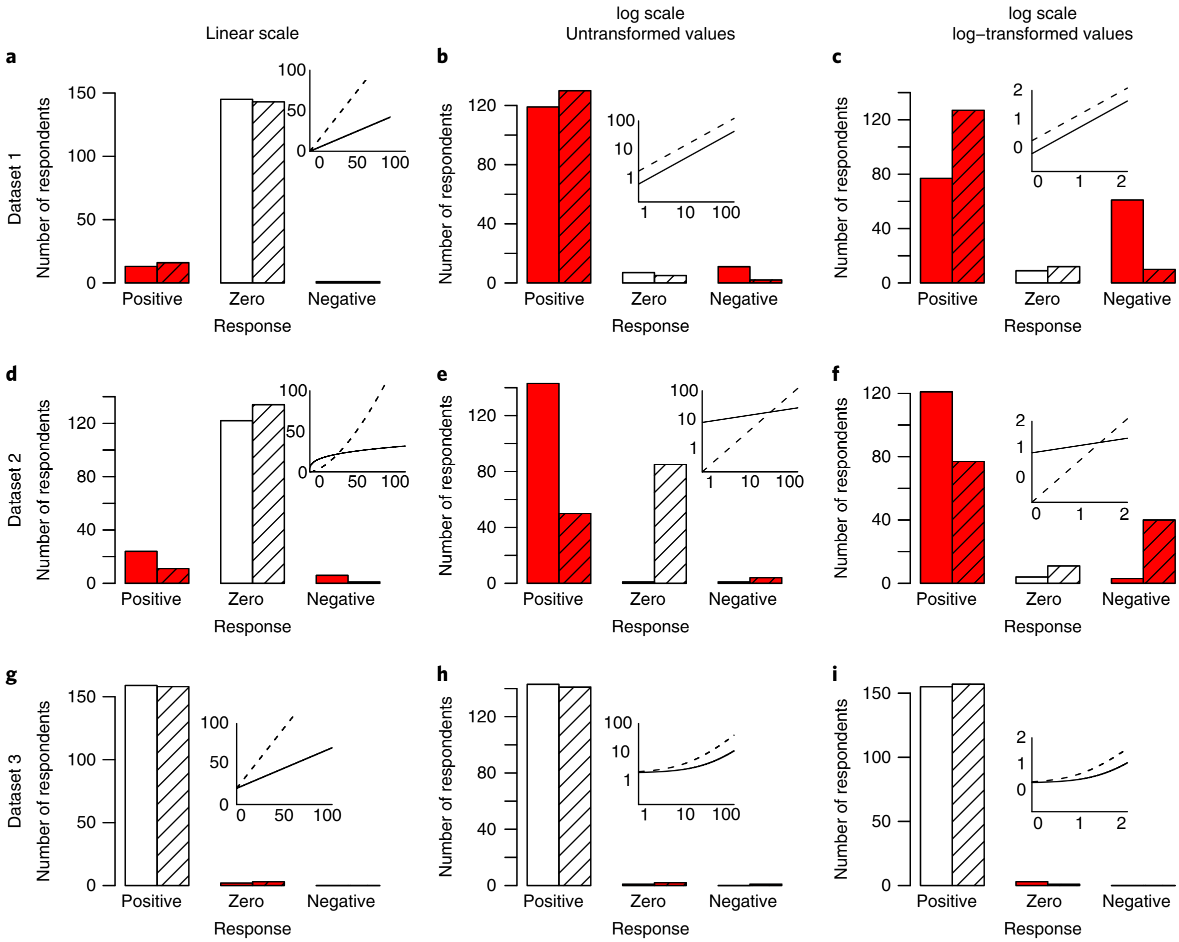

Fig. 4 | Survey responses about each population level at the edge of the habitat. a–i, Each bar shows the number of responses in each category for the rabbit (solid curves, solid bars) and chipmunk (dashed curves, hatched bars) populations. 'Wouldn't notice' responses are shown in Supplementary Table 1. Figure details as in Fig. 3. The full survey flow and items are in the Supplementary Information.

(if $b<0$) asymptote when plotted in log–log space, so are not visually misleading about the intercept (Fig. 4g–i).

The implications of Zeno's zero fallacy are not merely academic. Many data presented on log scales can reasonably be zero[21,22]. Authors have found a number of ways to present such data on logarithmic scales, all of which make sense but might exacerbate misconceptions about logarithms. For example, stream discharge[23], plant biomass[24] and nutrients[25] have been presented on a log scale, but with zero values for anything less than a particular order of magnitude. Fruit consumption was presented on a log scale on panels where there were no zeros, but on linear scales on panels where there were zeros[26]. The practice[27,28] of plotting $\log(x+1)$, which jettisons many of the good reasons for using logarithms (for example, that differences can no longer be interpreted as proportional or

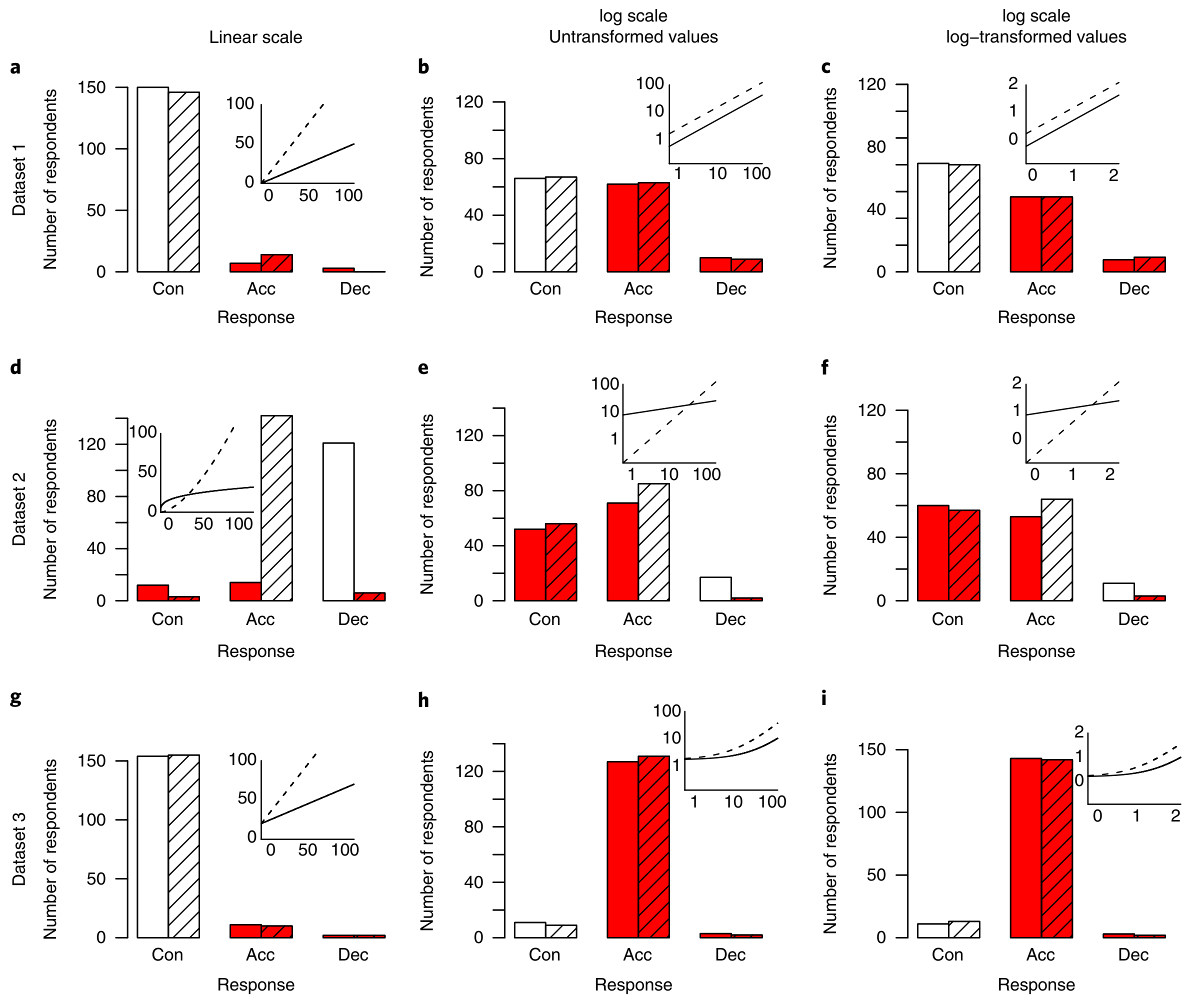

Fig. 5 | Survey responses about the manner in which each population changes with distance. a–i, Each bar shows the number of responses in each category for the rabbit (solid curves, solid bars) and chipmunk (dashed curves, hatched bars) populations. 'Wouldn't notice' responses are shown in Supplementary Table 1. Figure details as in Fig. 3. Acc, accelerating; Con, constant; Dec, decelerating. The full survey flow and items are in the Supplementary Information.

multiplicative changes), is a common way for data with zeros to be plotted on log scales. Even for datasets that do not contain zeros, there are many situations where it is reasonable to make an inference about what happens at zero. In the example on our survey, distance from the edge of a habitat ranged from 0.1 m to 100 m. Extrapolation from 0.1 m to 0 m seems justifiable for a question about population densities, but our results show that respondents extrapolate differently depending on whether the data are shown on logarithmic or linear scales.

Fourth and finally, we asked respondents about the manner in which a population changes. In our first and third datasets, both populations changed in a constant manner, but in our second dataset, the rabbit and chipmunk populations changed in decelerating and accelerating ways, respectively. When presented with linear–linear axes, 91% of respondents correctly identified these curvatures (Fig. 5a,d and g). In contrast, when the data were presented on log–log axes, responses were split between 'constant' (47%) and 'accelerating' (46%) for the first two datasets (Fig. 5b,c,e and f) and were nearly all (91%) 'accelerating' for the third dataset (Fig. 5h and i), regardless of the truth, suggesting three misconceptions (the three 'watch out for curves fallacies'; Table 1). The misconceptions are that all lines in log–log space are also lines in linear–linear space (which is only true if the log–log slope is parallel to the 1:1 line; $a = 1$), that all lines in log–log space curve upwards in linear–linear space (which is only true if the log–log slope is steeper than the 1:1 line; $a > 1$), and that curves in log–log space have similar curvature in linear–linear space (which is not necessarily true). Results from some of our additional survey items further support this conclusion. When asked to identify which of three graphs showed a constant increase, an accelerating increase, or a decelerating increase, respondents were correct 95% versus 16% of the time when viewing linear–linear versus log–log axes (Supplementary Fig. 2).

Each of these fallacies (Table 1) is understandable as a consequence of fundamental mathematical misconceptions. Four of the five fallacies are consistent with the 'linear extrapolation error'[12], the mistaken idea, common among students[12–15], that 'log' can be treated as a variable. Under that misconception, $\ln(y) = \ln(c) + a\ln(x)$ simplifies to $y = c + ax$ because the ln 'cancels', leading to the hand-hold fallacy, the first and third watch out for curves fallacies (all from interpreting the log–log slope a as the linear–linear slope) and Zeno's zero fallacy (interpreting the log–log intercept $\ln(c)$ as the linear–linear intercept). We hypothesize that the second watch out for curves fallacy (interpreting any upward-trending log–log line as accelerating) stems from correctly recalling that log–log lines mean power laws, but mistakenly thinking that all power laws curve upwards in linear–linear space. Because of these ties to fundamental misconceptions about logarithms themselves, we suspect that many readers also have misconceptions about log–linear and

Box 1 | Recommendations based on our results

Recommendations for educators

Ecology curricula should include explicit instruction on how to interpret log-scaled axes and equations. Our study, as well as other research on students' difficulty interpreting logarithms[15], indicates that misinterpretations of log-scaled data are not caused by carelessness, but rather by deeply held misconceptions. Thus, instructors should not merely remind students of the correct use and interpretation of logarithms, but directly combat misconceptions by creating cognitive conflict via refutational teaching. This method has students analyse why the wrong answer is wrong, rather than why the right answer is right. For example, instructors could provide students with an incorrect interpretation of a graph and ask, 'Why is this wrong?'. Research in psychology and science education supports refutational teaching as a means of reducing misconceptions and producing conceptual change[40–43], and it may be particularly relevant here because misconceptions about logarithms are so firmly held.

When to use log-scaled graphs

In general, we recommend that authors use logarithms like telescopes, which are also multiplicative tools. They can be useful, but only for visualizing certain things.

- Consider using log-scaled axes only if they illuminate one or more important features of your results. The papers in our bibliometric analysis probably used log-scaled axes for these reasons. Examples of important features include elasticities[2], conformation to power laws[3], or displaying meaningful variation across multiple orders of magnitude[1]. In these cases, though, authors should weigh the benefits of seeing these features against the costs of confused and annoyed readers. (We say 'annoyed' because 96% of our survey respondents who expressed an opinion always or usually preferred to see linear axes; Supplementary Table 5.)
- Do not use log-scaled axes simply because it takes effort to back-transform data. In this paper we focus on data presentation, not analysis; see refs [44–48] for discussion of log transformations in data analysis. When data are log-transformed for analysis, though, it is not essential to present them on a log scale. Data should only be presented on a log scale if there are good display reasons for doing so (see above). In our survey, 20% of our respondents listed conformation to statistical tests as the only reason they used log-scaled axes (Supplementary Table 6), but the advantage of convincing readers that your statistics are sound might be offset by misconceptions about your results.

How to mitigate misconceptions when you use log-scaled graphs

In our survey, 48% of respondents reported using logarithms to improve readability. In the light of our results it is tempting to make a joke about this, but we agree that there are many situations where logarithms do improve readability of certain features. In these situations we recommend the following steps to mitigate misconceptions.

- Consider also presenting your results in linear space, in additional figures[29], panels or insets. This retains the benefits of log scales while combating the misconceptions.
- Consider plotting the 1:1 line. This makes it easier to see curvature and relative changes, particularly for positive x–y relationships. In our bibliometric analysis only two papers[49,50], representing 1% of the x–y panels with more than one log axis, plotted the 1:1 line.
- Use untransformed values (such as 1, 10, 100) instead of log-transformed values (such as 0, 1, 2) for log-scaled presentation. Untransformed values do not present as much confusion about units[51], and do not depend on the base used for log transformation. They result in no more misconceptions about the shapes of relationships than log-transformed values (Figs. 3–5, Supplementary Figs. 1, 3–6). They are also what most readers prefer: 84% of respondents with an opinion always or usually preferred untransformed values on log-scaled graphs (Supplementary Table 5), even though only 27% of log-scaled axes presented untransformed values (Fig. 1).
- If you display log-transformed values, indicate the base used for log transformation. A telescope that does not list its magnification factor is not a useful scientific tool. In our bibliometric analysis, 63% of axes with log-transformed values did not list the base on the axis or in the caption (Fig. 1). Two bases—*e* and 10—are most common, narrowing the range of likely values, but any ambiguity clouds interpretation. Consider the log response ratio, which is a common[28,52–56] use of logarithms with good justification[57]. When untransformed values are plotted[54], there is no ambiguity about the numbers: 1 means no change, 2 means a factor of 2 increase. When log-transformed values are plotted, however, a value of 1 could mean that a treatment increases a response variable by a factor of approximately 2.7 (if base *e*) or a factor of 10 (if base 10). Compounding this issue, software packages (such as R; ref. [38]) often have a function 'log' that does not specify the base in the function name, so practitioners might not know which base they are using. (In R 'log' defaults to base *e*.)
- If you display log-transformed values, indicate the original units. Original units are essential to understanding log-transformed values, even though log-transformed values themselves are unitless[51]. A log-transformed value of 0 might mean 1 mg or 1 kg. In our bibliometric analysis, 45% of axes with log-transformed values did not report the original units on the axis or in the caption (Fig. 1). Only 31% of axes with log-transformed values reported both the log base and the original units on the axis or in the caption (Fig. 1). Some of these did list the base, the original units, or both somewhere in the text, but many did not. Therefore, the numbers on 69% of axes with log-transformed values are uninterpretable without significant effort on the part of the reader. There is still some information in those graphs—whether a relationship is positive or negative, for instance—but much is lost when the values are uninterpretable.

linear–log displays, although our survey did not test these display types explicitly.

We were curious to see whether ecologists' misunderstandings of logarithms also affected their understanding of equations, and whether they varied across different demographic groups. When respondents viewed equations instead of graphs (compare Supplementary Table 2 to Fig. 2), there were three major differences. First, respondents were much more likely to click 'wouldn't notice' on equations (mean of 24% across linear- and log-scaled equations) compared to graphs (3% across linear- and log-scaled graphs), as well as on log-scaled equations (30%) compared to linear-scaled equations (14%; Supplementary Table 1). Second, linear-scaled equations elicited more incorrect answers than linear-scaled graphs (Figs. 3–5, Supplementary Figs. 1, 3–6, Supplementary Table 3). Third, although misunderstandings of log-scaled equations were common, they were no more common than misunderstandings of log-scaled graphs (Figs. 3–5, Supplementary Figs. 1, 3–6, Supplementary Table 3). Across graphs and equations, demographic (Supplementary Table 4) splits revealed no major differences across gender, highest degree, career stage, or age (Supplementary Table 3). Somewhat surprisingly, respondents' reported levels of comfort with logarithms (Supplementary Table 5) were not predictive of their abilities to correctly interpret log-scaled graphs or equations (or linear-scaled graphs or equations either; Supplementary Fig. 7). This might reflect a Dunning–Kruger effect, where people who do understand logarithms know that they are complicated, and so rate their confidence lower, and many people who believe they understand logarithms well do not.

A side effect of using a survey to assess understanding is that respondents do not have the same context that they would in a real paper. Our results show that without the context of a full paper, linear–linear displays are correctly interpreted much more frequently than log–log displays. Would additional context narrow this gap? The degree to which additional context matters probably depends on the type of context and the level of depth with which readers read the paper. Simply reminding readers that axes are log-scaled, as is common but not ubiquitous (85%, Fig. 1), is unlikely to overcome misconceptions, according to our data. When axes or equations were displayed with the word 'log', the misconceptions were as strong as when 'log' was not displayed. Describing the signs of relationships (increase versus decrease) is also unlikely to help much, given that respondents glean this information from log–log figures and equations without a description (Supplementary Figs. 1, 3). We suspect that describing the nuances of relationships (for example, accelerating versus decelerating) would help. At the other end of the spectrum, displaying the same results in linear–linear scale figures[29] or panels could help overcome misconceptions, and it is precisely this sort of context that we recommend (Box 1).

Conducting a scientific project is hard work, and it is a shame if the outcome of all that work is misinterpreted. Our results suggest that displaying data on logarithmic scales often leads to five predictable fallacies in graphical interpretation by ecologists. Given the frequency of incorrect responses and the regularity with which ecologists are exposed to logarithms, we suspect that confusion about log-scaled data is common among many scientists, not just ecologists. We hope that by becoming aware of the potential pitfalls of presenting log-scaled data, taking steps to overcome these pitfalls and learning more about log-scaled data (see Box 1 for our recommendations), ecologists as well as scientists across many disciplines will present and interpret data more clearly.

Methods

Bibliometric analysis. We examined papers published in *Ecology* in 2015. *Ecology* publishes research across the subdisciplines and topics of ecology, covering a wide array of ecosystems, biomes, structure, function, level of complexity, taxa, approach and scale[30]. These topics are not represented in exactly equal frequencies[30] but the journal gives a broad representation of the field as a whole. The year 2015 was the most recent year that was fully available when we began the study. The foci within *Ecology* have changed over the decades in some ways (for example, the term 'ecosystem' became much more common after 1965), but not all (the distributions of major ecosystem types in the titles have remained relatively constant)[30]. Therefore, our bibliometric analysis is best viewed as assessing the contemporary state of ecology.

Our initial set of papers was research-focused papers, but not editorials, book reviews or commentaries. Of this initial set of 325 papers, our final set (see dataset in Supplementary Information) included only those papers that contained figures, which narrowed the group to 309 papers. These 309 papers included 254 'Articles', 41 'Reports', 6 'Notes', 5 'Concepts & Synthesis' papers and 3 'Centennial' papers.

For each paper, we recorded whether use of logarithms was reported (mentioned in the text or a caption, or displayed on an axis label). For each panel in each figure in each paper, we recorded how many axes there were, whether each axis was numerical versus categorical, whether each axis was log scale versus linear scale versus some other scale, whether the scale was noted, and whether the values on log-scaled axes were untransformed versus log-transformed. Many axes that we recorded as 'linear' incorporated some aspects of logarithms, such as ordinations that might have included log-transformed data, or coefficients from regressions on log-transformed data, so our estimate of the fraction of axes that are log-transformed might be conservative. For axes with log-transformed values, we recorded whether the base of the log transformation and the units of the untransformed data were mentioned on the figure itself (the axis label was the only place we found either piece of information) or in the figure caption. We did not seek to assess whether the authors of the papers in our bibliometric analysis interpreted their results correctly, as our focus was on readers' rather than authors' interpretations.

Survey. We developed a survey, following design advice from ref. [31], using Qualtrics software. The survey informed respondents that our goals were to assess what type of information ecologists glean from different forms of data presentation, and asked them to respond to the survey questions with the same degree of rigour they would apply to reading a paper. To decrease the time it would take to complete the survey, each respondent saw a randomized subset of the available items. The order in which each respondent viewed items (except for the opinion and demographic items, which were always at the end) was also randomized. The full survey, along with a description of its flow, is provided as Supplementary Information.

The survey contained six sections. The first (main) section presented graphs. Each main survey item began with "Consider this figure, which shows the relationship with distance from the edge of a habitat and population size for rabbits (solid) and chipmunks (dashed). In the answers, 'distance' refers to distance from the edge of a habitat (the horizontal axis)." Graphs were displayed below this introductory statement, followed by "For each prompt, please check the box that best completes the sentence. If I were looking at this figure, I would notice that …". The survey then asked respondents for information about each graph or equation and table they viewed: (1) whether the rabbit and chipmunk populations increase, decrease, or sometimes increase and sometimes decrease with distance; (2) whether the rabbit population, compared to the chipmunk population, always changes more steeply, always changes less steeply, always changes similarly steeply, or sometimes changes more steeply but sometimes changes less steeply with distance; (3) whether the rabbit and chipmunk populations are positive, zero or negative at the edge of the habitat; and (4) whether the rabbit and chipmunk populations change with distance in a constant, decelerating or accelerating manner. The option 'wouldn't notice' was available for all questions.

The second section was similar to the first, except that it presented the data as equations and tables instead of graphs. There were three types of equation and table presentation, analogous to the three types of graph (Supplementary Table 2): untransformed equations (analogous to linear-linear graphs), log-transformed equations with untransformed values (analogous to log–log graphs with untransformed values), and log-transformed equations with log-transformed values (analogous to log–log graphs with log-transformed values). Although we consider these to be the most comparable, we note that figures in some papers[18] display log-transformed values in equations but untransformed values in axis labels. The third section presented a series of three graphs and asked respondents to identify which increased in a constant versus decelerating versus accelerating manner. The fourth section, which we do not address here, asked about units and statistical significance.

The fifth section of the survey asked respondents a series of opinion questions about their preferences for viewing linear-scale versus logarithmic-scale axes, their preferences for viewing untransformed values (such as 1, 10 and 100, which in the survey we called 'linear values') or log-transformed values (such as 0, 1 and 2), why they used log-transformed axes and/or values, and about their comfort level with logarithms, the last of which used a Likert-type scale with seven possible answer categories. The sixth and final section requested demographic data.

We obtained Insitutional Review Board approval from Columbia University for our survey, and we have complied with all relevant ethical regulations. Informed consent was obtained from all participants. We distributed this online survey to the e-mail list of the Ecological Society of America (ESA) (about 8,000 addresses), with the help of the ESA office. The initial email was sent on 8 November 2016, and one follow-up email was sent on 21 November 2016. We stopped collecting data on 30 January 2017. A total of 988 ecologists answered at least one question on the survey (an approximately 12% response rate), 623 of which completed it (an approximately 8% response rate from the full list, 63% completion rate), for a response rate that is similar to other survey studies in the environmental literature (those in refs [32–36] range from 5–15%). The racial makeup of our survey respondents (0.5% Native American or Native Alaskan, 1% Black or African American, 4% Asian, 5% Hispanic, 89% White, among those who responded; Supplementary Table 4) was similar to the racial makeup of the ESA as a whole in 2005 (0.4% Native American, 1% Black, 5% Asian, 4% Hispanic, 89% White, among those who responded)[37]. The gender makeup of our survey respondents was more balanced (45% female; Supplementary Table 4) than the ESA as a whole was in 2003 (26% female)[37].

The respondents represented a range of highest degree and career stage, although most (69%) had a PhD. By profession, 35% of respondents were professors; 21% were research scientists in government, academia or industry; 10% were postdocs; 18% were PhD students; 4% were retired or emeritus; 3% were MA or MSc students; 1% were undergrads; and 2% were consultants. Although there are probably some differences between the makeup of our respondents and the ESA as a whole in terms of gender, career stage and age, we did not make any demographic adjustments in our analyses because there were no discernible differences across demographic groups in our results (Supplementary Table 3). Because each respondent saw a randomized subset of the items, sample sizes were typically in the low hundreds. All data preparation, plotting and analysis was done in R (ref. [38]). Our survey dataset is available as Supplementary Information, but with demographic information removed to maintain respondents' privacy.

Survey validity. Surveys can be subject to a number of forms of error, but our survey results and information from pilot rounds of the survey indicate that any such errors were minor. Our survey had a high enough response rate that low sample size was not an issue. Our results indicate that respondents took the survey seriously. The survey items that asked whether the rabbit and chipmunk populations increased, decreased, or sometimes increased and sometimes decreased with distance from the edge of the habitat (Supplementary Fig. 1) were the clearest survey item. When presented with this question, 98.6% of respondents correctly identified that the populations increased with distance from the edge, suggesting that respondents took the survey seriously.

Misinterpretation is a trickier potential source of bias to assess. The fact that 98.6% of answers to the clearest survey item were correct suggests that the overall setup of the questions was not misinterpreted, but it is possible that respondents interpreted other survey items differently from the way we intended. For our research questions here, misinterpretation of questions would only lead to bias if it applied differently to linear–linear scales versus log–log scales. We worded the questions identically for linear–linear versus log–log scales to minimize bias, but it remains possible that a question could be interpreted differently when viewing a graph on a log scale than on a linear scale.

The most likely possible bias is that, when presented with a linear-scaled graph, respondents might interpret a question as applying to the population itself, whereas when presented with a log-scaled graph, respondents might interpret the same question as applying to the log of a population, rather than the population itself. Consider the survey item "If I were looking at this figure, I would notice that the rabbit population changes with distance in …", where the response options were "a constant manner", "an accelerating manner", "a decelerating manner" and "wouldn't notice." Although some respondents might have interpreted the questions as applying to the log of populations, we claim that these are a minority of the 'incorrect' responses for two reasons.

The first reason is that feedback from our pilot of the survey directly confirmed that respondents were thinking about populations, rather than the logarithms of populations, albeit with a small sample size. In the pilot we asked five respondents to think aloud[39] as they were answering the questions. Their narration suggested that most of them were thinking of the populations, not the logs of populations. For example, their narration included the statements "Both increase. I don't think the log transformation will change that", "The solid population will change less with *x*. That wouldn't change with log transformation", "The linear scale one was easier because you don't need to think about log transformation", "The one that might seem linear was log–log" and "I don't really like log scale or log transformations because they're so much more difficult to interpret than regular numbers". In contrast, only one respondent said something that indicated s/he was not sure: "Asking about the population itself or the log of it?". Despite this initial uncertainty, this respondent ultimately decided to answer about the population itself.

The second reason is based on our survey results. If respondents thought the question was asking about the log of the population, they would respond to our fourth question (accelerating versus decelerating versus constant) by saying that straight lines on log–log graphs (Fig. 2b,c,e and f) are changing in a 'constant' manner. A large number of responses were 'accelerating' (Fig. 5b,c,e and f), even for relationships that were straight lines in both linear–linear and log–log space. These respondents were clearly thinking about population sizes, not the log of population sizes. There were also a large number of 'constant' responses (Fig. 5b,c,e and f), which could be either misinterpretations or misconceptions. Our pilot feedback suggested that misinterpretations were rare.

Beyond sample size, taking the survey seriously, and misinterpretation, another potential bias is that respondents might have Googled, used computational software, or otherwise gone beyond what they would typically do while reading a paper, in an effort to 'get the right answer'. This bias would lessen the evidence for misconceptions, and therefore would strengthen our conclusions that misconceptions are common. Similarly, non-responder bias, where potential respondents declined to take or finish the survey, seems more likely to affect those who are less comfortable with or less interested in logarithms. Our results suggest that there is not a strong relationship between comfort level and whether responses are correct, but if anything, such non-responder bias would be more likely to strengthen than to weaken our conclusions.

Data availability. All data analysed during the current study are included in this Article (and its Supplementary Information), with the following exception. Demographic information for survey respondents has been removed from the posted dataset because it could compromise research participant privacy.

Received: 18 October 2017; Accepted: 18 June 2018;
Published online: 16 July 2018

References

1. Vitousek, P. M. *Nutrient Cycling and Limitation: Hawai'i as a Model Ecosystem* (Princeton Univ. Press, Princeton, 2004).
2. Benton, T. G. & Grant, A. Elasticity analysis as an important tool in evolutionary and population ecology. *Trends Ecol. Evol.* **14**, 467–471 (1999).
3. Taubert, F. et al. Global patterns of tropical forest fragmentation. *Nature* **554**, 519–522 (2018).
4. Covington, A. K., Bates, R. G. & Durst, R. A. Definition of pH scales, standard reference values, measurement of pH and related terminology. *Pure Appl. Chem.* **57**, 531–542 (1985).
5. Richter, C. F. An instrumental earthquake magnitude scale. *Bull. Seismol. Soc. Am.* **25**, 1–32 (1935).
6. Pogson, N. Magnitudes of thirty-six of the minor planets for the first day of each month of the year 1857. *Mon. Not. R. Astron. Soc.* **17**, 12–15 (1856).
7. Cowell, F. A. *Measuring Inequality* (Oxford Univ. Press, New York, 2011).
8. Hastings, A. *Population Biology: Concepts and Models* (Springer, New York, 1997).
9. Krane, K. S. & Halliday, D. *Introductory Nuclear Physics* (Wiley, New York, 1988).
10. Bolker, B. M. *Ecological Models and Data in R* (Princeton Univ. Press, Princeton, 2008).
11. Sokal, R. R. & Rohlf, F. J. *Biometry* 3rd edn (W. H. Freeman, New York, 1995).
12. Matz, M. Towards a computational theory of algebraic competence. *J. Math. Behav.* **3**, 93–166 (1980).
13. Kaur, B. & Boey, H. P. S. Algebraic misconceptions of first year college students. *Focus Learn. Probl. Math.* **16**, 43–58 (1994).
14. Yen, R. Reflections on higher school certificate examinations: learning from their mistakes, High School Certificate 1998. *Reflections* **24**, 3–8 (1999).
15. Liang, C. B. & Wood, E. Working with logarithms: students' misconceptions and errors. *Math. Educ.* **8**, 53–70 (2005).
16. Preston, F. W. The canonical distribution of commonness and rarity: part I. *Ecology* **43**, 185–215 (1962).
17. Peters, R. H. *The Ecological Implications of Body Size* (Cambridge Univ. Press, Cambridge, 1986).
18. Rizzuto, M., Carbone, C. & Pawar, S. Foraging constraints reverse the scaling of activity time in carnivores. *Nat. Ecol. Evol.* **2**, 247–253 (2018).
19. Smith, B. & Wilson, J. B. A consumer's guide to evenness indices. *Oikos* **76**, 70–82 (1996).
20. Wright, I. J. et al. The worldwide leaf economics spectrum. *Nature* **428**, 821–827 (2004).
21. Zhou, X. H. & Tu, W. Confidence intervals for the mean of diagnostic test charge data containing zeros. *Biometrics* **56**, 1118–1125 (2000).
22. Tian, L. & Wu, J. Confidence intervals for the mean of lognormal data with excess zeros. *Biom. J.* **48**, 149–156 (2006).
23. Goodale, C. L. et al. Soil processes drive seasonal variation in retention of ^{15}N tracers in a deciduous forest catchment. *Ecology* **96**, 2653–2668 (2015).
24. Atwater, D. Z. & Callaway, R. M. Testing the mechanisms of diversity-dependent overyielding in a grass species. *Ecology* **96**, 3332–3342 (2015).
25. Taylor, P. G. et al. Organic forms dominate hydrologic nitrogen export from a lowland tropical watershed. *Ecology* **96**, 1229–1241 (2015).
26. McConkey, K. R., Brockelman, W. Y., Saralamba, C., & Nathalang, A. Effectiveness of primate seed dispersers for an "oversized" fruit *Garcinia benthamii*. *Ecology* **96**, 2737–2747 (2015).
27. Wallace, J. B., Eggert, S. L., Meyer, J. L. & Webster, J. R. Stream invertebrate productivity linked to forest subsidies: 37 stream-years of reference and experimental data. *Ecology* **96**, 1213–1228 (2015).
28. Albertson, L. K. & Allen, D. C. Meta-analysis: abundance, behavior, and hydraulic energy shape biotic effects on sediment transport in streams. *Ecology* **96**, 1329–1339 (2015).
29. Carey, J. C. et al. Temperature responses of soil respiration largely unaltered with experimental warming. *Proc. Natl Acad. Sci. USA* **113**, 13797–13802 (2016).
30. Gorham, E. & Kelly, J. A history of ecological research derived from titles of articles in the journal "*Ecology*," 1925–2015. *Bull. Ecol. Soc. Am.* **99**, 61–72 (2018).
31. Krosnick, J. A. & Presser, S. in *Handbook of Survey Research* 2nd edn (eds Marsden, P. V. & Wright, J. D.) 263–314 (Emerald Group, Bingley, 2010).
32. Lange, A., Vogt, C. & Ziegler, A. On the importance of equity in international climate policy: an empirical analysis. *Energy Econ.* **29**, 545–562 (2007).
33. Dannenberg, A., Sturm, B. & Vogt, C. Do equity preferences matter for climate negotiators? An experimental investigation. *Environ. Resour. Econ.* **47**, 91–109 (2010).
34. Lange, A., Löschel, A., Vogt, C. & Ziegler, A. On the self-interested use of equity in international climate negotiations. *Eur. Econ. Rev.* **54**, 359–375 (2010).
35. Kesternich, M. Minimum participation rules in international environmental agreements: empirical evidence from a survey among delegates in international climate negotiations. *Appl. Econ.* **48**, 1047–1065 (2016).

36. Dannenberg, A., Zitzelsberger, S. & Tavoni, A. Climate negotiators' and scientists' assessments of the climate negotiations. *Nat. Clim. Change* **7**, 437–443 (2017).
37. Ortega, S. et al. *Women and Minorities in Ecology II* (Ecological Society of America, 2006); https://www.esa.org/esa/wp-content/uploads/2012/12/wamieReport2006.pdf
38. R Development Core Team *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, 2013).
39. van Someren, M. W., Barnard, Y. F. & Sandberg, J. A. C. *The Think Aloud Method: A Practical Approach to Modelling Cognitive Processes* (Academic, London, 1994).
40. Tippett, C. D. Refutation text in science education: a review of two decades of research. *Int. J. Sci. Math. Educ.* **8**, 951–970 (2010).
41. Sinatra, G. M. & Broughton, S. H. Bridging reading comprehension and conceptual change in science education: the promise of refutation text. *Read. Res. Q.* **46**, 374–393 (2011).
42. Lassonde, K. A., Kendeou, P. & O'Brien, E. J. Refutation texts: overcoming psychology misconceptions that are resistant to change. *Scholarsh. Teach. Learn. Psychol.* **2**, 62–74 (2016).
43. Kowalski, P. & Taylor, A. K. Reducing students' misconceptions with refutational teaching: for long-term retention, comprehension matters. *Scholarsh. Teach. Learn. Psychol.* **3**, 90–100 (2017).
44. O'Hara, R. B. & Kotze, D. J. Do not log-transform count data. *Methods Ecol. Evol.* **1**, 118–122 (2010).
45. Wilson, J. B. Priorities in statistics, the sensitive feet of elephants, and don't transform data. *Folia Geobot.* **42**, 161–167 (2007).
46. Feng, C. et al. log transformation: application and interpretation in biomedical research. *Stat. Med.* **32**, 230–239 (2013).
47. Ives, A. R. For testing the significance of regression coefficients, go ahead and log-transform count data. *Methods Ecol. Evol.* **6**, 828–835 (2015).
48. Warton, D. I., Lyons, M., Stoklosa, J. & Ives, A. R. Three points to consider when choosing a LM or GLM test for count data. *Methods Ecol. Evol.* **7**, 882–890 (2016).
49. Neuheimer, A. B. et al. Adult and offspring size in the ocean over 17 orders of magnitude follows two life history strategies. *Ecology* **96**, 3303–3311 (2015).
50. Luo, Z., Wang, E. & Smith, C. Fresh carbon input differentially impacts soil carbon decomposition across natural and managed systems. *Ecology* **96**, 2806–2813 (2015).
51. Matta, C. F., Massa, L., Gubskaya, A. V. & Knoll, E. Can one take the logarithm or the sine of a dimensional quantity or a unit? Dimensional analysis involving transcendental functions. *J. Chem. Educ.* **88**, 67–70 (2011).
52. Deng, Q. et al. Down-regulation of tissue N:P ratios in terrestrial plants by elevated CO_2. *Ecology* **96**, 3354–3362 (2015).
53. Kempel, A. et al. Herbivore preference drives plant community composition. *Ecology* **96**, 2923–2934 (2015).
54. LeBauer, D. S. & Treseder, K. K. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology* **89**, 371–379 (2008).
55. Yu, Q. et al. Stoichiometric homeostasis predicts plant species dominance, temporal stability, and responses to global change. *Ecology* **96**, 2328–2335 (2015).
56. Yuan, Z. Y. & Chen, H. Y. H. Negative effects of fertilization on plant nutrient resorption. *Ecology* **96**, 373–380 (2015).
57. Hedges, L. R., Gurevitch, J. & Curtis, P. S. The meta-analysis of response ratios in experimental ecology. *Ecology* **80**, 1150–1156 (1999).

Acknowledgements

We thank all the respondents whose efforts created these data, the volunteers who participated in pilot surveys and K. McCarter for helping to distribute the survey to the Ecological Society of America member list. Feedback from the Menge laboratory and J. Lubchenco improved this work.

Author contributions

All authors conceived of the project. D.N.L.M. and A.C.M. designed the survey. T.A.B., A.W.Q., N.B.S., B.N.T. A.A.W. and D.N.L.M. conducted the bibliometric analysis. D.N.L.M. analysed data and wrote the paper. All authors edited the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41559-018-0610-7.

Correspondence and requests for materials should be addressed to D.N.L.M.

nature research

Corresponding author(s): Duncan Menge

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistical parameters

When statistical analyses are reported, confirm that the following items are present in the relevant location (e.g. figure legend, table legend, main text, or Methods section).

n/a	Confirmed	
☐	☒	The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	An indication of whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistics including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated
☒	☐	Clearly defined error bars *State explicitly what error bars represent (e.g. SD, SE, CI)*

Our web collection on statistics for biologists may be useful.

Software and code

Policy information about availability of computer code

Data collection	We used qualtrics to collect our survey data.
Data analysis	We used R to subset, plot, and analyze the data.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data analyzed during the current study are included in this published article (and its supplementary files), with the following exception. Demographic information for survey respondents have been scrubbed from the posted dataset because they could compromise research participant privacy.

Field-specific reporting

Please select the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/authors/policies/ReportingSummary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We conducted a quantitative bibliometric analysis of all papers published in the journal Ecology in 2015 and a quantitative survey of ecologists. The survey asked respondents to evaluate a series of graphs and equations, to describe their level of comfort with logarithms, to express a series of preferences about data presentation, and to explain when and why they presented data on logarithmic scales.
Research sample	Given that our questions concerned ecologists, we focused on ecologists. Our sample was taken from the Ecological Society of America's email list: The survey was sent to the entire list, 12% (988) of the list responded, and 63% (623) of those who responded completed the survey. The demographic makeup of our respondents is detailed in Supplementary Table 4. The racial makeup of our sample is very similar to ESA as a whole, whereas the gender makeup of our sample (45% female, 55% male, <1% other) is more balanced than ESA as a whole.
Sampling strategy	The survey was emailed to the entire list of the Ecological Society of America. 12% (988) of the list responded, and 63% (623) of those who responded completed the survey. Because our study was a survey, it was not possible to predict our sample size exactly. The population of interest was all ecologists, so we determined that sending the survey to all ecologists on the Ecological Society of America email list (which has ~8000 members) would be sufficient. The response rate (around 10%) was more than sufficient to make the comparisons we wanted to make.
Data collection	We used Qualtrics software for data collection. We do not know the identity of the respondents, only the demographic information they provide.
Timing	ESA sent out the first email on November 8, 2016, and sent a reminder email on November 21, 2016. We stopped collecting data on January 30, 2017.
Data exclusions	We did not exclude any data.
Non-participation	The survey was sent to the entire ESA email list (~8000 addresses), 12% (988) of the list responded, and 63% (623) of those who responded completed the survey.
Randomization	Respondents were allocated into different experimental randomly, using the randomization capabilities within the Qualtrics software.

Reporting for specific materials, systems and methods

Materials & experimental systems

n/a	Involved in the study
☒	☐ Unique biological materials
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	Our survey involved human research participants. We obtained IRB approval from Columbia University. We have included a table of demographic information as a Supplementary Table. A few of the relevant demographics: age ranged from 20-84, with a mean of 43 years old. The sample was 45% female, 55% male, and <1% transgender or other.
Recruitment	The population of interest was all ecologists, so we determined that sending the survey to all ecologists on the Ecological Society

Recruitment

of America email list (which has ~8000 members) would be sufficient. The response rate (623 completed the survey) was more than sufficient to make the comparison we wanted to make. It is possible that the subset of the entire list that took the survey was biased, but our guess is that such bias would make our findings more conservative. For instance, if there was a bias toward people who enjoyed math problems about logarithms, the entire population would likely show even lower understanding of logarithms than we found.