Ecology Letters

WILEY

ECOLOGY LETTERS

SYNTHESIS

Converting and Constructing Effect Sizes With the Response Ratio

Marc J. Lajeunesse

Department of Integrative Biology, University of South Florida, Tampa, Florida, USA

Correspondence: Marc J. Lajeunesse (lajeunesse@usf.edu)

Received: 28 August 2025 | **Revised:** 26 January 2026 | **Accepted:** 29 January 2026

Editor: Alfredo Sánchez-Tójar

Keywords: A/B testing | common currency | constructable effect sizes | geometric mean theorem | research synthesis | secondary analysis

ABSTRACT

Data synthesised and published as response ratios in ecology (lnRR, or ratio of means, RoM) remain isolated from broad secondary analyses because they cannot be converted to other effect size metrics. Here I address this lack of data interoperability by developing a conversion to the widely used Hedges' *d* (standardised mean difference, SMD). This conversion is practical and near exact—as long as assumptions of homogeneity of variances are met, Hedges' *g* correction is used to adjust for small-sample bias, and only additive and not multiplicative ecological processes are converted. I then generalise this conversion with abstract algebra to develop additional opportunities to reuse effect sizes—first by stating the response ratio as a geometric construction of Pythagorean means, and then *d* as a proportional compass-and-straightedge construction of the response ratio. Constructability is a new pathway of interoperability for effect sizes, and without collecting new data, allows for the response ratio and *d* to be repurposed into relative change datatypes such as the arithmetic, harmonic, geometric, quadratic and logarithmic means. Much of what has been synthesised in ecology is only available as response ratios, and I hope these conversions increase their value post-publication and facilitate reuse for bolder, more comprehensive meta-analyses.

1 | Introduction

Common effect size metrics like the standardised mean difference (Hedges' *d*) and the correlation coefficient (*r*) are highly reusable because they can be extracted from different published meta-analyses, converted to a common currency (Arnqvist and Wooster 1995; Pustejovsky 2014), and analysed with a single, cohesive synthesis (e.g., Bergquist et al. 2023). This data interoperability is possible because both *d* and *r* are statistical expressions of the linear model (Rosenthal 1994; but see McGrath and Meyer 2006). However, ecologists have a strong culture of research synthesis with the response ratio (lnRR, or the ratio of means; Hedges et al. 1999), with many celebrated applications at all ecological scales (e.g., Cardinale et al. 2006; Parker et al. 2006; Elser et al. 2007; Hoeksema et al. 2010). But unlike *d* or *r*, there are no conversions for the response ratio or even a clear link with the linear model to help guide interoperability (Lajeunesse 2011).

The lack of interoperability has been fatal for reuse. Early attempts to broadly harvest effect sizes from ecological meta-analyses excluded those published as response ratios due to a lack of conversions (e.g., Møller and Jennions 2002). More recent secondary analyses, sometimes synthesising hundreds of thousands of published ecological studies, were limited to apples-to-apples comparisons because a common currency between *d*, *r* and response ratios could not be achieved (e.g., Hillebrand and Gurevitch 2014; Yang et al. 2022, 2023; Fox 2022; Costello and Fox 2022). When effect sizes are excluded or isolated this way, ecologists are significantly constrained in what they can achieve with research synthesis. It fragments knowledge, preventing holistic generalisations of ecological phenomena (Lajeunesse 2010; Spake et al. 2022). It erodes both the statistical power of secondary analyses and the robust decomposition of multiple sources of heterogeneity by analysing data separately rather than collectively (Lajeunesse 2013a; Hillebrand and Gurevitch 2014). More practically, a divided literature limits the clear and comprehensive views needed for informing policy (Haddaway and Pullin 2014). Given that nearly half of everything synthesised in ecology is available only as response ratios (Nakagawa and Santos 2012), and that reusing published effect sizes is perhaps our best shot to unify ecology into a single, cohesive synthesis, then addressing interoperability is critical to achieving domain-level understanding of the vast and diverse study outcomes published by ecologists (e.g., Barto and Rillig 2012; Senior et al. 2016; Bebout and Fox 2024).

Here I address the data fragmentation caused by the widespread use of response ratios by developing a practical conversion to the broadly used, cross-disciplinary Hedges' d effect size metric (Hedges 1981, 1982). This conversion is the missing link to full interoperability, as d has established relationships and conversions to all major effect size metrics (e.g., r, Odd's ratios; see Lipsey and Wilson 2001; Lajeunesse 2013b). I begin with a walkthrough of the conversion using algebra—it offers many new insights about the response ratio and d. I follow with Monte Carlo simulations, to assess sensitivity to assumption violations, and a worked example of converting published effect size data. I then place this conversion within a broader geometric framework to formalise a new type of data interoperability for mean-based effect sizes, focusing on their compass-and-straightedge constructability (sensu Wantzel 1837; see Box 1). With this framework, I then show how to repurpose effect sizes into a broad diversity of relative change datatypes (Törnqvist et al. 1985), and how the response ratio itself can be used to repurpose d. Finally, although these conversions offer opportunities to unify previously isolated datasets (such as the data used by Hillebrand and Gurevitch 2014), there are also many opportunities for misuse. Therefore, rather than focusing on how to analyse unified effect size data (which requires no new methodology beyond conventional meta-analysis), I end with a discussion on the diversity of models needed to support a common currency of effect sizes, while emphasising how these models need careful attention to avoid misrepresenting the outcomes of ecological studies. Ultimately, the value of response ratios will not be realised until interoperability is addressed, and I hope my conversions offer a fresh start to revisit, reuse, or even repurpose these valuable data.

2 | Algebraic Approach for Converting Response Ratios

Conventional approaches to converting effect sizes typically leverage a statistical model to establish a relationship between metrics. For example, d's can be converted to Odd's ratios (OR) with a scale transformation because a Normal distribution can be approximated with a Logistic distribution ($\ln\mathrm{OR} \approx d\left[\pi / \sqrt{3}\right]$; Chinn 2000). Similarly, d can be converted to a (point-biserial) correlation coefficient because d and r are expressions of a bivariate linear model ($r \approx d / \sqrt{d^2 + 4}$; Cohen 1988; Pustejovsky 2014). Conversions are not exact but can be accurate when effects have large, balanced sample sizes (Poom and af Wåhlberg 2022). However, similar conversion approaches are not possible with the response ratio because the metric is not derived from a test statistic (as d is to a t-test, or as lnOR is to a chi-squared test), nor a conventional statistical model (as r is to linear regression). Therefore, here I outline how to convert the response ratio to d algebraically, and in the succeeding section, I outline a geometric framework to explain why it works as a non-statistical conversion. Finally, recently Murad et al. (2019) suggested single-imputation via regression as an approach to convert response ratios to d. I therefore compare this method with my conversion to assess the validity of these practices.

Before proceeding, there are two risks to keep in mind when converting response ratios. First, response ratios are sometimes chosen explicitly to model multiplicative effects—such as when variability scales proportionally with the mean, or equivalently, when treatment effects are proportional to baseline levels (e.g., effects of elevated CO_2 on plant biomass; Hedges et al. 1999). Therefore, when converting response ratios to d, there is the risk of misrepresenting the proportional nature of the underlying ecological processes with additive (linear scale) effects (as modelled by d; also see Discussion: Conceptual models). Second, the variance and sample sizes of the response ratio are needed to convert to d. Fortunately, these are often reported in published meta-analyses because they are used to weight effects (Hedges and Olkin 1985). However, because there are multiple ways to calculate the variance of response ratios, each with different assumptions (see Hedges et al. 1999), careful attention is needed to identify which calculation method was used prior to conversions to help understand the risk of bias when converting effects (see Best practices for converting response ratios).

2.1 | Working Definitions of lnRR and d

Mean-based effect sizes like the response ratio (lnRR; Hedges et al. 1999; Friedrich et al. 2008) and d (Hedges 1981) quantify the direction and magnitude of study outcomes by estimating a change between a treatment (T) and control (C) mean ($\overline{X}$). Changes between $\overline{X}_\mathrm{T}$ and $\overline{X}_\mathrm{C}$ are stated either as a natural log-transformed ratio of means:

$$\ln\mathrm{RR} = \ln\left(\overline{X}_\mathrm{T} / \overline{X}_\mathrm{C}\right) \tag{1}$$

or as a standardised mean difference (SMD), as in:

$$\text{Hedges}'d = \left(\overline{X}_\mathrm{T} - \overline{X}_\mathrm{C}\right) / s \tag{2}$$

where s is the sample size (N) weighted pooled standard deviations (SD) of $\overline{X}_\mathrm{T}$ and $\overline{X}_\mathrm{C}$:

$$s = \sqrt{\frac{\left(N_\mathrm{T} - 1\right)\left(\mathrm{SD_T}\right)^2 + \left(N_\mathrm{C} - 1\right)\left(\mathrm{SD_C}\right)^2}{N_\mathrm{T} + N_\mathrm{C} - 2}} \tag{3}$$

When sample sizes are low ($N < 15$), both metrics poorly estimate the expected effect and therefore small-sample corrections are needed to minimise bias. Lajeunesse (2015) reported a delta-method (Δ) correction for the response ratio:

$$\ln\mathrm{RR}^\Delta = \ln\mathrm{RR} + {}^1\!/_2\left[\frac{\left(\mathrm{SD}_T\right)^2}{N_\mathrm{T}\overline{X}_\mathrm{T}^2} - \frac{\left(\mathrm{SD}_C\right)^2}{N_\mathrm{C}\overline{X}_\mathrm{C}^2}\right] \tag{4}$$

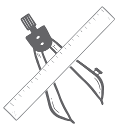
and Hedges (1981) developed a correction often referred to as:

BOX 1 | Constructability as a model to understand effect sizes.

Using the principles of constructability from abstract algebra, effect size metrics like the response ratio (RR) and Hedges' d can be visualised geometrically as classical compass-and-straightedge constructions (see Figures 3–5). Visualising the geometry behind the algebra of effect sizes is a way to model links between metrics and other measures of relative change (Table 1). Below are key terms, tools and rules of the method.

Abstract algebra. A branch of pure mathematics that aims to understand algebraic structures. Applications include modelling the algebra of networks and phylogenetic tree topologies (see Macauley and Youngs 2020). I borrow methods here to model, or more specifically geometrically 'construct', the algebraic structure of effect sizes.

Constructability. A set of simple tools and rules for constructing numbers with line segments—such as the expected magnitude and direction of an effect size given the **algebraic operations** within a metric. These include the quotient of the response ratio ($\overline{X}_T / \overline{X}_C$) and mean difference of Hedges' d ($\overline{X}_T - \overline{X}_C$). Establishing constructability helps prove that effect size metrics are linked geometrically and have proportional effect sizes (Figures 3 and 5), despite using different algebraic operations to define effects.



Tools of constructability. Constructability can be established if the algebra behind an effect size metric can be defined geometrically using only circles and line segments drawn by a ***compass*** and ***straightedge***. (Just like the classic Greek mathematicians!).

Basic rules of constructability. First, draw a line segment of unit length (i.e., length of one). This forms the basis of the construction and ensures a common scale with other constructions. Second, define new line segments with lengths equalling the raw study parameters used to compute an effect size—such as the $\overline{X}_C$ and $\overline{X}_T$ used in the response ratio (Equation 1). The next step is to draw these lengths in ways that **encode algebraic operations** with the rule that any additional lengths must be obtained by intersections (i.e., points where lines or circles meet) drawn with a ***compass*** or ***straightedge***. For example, starting from the unit length, draw a line segment of length $\overline{X}_C$, then starting from the unit length again, draw a parallel length of $\overline{X}_T$. The segment distance between the endpoints of $\overline{X}_C$ and $\overline{X}_T$ encodes subtraction as a geometric construct ($\overline{X}_T - \overline{X}_C$). An illustration of geometric subtraction is found along the height of the triangle in Figure 3c.

Advanced rules of constructability. Intersecting lines and circles drawn with a ***compass*** and ***straightedge*** can often create ***triangular constructions***—these have additional geometric properties useful to **encode operations** like quotients (Figure 3), multiplications and square roots (Figure 4). For example, in a right triangle, the height to the hypotenuse is the geometric mean of the two segments (Table 1). This property is known as the ***Geometric Mean Theorem*** and can be used to construct square root operations. Further, given that right triangles can be constructed on a circle because any angle within a semi-circle is a right angle (Figures 4 and 5), this property is known as ***Thales' theorem*** which helps encode **proportionality** (i.e., lengths that maintain a specific ratio relative to other lengths) and **similarity** (i.e., triangles with different lengths but with equal corresponding angles).

Pythagorean means. These are established ***triangular constructions*** that emerge when the ***Geometric Mean Theorem*** and ***Thales' theorem*** are combined under a single geometric framework—these include arithmetic, geometric and harmonic means (see diversity of measures of relative change in Table 1). These means are interconnected through **proportionality** and **similarity**, and within the context of constructing effect sizes, their interconnectedness helps construct the geometric scaling needed to model the proportionality between the response ratio and d (see Figure 5).

$$\text{Hedges}'g = d\left(1 - \frac{3}{4\left[N_T + N_C\right] - 9}\right) \qquad (5)$$

These corrections are important and should be used, but in the context of establishing a conversion between lnRR and d, they do not offer insight to this goal. I therefore focus on uncorrected versions (Equations 1 and 2) with the caveat that later I assess if corrections are needed. Finally, numerical values of lnRR are rarely interpreted directly but converted to percent (%) change (Pustejovsky 2015, 2018):

$$\%\,\text{change} = 100 \times \left(e^{\ln RR} - 1\right) \qquad (6)$$

2.2 | lnRR and *d* Estimate a Common Effect

Here I show that lnRR and d can be stated to have a common effect as the mean difference (MD) or $\overline{X}_T - \overline{X}_C$. Returning to the two metrics:

$$\ln RR = \ln\left(\overline{X}_T / \overline{X}_C\right) \text{ and } d = \left(\overline{X}_T - \overline{X}_C\right) / s$$

and assuming that $\overline{X}_C$ represents the unmanipulated, baseline for comparison and $\overline{X}_T$ the treatment or manipulated group. Operationally, and following the definition of a causal effect by Rubin (1974), the treatment mean can be decomposed as a mean effect (E) plus the control group or $\overline{X}_T = \overline{X}_E + \overline{X}_C$. Revising each metric given this decomposition:

$$\ln RR = \ln\left(\frac{\overline{X}_E + \overline{X}_C}{\overline{X}_C}\right) \text{ and } d = \frac{\left(\overline{X}_E + \overline{X}_C\right) - \overline{X}_C}{s}$$

and then isolating the mean effect in each, gives:

$$\overline{X}_E = \overline{X}_C e^{\ln RR} - \overline{X}_C \text{ and } \overline{X}_E = ds$$

Finally, after reintroducing the effect size metrics of Equations (1) and (2) back into these $\overline{X}_E$, the underlying effect of both ln RR and d can be restated as a mean difference:

$$\overline{X}_E = \overline{X}_C e^{\ln RR} - \overline{X}_C = \overline{X}_C e^{\ln(\overline{X}_T/\overline{X}_C)} - \overline{X}_C = \overline{X}_T - \overline{X}_C = \mathrm{MD}$$

and

$$\overline{X}_E = ds = \left(\left[\overline{X}_T - \overline{X}_C\right]/s\right)s = \overline{X}_T - \overline{X}_C = \mathrm{MD}$$

2.3 | Common Effect but Different Scales

The common effect between lnRR and d as a mean difference ($\mathrm{MD} = \overline{X}_T - \overline{X}_C$) can be used to establish a conversion. Rather than using lnRR (Equation 1), I use an equivalent form that borrows the MD used when converting lnRR to percent change (Equation 6):

$$\mathrm{lnRR} = \ln\left(\frac{\mathrm{MD}}{\overline{X}_C} + 1\right) = \ln\left(\frac{\overline{X}_T - \overline{X}_C}{\overline{X}_C} + 1\right) \text{ and } d = \frac{\overline{X}_T - \overline{X}_C}{s}$$

Given this common effect, $\overline{X}_T - \overline{X}_C$ can now be isolated in each metric:

$$\overline{X}_T - \overline{X}_C = \left(e^{\mathrm{lnRR}} - 1\right)\overline{X}_C \text{ and } \overline{X}_T - \overline{X}_C = ds$$

And then combined and rearranged to state each in terms of one another as follows:

$$\mathrm{lnRR} = \ln\left(d\frac{s}{\overline{X}_C} + 1\right) \text{ and } d = \left(e^{\mathrm{lnRR}} - 1\right)\left(\frac{s}{\overline{X}_C}\right)^{-1}$$

Two insights about scaling emerge from these conversions. First, the natural log (ln) does not introduce scaling dependence with d, and thus conversions can be simplified and exponentiated into the raw ratio of means ($e^{\mathrm{lnRR}} = \mathrm{RR}$) prior to conversions:

$$\mathrm{RR} = d\frac{s}{\overline{X}_C} + 1 \tag{7}$$

and

$$d = (\mathrm{RR} - 1)\left(\frac{s}{\overline{X}_C}\right)^{-1} \tag{8}$$

Note that exponentiating lnRR this way does not introduce bias to the transformed effect size, this is only a problem when a pooled lnRR (meta-analysis average of multiple log ratios) is back-transformed (a bias known as Jensen's inequality; Rao 1973).

Second, the ratio of the pooled standard deviation to the control mean, or $s/\overline{X}_C$, is crucial for converting the ratio scales (sometimes referred to as standardizers) of each metric: $1/\overline{X}_C$ for RR and $1/s$ for d. This $s/\overline{X}_C$ is a variant of the coefficient of variation that assumes that the standard deviations of $\overline{X}_T$ and $\overline{X}_C$ are homogeneous (equal) and are pooled with a sample-size weighted average (Hedges 1982; Nakagawa et al. 2023). Pooling achieves a more precise estimate of the standard deviation but yields a less performing metric should the treatment and control standard deviations differ (Delacre et al. 2021). Finally, $\mathrm{RR} - 1$ will equal d when $\overline{X}_C = s$.

2.4 | Approximating Scales to Establish a Conversion

Conversion Equations (7) and (8) are exact if $s/\overline{X}_C$ is known. However, assuming that raw study parameters like means ($\overline{X}$) and standard deviations (SD) are unavailable (e.g., not reported in a published meta-analysis), then it is still possible to approximate $s/\overline{X}_C$ with lnRR, its variance used as weights for meta-analysis, and the sample sizes for the treatment (N_T) and control groups (N_C). For this approximation to work, a key assumption about the homogeneity of variances needs to be made. Since d assumes that the *SD*'s of $\overline{X}_T$ and $\overline{X}_C$ are homogeneous (Hedges 1981), but the conventional and most used form of lnRR does not (Hedges et al. 1999; or see Best practices for converting response ratios), then we need to coerce this assumption onto lnRR for a conversion to d. Consequently, the conversion will no longer be exact when $\mathrm{SD}_T \neq \mathrm{SD}_C$.

Starting with the assumption that $\mathrm{SD}_T = \mathrm{SD}_C$, and that the standard deviations in the variance (var) of lnRR can be substituted with their pooled weighted average (s) from Equation (3); assumptions also applied in Hedges et al. (1999):

$$\mathrm{var(lnRR)} = \frac{(\mathrm{SD}_T)^2}{N_T\overline{X}_T^2} + \frac{(\mathrm{SD}_C)^2}{N_C\overline{X}_C^2} \approx \frac{(s)^2}{N_T\overline{X}_T^2} + \frac{(s)^2}{N_C\overline{X}_C^2}$$

Raw study parameters can then be removed by multiplying with $\overline{X}_C^2$ and substituting $\overline{X}_C/\overline{X}_T$ with $1/\mathrm{RR}$ to give:

$$\mathrm{var(lnRR)} \cdot \overline{X}_C^2 \approx \frac{(s)^2\overline{X}_C^2}{N_T\overline{X}_T^2} + \frac{(s)^2\overline{X}_C^2}{N_C\overline{X}_C^2} = \frac{(s)^2}{N_T RR^2} + \frac{(s)^2}{N_C}$$

and with some rearrangement to isolate $s/\overline{X}_C$, an approximation emerges:

$$\frac{s}{\overline{X}_C} \approx \sqrt{\frac{N_T N_C \mathrm{var(lnRR)}}{N_T/\mathrm{RR}^2 + N_C}} \tag{9}$$

Multiplying the response ratio by $s/\overline{X}_C$ gives $s/\overline{X}_T$, while adding $\overline{X}_C/s$ to d gives $\overline{X}_T/s$. If both the response ratio and d were reported for the same study outcome, then $(\mathrm{RR}-1)/d = s/\overline{X}_C$. This approximation of $s/\overline{X}_C$ (Equation 9), as well as the recovery of $s/\overline{X}_T$, may offer an alternative source for the coefficient of variation needed to impute the variances of response ratios with missing SD (see Nakagawa et al. 2023). Finally, with an additional assumption that sample sizes are equal ($N = N_T = N_C$), a simpler approximation that omits but transfers sample sizes onto $s/\overline{X}_C$ is also available:

$$\frac{s}{\overline{X}_C\sqrt{N}} \approx \sqrt{\frac{\mathrm{var(lnRR)}}{1/\mathrm{RR}^2 + 1}} \tag{10}$$

and is equivalent to the ratio of the pooled standard error (SE) to the control mean.

2.5 | Converting RR to Hedges' *d*

Given the working estimate of $s/\overline{X}_C$, it is now possible to convert RR to an approximate d (hereafter $\tilde{d}$) by combining Equations (8) and (9) to give:

$$\vec{d} = (\mathrm{RR} - 1)\sqrt{\frac{N_{\mathrm{T}} / \mathrm{RR}^2 + N_{\mathrm{C}}}{N_{\mathrm{T}} N_{\mathrm{C}} \mathrm{var}(\ln \mathrm{RR})}} \tag{11}$$

No conversion is necessary for estimating the variance of $\vec{d}$ given that, for better or for worse (see Hamman et al. 2018), Equation (11) can be used directly to estimate its large-sample variance (Hedges 1981):

$$\mathrm{var}\left(\vec{d}\right) = \frac{1}{N_{\mathrm{T}}} + \frac{1}{N_{\mathrm{C}}} + \frac{\vec{d}^{\,2}}{2\left(N_{\mathrm{T}} + N_{\mathrm{C}}\right)} \tag{12}$$

2.6 | Converting RR to a Two-Sample *t*-Test

The response ratio can also be converted to an approximate two-sample, equal-variance t-test ($\vec{t}$) when combining Equations (8) and (10) to give:

$$\vec{t} = (\mathrm{RR} - 1)\sqrt{\frac{1 / \mathrm{RR}^2 + 1}{2\mathrm{var}(\ln \mathrm{RR})}} \tag{13}$$

This $\vec{t}$ has degrees of freedom (df) equaling $\mathrm{df} = N_{\mathrm{T}} + N_{\mathrm{C}} - 2$ and a p-value calculated in base R (R Core Team 2024) with 2 * pt(−abs(t), df) or in MS Excel with 2 * T.DIST(−ABS(t), df, TRUE). In turn, and following Rosenthal's (1994) conversion of a t-test to d, this $\vec{t}$ approximated from the response ratio can be converted to $\vec{d}$ when sample sizes are assumed equal ($N = N_{\mathrm{T}} = N_{\mathrm{C}}$):

$$\vec{d} = \vec{t}\sqrt{2 / N} \tag{14}$$

and when sample sizes are unequal:

$$\vec{d} = \vec{t}\sqrt{1 / N_{\mathrm{T}} + 1 / N_{\mathrm{C}}} \tag{15}$$

3 | Performance of Converting RR to *d*

Here I use Monte Carlo simulations to assess the accuracy of converting RR to d (Equation 11) for a given sample size (N), whether small-sample corrections are needed, and finally the sensitivity to violations of the assumption of homogeneity of variances (see Approximating scales to establish a conversion). Simulations quantify accuracy as the bias ($\mathrm{bias} = \delta - \mathbb{S}$) between the expected mean of the standardised mean difference between treatment and control groups, or $\delta = \left(\mu_{\mathrm{T}} - \mu_{\mathrm{C}}\right) / \sigma_p^2$, and the mean of randomly simulated effect sizes ($\mathbb{S}$). Here, μ_{T} and μ_{C} are independent and normally distributed ($\mathcal{N}$) population means and σ_p^2 the expected pooled (p) and weighted average of the two population standard deviations (σ^2) of μ_{T} and μ_{C} (see Equation 3). Raw study parameters (e.g., $\overline{X}$, SD) used to simulate and calculate effect sizes ($\mathbb{S}$) were randomly generated by sampling and averaging ith number of N samples from $X_i \sim \mathcal{N}\left(\mu, \sigma^2\right)$. Simulations manipulated sample sizes of effect sizes from $N = 4$–36, magnitudes of μ from 15 to 25 yielding a standardised mean difference spanning −1 to 1 and lnRR from −0.51 to 0.51, and also manipulated whether standard deviations were equal ($\sigma_{\mathrm{T}}^2 = \sigma_{\mathrm{C}}^2$ with $\sigma^2 = 5$) or unequal ($\sigma_{\mathrm{T}}^2 \neq \sigma_{\mathrm{C}}^2$ with $5 \neq 4$ and $5 \neq 3$). For each $\mathbb{S}$ simulated in base R (R Core Team 2024), 500,000 effect sizes were randomly generated (> 722 million used to assess bias), and the simulated $\mathbb{S}$ include d (Equation 2) and g (Equation 5) given they are estimators of δ (Hedges 1981), RR (Equation 1) and RR^{Δ} (Equation 4), and finally RR converted to d (Equation 11). Note I accent effect sizes derived from conversions with an arrow (i.e., d converted from RR becomes $\vec{d}$). The usefulness of small-sample corrections was assessed before conversions (RR^{Δ} to d which will be $\vec{d}^{\,\Delta}$), after ($\vec{d}$ corrected giving $\vec{g}$), and both before and after ($\vec{d}^{\,\Delta}$ corrected giving $\vec{g}^{\,\Delta}$). Finally, I compare $\mathbb{S}$ from the RR-to-d conversion suggested by Murad et al. (2019) with $d = \ln \mathrm{RR} / 0.392$.

When assumptions of homogeneity are met, response ratios converted to d with Equation (11) are near equivalent to d (compare d to $\vec{d}$ when $\sigma_{\mathrm{T}}^2 = \sigma_{\mathrm{C}}^2$ in Figure 1). Hedges' d overestimates effects at small sample sizes (d in Figure 1; also see Hedges 1981), so does the conversion of RR to d ($\vec{d}$ in Figure 1), but the small-sample correction by Hedges (1981) largely removes this bias from both d and $\vec{d}$ (see g and $\vec{g}$ in Figure 1), while Lajeunesse's (2015) correction offered only marginal gains when applied to RR prior to conversions ($\vec{d}^{\,\Delta}$ and $\vec{g}^{\,\Delta}$ in Figure 1). When assumptions of homogeneity are violated ($\sigma_{\mathrm{T}}^2 \neq \sigma_{\mathrm{C}}^2$), conversions of RR to d ($\vec{d}$) or RR to g ($\vec{g}$) will underestimate effects when the treatment group is more variable than the control group ($\sigma_{\mathrm{T}}^2 > \sigma_{\mathrm{C}}^2$), and will overestimate effects when the control arm is more variable (see $\vec{d}$ and $\vec{g}$ in cases when $\sigma_{\mathrm{T}}^2 < \sigma_{\mathrm{C}}^2$ in Figure 1). Finally, the linear transformation by Murad et al. (2019) converts RR to d only under null conditions when $\overline{X}_{\mathrm{T}}$ is near equivalent to $\overline{X}_{\mathrm{C}}$.

4 | A Worked Example

Here I reanalyse data from Rinehart and Hawlena (2020) to offer a numerical example and to assess how conversions perform with real study outcomes extracted from the literature. This meta-analysis reported all the raw study parameters needed to calculate lnRR and d for 339 study outcomes (i.e., $\overline{X}$, SD, N). I start with a conversion example, follow with a comparison across study outcomes, and end with a meta-analysis of effects.

Rinehart and Hawlena (2020) extracted blood glucose levels from Brachetta et al. (2016) with the goal to combine and contrast stoichiometry effects in the presence and absence of predators. Glucose levels for tuco-tuco rodents when cats were absent and present were, respectfully, $\overline{X}_{\mathrm{T}} = 88.97$ mg/mL ($\mathrm{SD}_{\mathrm{T}} = 26.27$, $N_{\mathrm{T}} = 18$) and $\overline{X}_{\mathrm{C}} = 64.0$ mg/mL ($\mathrm{SD}_{\mathrm{C}} = 20.39$, $N_{\mathrm{C}} = 12$). With these study parameters, effect sizes were $d = 1.0306$ with $\mathrm{var}(d) = 0.1566$ (Equations 2 and 12), $\ln \mathrm{RR} = 0.3283$ with $\mathrm{var}(\ln \mathrm{RR}) = 0.0133$ (Equation 1), and after exponentiating lnRR to $\mathrm{RR} = 1.3886$, this response ratio converted to d using Equation (11) equalled $\vec{d} = 1.0585$ with a $\mathrm{var}\left(\vec{d}\right) = 0.1576$. Finally, the t-test using the raw study parameters was similar to the t-test converted from RR using Equation (13): $t = 2.7765$ ($\mathrm{df} = 28$, $p = 0.0097$) versus $\vec{t} = 2.9347$ ($\mathrm{df} = 28$, $p = 0.0066$). Note that conversions are not exact given the differences between SD_{T} and SD_{C}, and the imbalance of N among experimental groups.

Now focusing on the entire dataset and comparing effect sizes from all study outcomes with the caveat that one was excluded due to having $\mathrm{SD} = 0$ (which leads to division by zero in d) and another two were excluded due to having negative ratios (which cannot be log transformed to lnRR), the correlation between RR and d was high but discrepancies between metrics increased with magnitude of effect ($r = 0.537$, $R^2 = 0.288$, $k = 336$; Figure 2a). This slanted 'bowtie' relationship between RR and d with the knot centered around null effects was also observed by Hillebrand and Gurevitch (2014) (their $r = 0.426$, $R^2 = 0.182$, $k = 838$) and Friedrich et al. (2011) (their lnRR vs. d: $r = 0.787$, $R^2 = 0.62$, $k = 232$). I explain this correlation in Section 6.4. However, after converting RR to d with Equation (11), the correlation between $\vec{d}$ and d was strong with fewer discrepancies among effect sizes ($r = 0.969$, $R^2 = 0.938$, $k = 336$; Figure 2b)—which was more performant than the conversion via imputation by Murad et al. (2019) that preserved the bowtie pattern ($r = 0.627$, $R^2 = 0.394$, $k = 336$; Figure 2c). The bowtie discrepancies were removed when converting RR into d with Equation (11), but among those that remained (see Figure 2b), many can be explained by idiosyncrasies among study outcomes. For example, $\vec{d}$'s that differed from d by 0.05 magnitude of effect were due to having imputed zero standard errors (SE) with 0.01 SD in the dataset ($k = 1$), having double negative outcomes between treatment and control arms ($k = 2$), and having outcomes with SD's that were at least 2 times more variable than the other ($k = 9$). Excluding these outliers ($k = 12$) resulted in d versus $\vec{d}$ correlation of $r = 0.997$ ($R^2 = 0.995$, $k = 324$; Figure 2d). Finally, the conversion of RR to t to d (Equation 13 then Equation 15) had near equivalent values to those from the direct RR to d conversion (Equation 15 vs. Equation 11: $r = 0.998$, $k = 336$).

Finally, using the *metafor* package for *R* (v. 4.6-0; Viechtbauer 2010), the pooled effect of d was similar to the pooled $\vec{d}$ (pooled $d = 0.052$, 95% CI [−0.048, 0.152], $\tau^2 = 0.676$, $k = 324$; pooled $\vec{d} = 0.045$, 95% CI [−0.056, 0.147], $\tau^2 = 0.704$, $k = 324$). Results are from separate random-effects meta-analyses using a restricted maximum likelihood estimate of the between-study variance (τ^2).

5 | Best Practices for Converting Response Ratios

Given the simulations and worked example, I recommend the Hedges' g small-sample correction after converting response

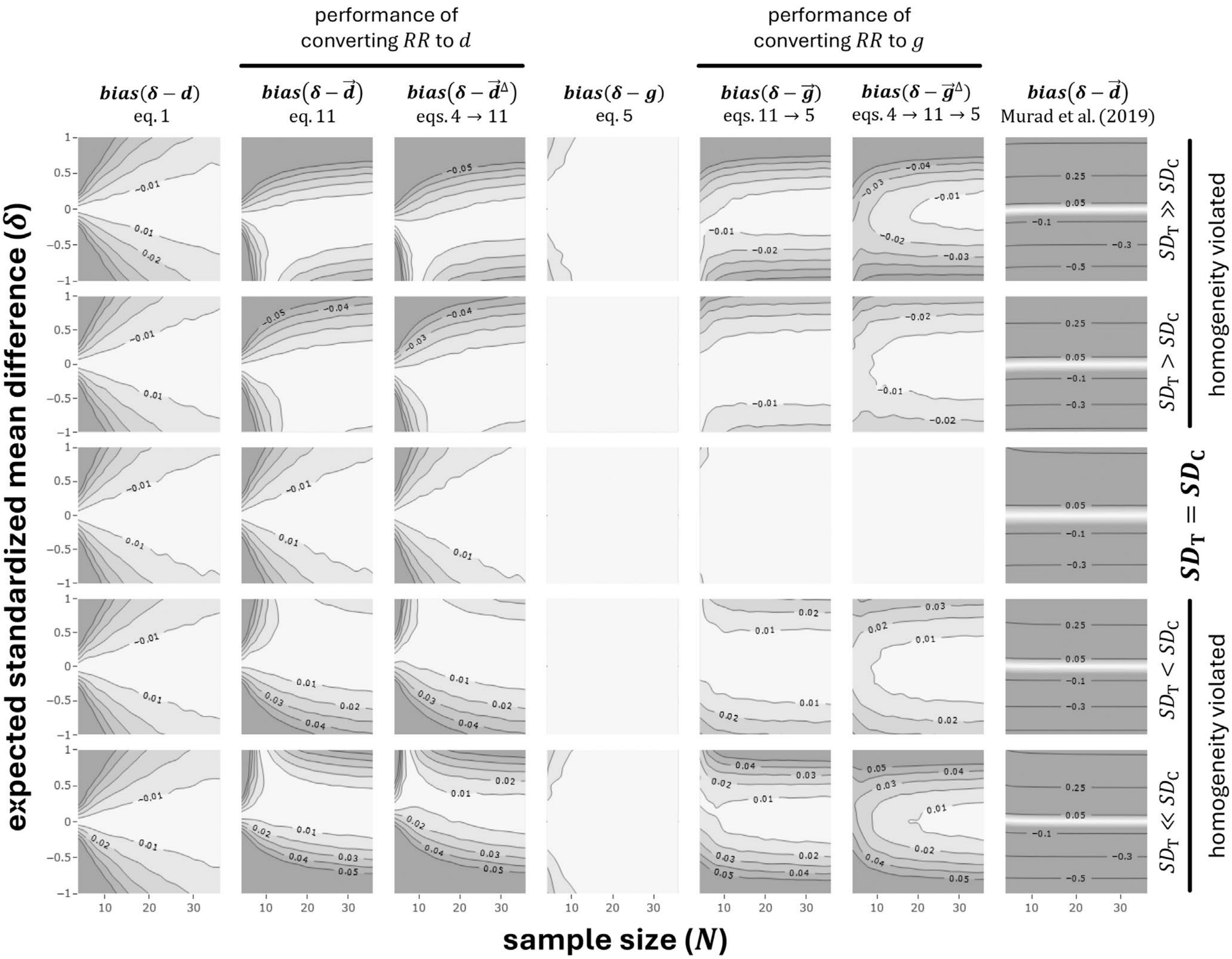


FIGURE 1 | Results from Monte Carlo simulations assessing bias when converting response ratios (RR, Equation 1) to Hedges' d effect sizes (Equation 2). Grey shading and contour lines define the magnitude of bias (e.g., relative difference from the expected standardised mean difference or δ; see Section 3), with white indicating little bias, when comparing conversions, small-sample uncorrected and corrected effect size metrics, and changes to the assumption of homogeneity in the standard deviations (SD) of the treatment and control means. Here RR converted to d is $\vec{d}$, small-sample corrected RR to d is $\vec{d}^{\Delta}$, small-sample corrected d is g, and small-sample corrected RR converted to small-sample corrected d is $\vec{g}^{\Delta}$. Finally, also presented is the bias from Murad et al. (2019) imputation conversion: $d = \text{lnRR} / 0.392$.

ratios to *d*. This correction is straightforward (Equation 5; but also see Hedges 1981) when converting lnRR to *g*:

$$\vec{g} = \vec{d}\left(1 - \frac{3}{4\left[N_T + N_C\right] - 9}\right) = \left(e^{\ln RR} - 1\right)\sqrt{\frac{N_C + N_T/e^{2\ln RR}}{N_T N_C \operatorname{var}(\ln RR)}}\left(1 - \frac{3}{4\left[N_T + N_C\right] - 9}\right) \tag{16}$$

which has a variance of:

$$\operatorname{var}(\vec{g}) = \frac{1}{N_T} + \frac{1}{N_C} + \frac{\vec{d}^{\,2}}{2(N_T + N_C)}\left(1 - \frac{3}{4\left[N_T + N_C\right] - 9}\right)^2 = \frac{1}{N_T} + \frac{1}{N_C} + \frac{\vec{g}^{\,2}}{2(N_T + N_C)} \tag{17}$$

The correction is needed because *d* overestimates the effect at small samples (see Figure 1), and given that low *N* is common among ecological studies (Jennions and Møller 2003), it will help shrink *d* into *g* which will be closer to expected value when sample sizes are large.

My simulations also made it clear that conversions are more sensitive to violations of homogeneity of variances than both *d* or *g* under similar heteroscedastic conditions (e.g., compare $\vec{d}$ and *d* under non-equivalent SD's in Figure 1). However, addressing this issue is challenging and perhaps has no solution given there are no clear ways to assess the equality of variances after effect sizes are calculated or published without raw study parameters (e.g., means, SD, samples sizes). However, violating the assumption of homogeneity of variances (homoscedasticity) is a perennial concern for mean-based effect sizes, where some have argued that generally the SD of the control and treatment groups are more likely to differ than not (Grissom and Kim 2001; Senior et al. 2020), while others have further argued to move entirely past this assumption and always use equations of the standardised mean difference that do not homogenise variances (e.g., Aoki 2020; Delacre et al. 2021).

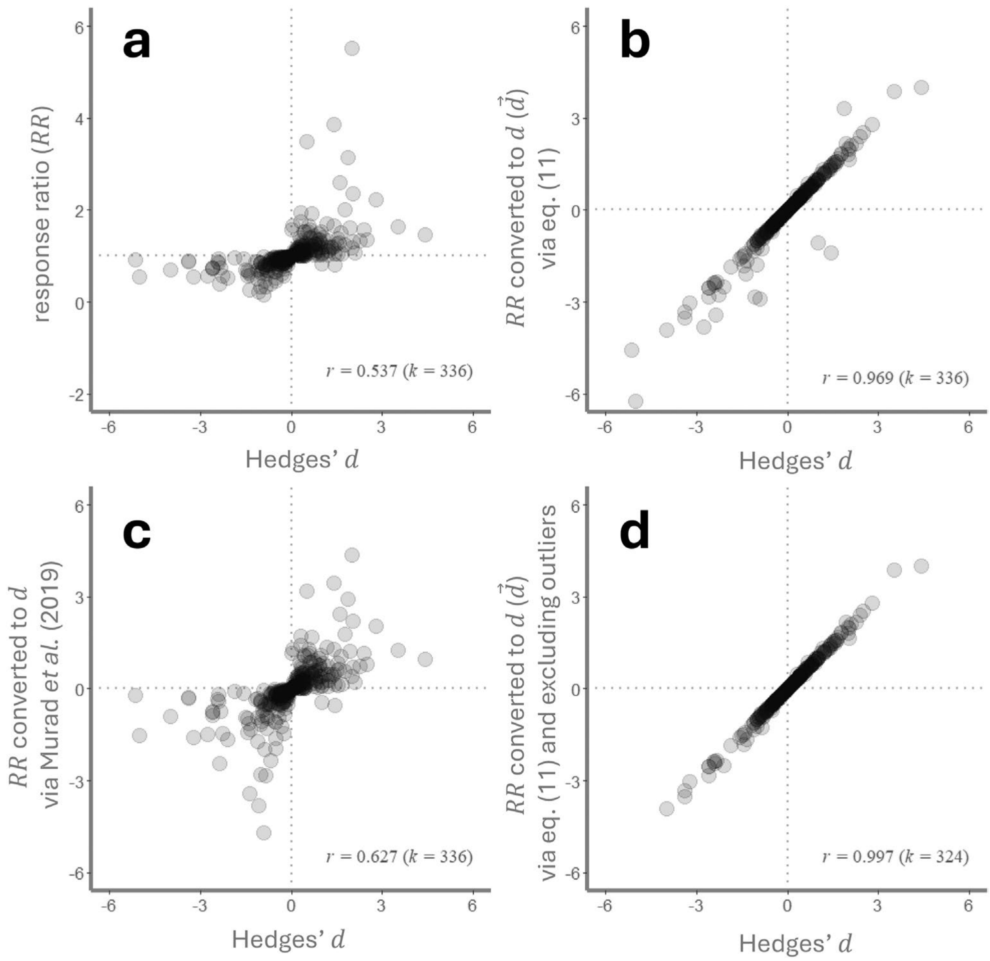


FIGURE 2 | Example of converting response ratios to Hedges' *d* effect sizes using the published dataset of 336 study outcomes from Rinehart and Hawlena (2020). Panel (a) is the relationship between Hedges' *d* effect sizes and response ratios (RR) using the same raw study parameters (i.e., means, standard deviations, sample sizes), (b) the relationship between *d* and RR's converted to *d* (i.e., $\vec{d}$) with Equation (11), (c) the relationship *d* and RR's converted to *d* using the linear transformation suggested by Murad et al. (2019); ($d = \ln RR / 0.392$) and finally (d) the relationship between *d* and RR's converted to $\vec{d}$ with Equation (11) but also excluding conversion outliers (see exclusion criteria in section: A worked example). Also presented are the correlation coefficients (*r*) of each of these pairwise relationships among effect size metrics and the total number of effect sizes (*k*). An *r* value close to 1 indicates a near exact conversion among metrics.

However, since the accuracy of the RR-to-d conversion is dependent on the variance of the response ratio, it is important to identify the equation used to estimate var(lnRR) prior to conversions to help understand the risk with heteroscedasticity. The most widely used variance equation for the log response ratio is the unequal-variance version ($\sigma_T^2 \neq \sigma_C^2$; Equation 1 from Hedges et al. 1999):

$$\frac{(\mathrm{SD_T})^2}{N_\mathrm{T}\overline{X}_\mathrm{T}^2} + \frac{(\mathrm{SD_C})^2}{N_\mathrm{C}\overline{X}_\mathrm{C}^2}$$

less commonly used, there is the equal-variance version ($\sigma_T^2 \neq \sigma_C^2$; unnumbered in Hedges et al. 1999):

$$\frac{(s)^2}{N_\mathrm{T}\overline{X}_\mathrm{T}^2} + \frac{(s)^2}{N_\mathrm{C}\overline{X}_\mathrm{C}^2} = (s)^2\left(\frac{1}{N_\mathrm{T}\overline{X}_\mathrm{T}^2} + \frac{1}{N_\mathrm{C}\overline{X}_\mathrm{C}^2}\right)$$

and finally, a reduced form (Hedges and Olkin 1985; Adams et al. 1997; Lajeunesse 2013b), which assumes that the coefficient of variation among treatment and control groups equal one (Nakagawa et al. 2023), is also sometimes used:

$$\frac{1}{N_\mathrm{T}} + \frac{1}{N_\mathrm{C}} = \frac{N_\mathrm{T} + N_\mathrm{C}}{N_\mathrm{T} N_\mathrm{C}}$$

The equal-variance version of var(lnRR) will yield exact conversions from RR-to-g, the unequal-variance version will yield similar but inaccurate conversions when $\sigma_T^2 \neq \sigma_C^2$ (see Figure 1), but the sample-size reduced form should not be used with conversions as it will significantly overestimate g. Consequently, effect size data originating from meta-analyses that use the N-reduced form as weights for meta-analysis may not be reusable—perhaps this is the grand trade-off of synthesising research with poorly reported standard deviations (see Lajeunesse and Forbes 2003; also see example Hargreaves et al. 2020). Note there may be a pathway for reuse at the synthesis-level using the Nakagawa et al. (2023) imputation method for missing RR weights and then using these imputed weights for conversions.

Once effect sizes are converted, secondary analysis of mixtures of g and $\vec{g}$ should expect more variability among $\vec{g}$ due to the conversion being more sensitive to $\sigma_T^2 \neq \sigma_C^2$ violations than g (see Figure 1). To assess the severity of this violation, and similar to previous best practices for handling converted effect sizes (see Lajeunesse 2013b), a sensitivity analysis that pools and contrasts g and $\vec{g}$ should be completed prior to the synthesis to justify that together they represent a homogeneous group and why these effects can be pooled in a single meta-analysis.

6 | Geometric Extension for Converting Response Ratios

When deriving the RR-to-d conversion (see Common effect but different scales), rather than using the ratio of means ($\overline{X}_\mathrm{T}/\overline{X}_\mathrm{C}$), I used an alternate but equivalent form of the response ratio based on the mean difference ($\overline{X}_\mathrm{T} - \overline{X}_\mathrm{C}$; Equation 6). This was intuitive because the mean-difference is a defining component of d

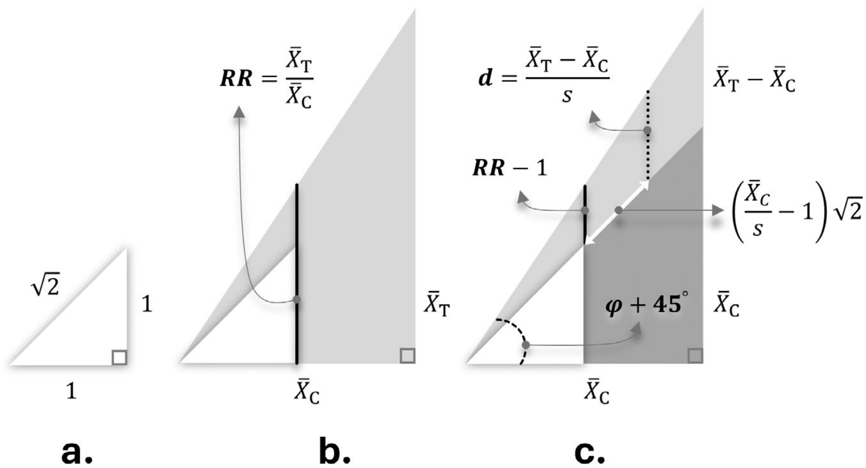


FIGURE 3 | The response ratio (RR) as a constructible effect size metric and Hedges' d as a straightedge construction of RR. For further context of this constructable framework see Box 1. (a) A white isosceles right triangle with legs of unit length and hypotenuse equal to $\sqrt{2}$ that allows for effect size metrics to be constructed geometrically and algebraically (see Herstein 1996). (b) Given the light-grey right triangle with legs of magnitude $\overline{X}_\mathrm{C}$ and $\overline{X}_\mathrm{T}$ (control and treatment means), RR will equal the length of a line segment parallel to $\overline{X}_\mathrm{T}$ but intersects with $\overline{X}_\mathrm{C}$ at one unit length from the start of $\overline{X}_\mathrm{C}$. Finally, (c) Hedges' d constructed with the geometry of RR. Using only straightedge rules, and upscaling the unit triangle (a) to have legs equal to $\overline{X}_\mathrm{C}$ (becoming the dark-grey right triangle of c), d will equal the dashed line segment parallel and proportional to both the mean difference ($\overline{X}_\mathrm{T} - \overline{X}_\mathrm{C}$) and RR − 1. The magnitude and direction of RR will remain fixed relative to $\overline{X}_\mathrm{T} - \overline{X}_\mathrm{C}$, because it is anchored to the unit triangle defined in (a), but the magnitude of d will scale with the pooled standard deviation (s; see Equation 3)—emphasised here by the white double arrow line equalling $\sqrt{2}(\overline{X}_\mathrm{C}/s - 1)$ which models the RR-to-d scaling factor defined in the section: Common effect but different scales. Presented is the case when $s > 1$, but when $s < 1$, the dashed line representing d will remain parallel and proportional to $\overline{X}_\mathrm{T} - \overline{X}_\mathrm{C}$ and RR − 1, but will be found on the right side of $\overline{X}_\mathrm{T} - \overline{X}_\mathrm{C}$ following the extended path traced by the hypotenuse of the dark-grey right triangle (not presented). This hypotenuse of the dark-grey triangle equalling $\overline{X}_\mathrm{C}\sqrt{2}$ also defines the directionality and null value of effect sizes. For example, given that RR − 1, d and $\overline{X}_\mathrm{T} - \overline{X}_\mathrm{C}$ are all proportional with a common angle (φ), then $\varphi = 0^\circ$ is the null boundary (e.g., where $\overline{X}_\mathrm{T} = \overline{X}_\mathrm{C}$, RR = 1, $d = 0$), but when φ is negative (i.e., rotating clockwise below either hypotenuse of the dark grey triangle or the unit triangle), then $\overline{X}_\mathrm{T} < \overline{X}_\mathrm{C}$, RR < 1 and d will be a negative. In the example presented here, φ is positive.

(Equation 2). However, although never used in practice, the response ratio also has many other equivalent forms based on mean sums ($\overline{X}_\text{T} + \overline{X}_\text{C}$), products ($\overline{X}_\text{T}\overline{X}_\text{C}$), or even sum–differences:

$$\ln\text{RR} = \ln\left(\frac{\overline{X}_\text{T}}{\overline{X}_\text{C}}\right) = \ln\left(\frac{\overline{X}_\text{T} - \overline{X}_\text{C}}{\overline{X}_\text{C}} + 1\right) = \ln\left(\frac{\overline{X}_\text{T} + \overline{X}_\text{C}}{\overline{X}_\text{C}} - 1\right)$$
$$= 2\ln\left(\frac{\sqrt{\overline{X}_\text{T}\overline{X}_\text{C}}}{\overline{X}_\text{C}}\right) = 2\text{atanh}\left(\frac{\overline{X}_\text{T} - \overline{X}_\text{C}}{\overline{X}_\text{T} + \overline{X}_\text{C}}\right) \tag{18}$$

with atanh being the inverse hyperbolic tangent function. This diversity of ways to stating the response ratio points to additional conversion opportunities because they contain different ways to model the relationship between $\overline{X}_\text{T}$ and $\overline{X}_\text{C}$. For example, rather than converting response ratio as a mean difference, the $\overline{X}_\text{T} + \overline{X}_\text{C}$ in the sum version of the response ratio can be used to estimate the arithmetic mean of the control and treatment groups, or $(\overline{X}_\text{T} + \overline{X}_\text{C})/2$.

Here I outline a generalisable framework to take advantage of these conversion opportunities and offer context to the metric scaling between RR and d. For example, given that the response ratio can be defined with a sum, difference, product, or ratio (quotient), as well as with arithmetic and geometric means (Equation 18), then these features indicate that this effect size metric is algebraically and geometrically constructable (sensu Wantzel 1837; but see Hungerford 1997). Constructability is a key field of abstract algebra where numerical values can be restructured using simple rules based on a compass (circle building algebra or geometry) and straightedge (line building). I outline the tools and rules of this method in Box 1. For mean-based effect sizes, this framework allows for effect sizes to be restated into different constructs of effects (e.g., different pairwise relationships between $\overline{X}_\text{T}$ and $\overline{X}_\text{C}$) with simple rules and without additional information.

In the following sections, I apply these rules to first define the response ratio as a constructable effect size metric. Then I show how this constructability can be used to repurpose response ratios into a diversity of pairwise relationships between $\overline{X}_\text{T}$ and $\overline{X}_\text{C}$, and finally extend this method to show that d itself is a compass-and-straightedge construction (proportional extension) of the response ratio. Formal proofs of constructability are tedious and beyond the scope of this paper. Therefore, I rely on visual/graphical proofs using geometric diagrams to layout the framework. Finally, all the tools and rules of the approach applied below are outlined in Box 1.

6.1 | RR Is a Constructable Effect Size Metric

The response ratio is a constructible effect size metric with a magnitude and direction that can be stated geometrically. Starting with two lengths with magnitudes $\overline{X}_\text{C}$ and $\overline{X}_\text{T}$, and each forming the legs of a right triangle, RR will equal the length of an intersecting line that is parallel to $\overline{X}_\text{T}$, but is one unit length away from the start of $\overline{X}_\text{C}$ (see visual proof in Figure 3a,b). This construction reveals two key properties: (1) the smaller (white) right triangle with an altitude of RR and unit base is proportional to the larger (light-grey) triangle with altitude $\overline{X}_\text{T}$, and (2) these two triangles are similar and have identical angles.

These properties of proportionality and similarity (Box 1) reveal that the response ratio can be constructed using strictly straightedge rules (following Hungerford 1997) starting with a unit length (Figure 3b), and following Thales' intercept theorem (i.e., segments generated by parallel lines are proportional; Box 1; also Berger 2009, 49), can be converted into other segments of length $\overline{X}_\text{T} + \overline{X}_\text{C}$, $\overline{X}_\text{T} - \overline{X}_\text{C}$, or $\overline{X}_\text{T}\overline{X}_\text{C}$ (see Figure 4). Further extending this constructability with the geometric mean theorem (which introduces circle diameters as a compass to construct the square roots needed for Pythagorean means; Berger 2009, Chapter 10; Box 1), the response ratio can be fully repurposed via compass-and-straightedge into a broad diversity of measures of relative change between $\overline{X}_\text{T}$ and $\overline{X}_\text{C}$ (see Figure 4). These include all the classic Pythagorean means (arithmetic, geometric, harmonic, quadratic means; see Box 1), which I rank from smallest to largest mean-values based on well-known geometric inequalities among means when $\overline{X}_\text{T} \neq \overline{X}_\text{C}$ (Bullen 2003):

$$\begin{array}{ccccccccc}
\text{Harmonic mean (HM)} & < & \text{Geometric mean (GM)} & < & \text{Logarithmic mean (LM)} & < & \text{Arithmetic mean (AM)} & < & \text{Quadratic mean (QM)} \\
\underbrace{\text{HM}(\overline{X}_\text{T},\overline{X}_\text{C})/\overline{X}_\text{C}}_{\frac{2\text{RR}}{\text{RR}+1}} & & \underbrace{\text{GM}(\overline{X}_\text{T},\overline{X}_\text{C})/\overline{X}_\text{C}}_{\sqrt{\text{RR}}} & & \underbrace{\text{LM}(\overline{X}_\text{T},\overline{X}_\text{C})/\overline{X}_\text{C}}_{\frac{\text{RR}-1}{\ln\text{RR}}} & & \underbrace{\text{AM}(\overline{X}_\text{T},\overline{X}_\text{C})/\overline{X}_\text{C}}_{\frac{\text{RR}+1}{2}} & & \underbrace{\text{QM}(\overline{X}_\text{T},\overline{X}_\text{C})/\overline{X}_\text{C}}_{\sqrt{\frac{\text{RR}^2+1}{2}}}
\end{array} \tag{10}$$

Note that all conversions are standardised—here the standardizer or scaling factor is the same as in the response ratio (i.e., $1/\overline{X}_\text{C}$). Unfortunately, reverse-engineering the measurement scale of $\overline{X}_\text{C}$ is not possible after effect sizes are calculated. Table 1 itemises how to repurpose RR.

6.2 | Hedges' *d* as a Construction of RR

There are many ways to state d as a compass-and-straightedge construction of the response ratio. Using only straightedge rules (see Box 1), and upscaling the unit triangle (Figure 3a) to have legs equal to $\overline{X}_\text{C}$ (depicted as the dark grey right triangle in Figure 3c), this new triangle divides the geometry of the response ratio into a segment that outlines the mean difference ($\overline{X}_\text{T} - \overline{X}_\text{C}$; see Figure 3c). Since the line segment representing RR − 1 is parallel to $\overline{X}_\text{T} - \overline{X}_\text{C}$, and both share the same angle (φ), then following Thales' theorem (Box 1), RR − 1 will be

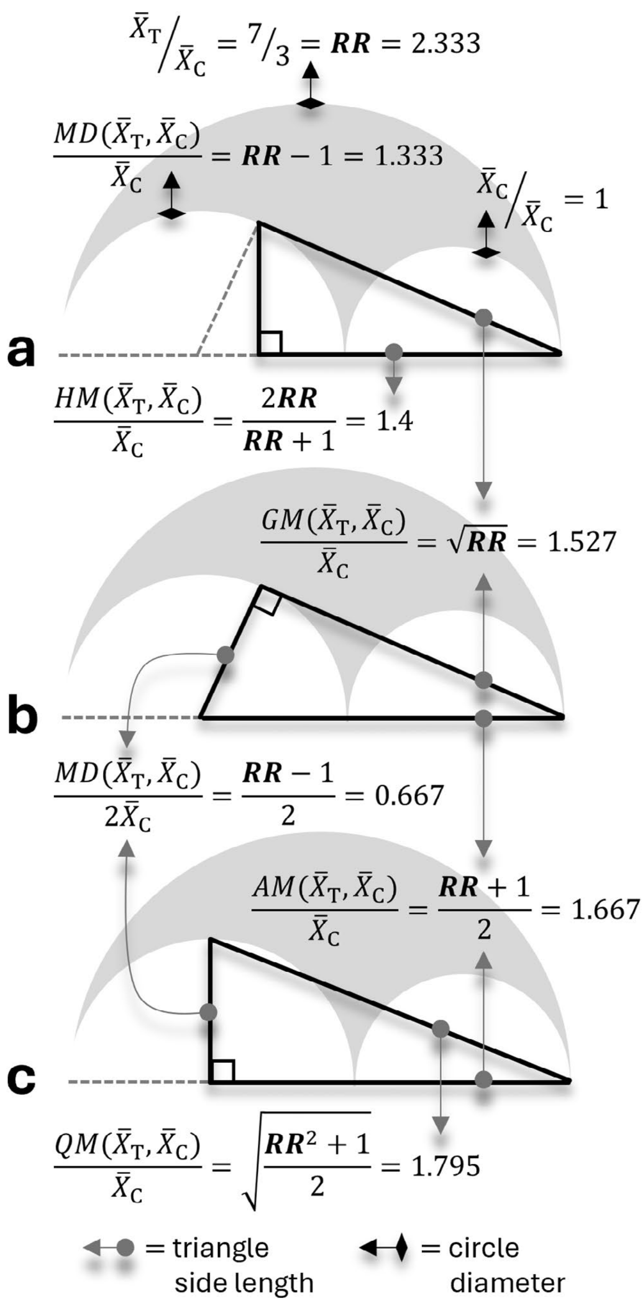


FIGURE 4 | Geometries of the response ratio (RR) constructed with a diversity of Pythagorean means (Box 1). Figure 3 constructed RR strictly with a straightedge. Here all examples use both a compass and straightedge rules to construct RR giving a geometry based on an outer circle with a diameter of the treatment/manipulated mean ($\bar{X}_T$) that is tangent to two congruent and packed (kissing) circles with diameters of size $\bar{X}_C$ (control/reference mean) and the mean difference (MD) between the treatment and control groups or $\bar{X}_T - \bar{X}_C$. The response ratio expresses this geometry on the $1/\bar{X}_C$ scale and thus all three circle diameters are standardised by the control mean. Geometries of RR stated as tangent circles (i.e., Soddy circles) of a right-angle triangle using (a) geometric ($GM = \sqrt{\bar{X}_T \bar{X}_C}$) and harmonic ($HM = 2\bar{X}_T\bar{X}_C / [\bar{X}_T + \bar{X}_C]$) means of $\bar{X}_T$ and $\bar{X}_C$, (b) geometric (GM) and arithmetic means ($AM = [\bar{X}_T + \bar{X}_C]/2$) and (c) arithmetic (AM) and quadratic means ($QM = \sqrt{[\bar{X}_T^2 + \bar{X}_C^2]/2}$). Also provided is a numerical example when $\bar{X}_T = 7$ and $\bar{X}_C = 3$. Finally, this geometry is not possible with lnRR as the logarithmic scale will yield transcendental values that are not practical for devising constructions or conversions. See Table 1 for additional context.

proportional to $\bar{X}_T - \bar{X}_C$ by a scaling factor of $1/\bar{X}_C$. Hedges' d is also proportional to $\bar{X}_T - \bar{X}_C$, but with a scaling factor of $1/s$, and therefore d and RR − 1 are proportional with a scaling factor of $\bar{X}_C / s$ (see Figure 3c). This geometry explains why RR can be restated algebraically into $\bar{X}_T - \bar{X}_C$ (see lnRR and d estimate a common effect), why d is a proportional construction of both $\bar{X}_T - \bar{X}_C$ and RR (Figure 3c), and why the algebraic conversion between RR and d works (see Common effect but different scales).

6.3 | Retrieving a Grand Mean From RR and *d*

Similar to how the response ratio can be repurposed geometrically (see Figure 4), the constructability of d relative to RR offers analogous opportunities. For example, stating d as a compass-and-straightedge construction of RR (Figure 5), we can retrieve the standardised arithmetic mean—which is the mean of means (i.e., $AM = \bar{X} = [\bar{X}_T + \bar{X}_C]/2$) or the grand (overall) mean when ith number of samples (X_i) from $\bar{X}_T$ and $\bar{X}_C$ are pooled. As in the case when effect sizes quantify size differences between males and females, but an estimate of overall size is needed. This standardised grand mean can be extracted from RR when sample sizes are equal ($N_T = N_C$) and unequal ($N_T \neq N_C$) with:

$$\frac{\text{AM}(\bar{X}_T, \bar{X}_C)}{\bar{X}_C} = \underbrace{\frac{\text{RR}+1}{2}}_{\text{if } N_T = N_C} \text{ or } \underbrace{\frac{N_T\text{RR}+N_C}{N_T+N_C}}_{\text{if } N_T \neq N_C} = \frac{\sum_i^{N_T} X_i + \sum_i^{N_C} X_i}{N_T+N_C}\left(\frac{1}{\bar{X}_C}\right)$$

and will have a variance approximated as:

$$\text{var}\left[\frac{\text{AM}(\bar{X}_T, \bar{X}_C)}{\bar{X}_C}\right] \approx \sqrt{\frac{N_T N_C}{N_T+N_C-1}\left[\frac{(\text{RR}-1)^2}{N_T+N_C} + \frac{\text{RR}^2(N_T+N_C-2)\text{var}(\ln\text{RR})}{N_T/\text{RR}^2+N_C}\right]}$$

A standardised grand mean can also be retrieved from d using RR as follows:

$$\frac{\text{AM}(\bar{X}_T, \bar{X}_C)}{s} = \underbrace{d\left[\frac{\text{RR}+1}{2(\text{RR}-1)}\right]}_{\text{if } N_T = N_C} \text{ or } \underbrace{d\left[\frac{N_T\text{RR}+N_C}{(N_T+N_C)(\text{RR}-1)}\right]}_{\text{if } N_T \neq N_C}$$

$$= \frac{\sum_i^{N_T} X_i + \sum_i^{N_C} X_i}{N_T+N_C}\left(\frac{1}{s}\right)$$

or approximated strictly with RR using Equation (11):

$$\frac{\text{AM}(\bar{X}_T, \bar{X}_C)}{s} \approx \left(\frac{\text{RR}+1}{2}\right)\sqrt{\frac{N_T/\text{RR}^2+N_C}{N_T N_C \text{var}(\ln\text{RR})}}$$

Table 1 outlines other retrievable assets from d.

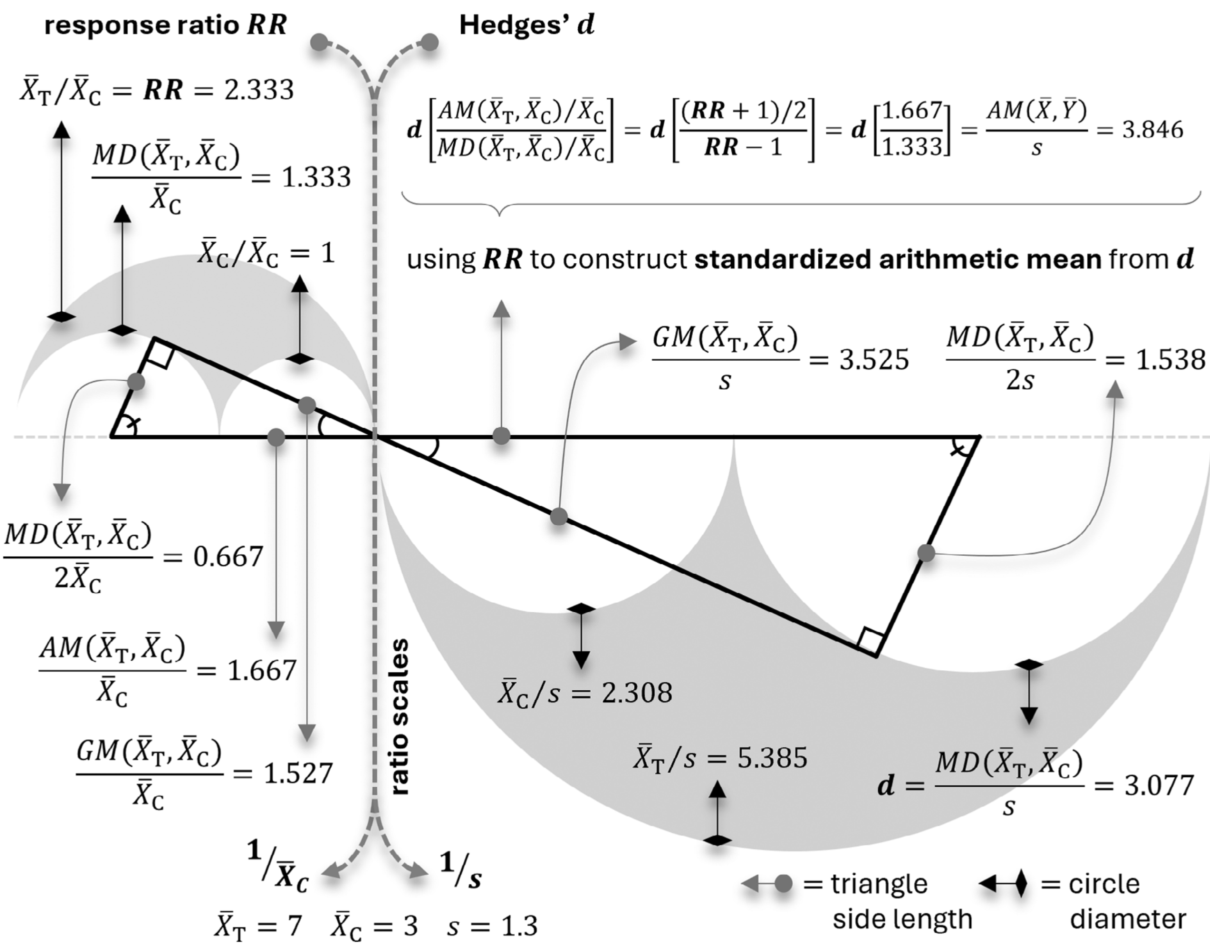


FIGURE 5 | The shared geometry of the response ratio (RR) and Hedges' d effect sizes with a numerical example on how to use RR to construct d. Both metrics have the same tangent circle construction presented in Figure 4; however, the geometry of Hedges' d is standardised on the $1/s$ scale with s being the pooled standard deviation of $\bar{X}_T$ and $\bar{X}_C$ (Equation 3), while RR is standardised on the $1/\bar{X}_C$ scale. Because RR expresses the control group $\bar{X}_C$ as a unit length ($\bar{X}_C/\bar{X}_C = 1$), the metric is constructable given the geometric mean theorem (Berger 2009; but see Box 1), and can be used as a compass-and-straightedge constructor for d (standardised mean difference between $\bar{X}_T$ and $\bar{X}_C$, MD$/s$ or $[\bar{X}_T - \bar{X}_C]/s$). Constructability allows for retrieval of Pythagorean means from d (Box 1), like the inlaid example of using RR to retrieve the standardised arithmetic mean (AM; see Section 6.3). Also presented are how the geometric mean (GM) and the mean difference (MD) relate to RR and d under a single geometric framework. The standardised AM, GM and MD, along with other mean assets, are defined in Table 1.

6.4 | Insights on the Bowtie Relationship Between RR and d

The bowtie relationship between RR and d is pathological of strong dependencies between metrics. The most overt is the direction of the effect. Since d and RR will never have opposing directionality, this binds effect sizes to the lower-left bow when $\bar{X}_T < \bar{X}_C$ (1st quadrant delineated by the null values of d and RR; see Figure 2a) and the top-right bow when $\bar{X}_T > \bar{X}_C$ (3rd quadrant). Consequently, there will always be a strong positive correlation between RR and d (Figure 2a). Opportunistically, this expectation of common directionality can be used to detect outliers, should d and RR (or lnRR) be jointly reported in a single meta-analysis, as effect sizes occupying the 2nd and 4th quadrants (see outliers in Figure 2b) likely originate from inappropriate values of $\bar{X}_T$ and $\bar{X}_C$ (e.g., when both are negative). Directionality here does not necessarily imply positive or negative values; but rather how $\bar{X}_T < \bar{X}_C$ or $\bar{X}_T > \bar{X}_C$ relate to the null. Finally, note that the expectation of common directionality can break down when multiple effect sizes are pooled, because the different weighting between metrics combined with different between-study variance estimates can result in divergent magnitudes and directions of synthesis-level effects. This is why pooled d and lnRR can have conflicting effects (see Spake et al. 2021).

The magnitudes of RR and d also have dependencies. First, d will always be proportional to RR − 1 (see Figure 3c), and this proportionality dependence further strengthens under null conditions when $\bar{X}_T$ becomes near equivalent to $\bar{X}_C$. This dependency structures the distribution of effects around the knot in the bowtie pattern between RR and d (Figure 2a). Second, and not independent of proportionality nor directionality, is that the correlated relationship between RR and d is fully dependent on all of the underlying study parameters used to compute effects (i.e., N, SD and $\bar{X}$; Lajeunesse and Forbes 2003). These common study parameters translate (via Delta method) into a complex covariance (cov) structure between metrics:

TABLE 1 | Retrievable datatypes from the response ratio (RR) and Hedges' d effect sizes given their constructable geometry (see Figures 3–5; but see methods in Box 1).

Datatype	Unstandardised definition $f(\bar{X}_T,\bar{X}_C)$	Response ratio (RR) $\frac{f(\bar{X}_T,\bar{X}_C)}{\bar{X}_C}$	Hedges' d $\frac{f(\bar{X}_T,\bar{X}_C)}{s}$	Details
Treatment mean	$\bar{X}_T$	RR (Equation 3)	$\frac{d}{1-1/\mathrm{RR}}$	Isolates means for use in single arm studies on continuous outcomes
Control mean	$\bar{X}_C$	1	$\frac{d}{\mathrm{RR}-1}$	
Mean difference (MD)	$\bar{X}_T-\bar{X}_C$	$\mathrm{RR}-1$	d (Equation 4)	
Mean sum (MS)	$\bar{X}_T+\bar{X}_C$	$RR+1$	$\frac{d(RR+1)}{RR-1}$*	Cumulative effects
$\bar{X}_T$ sum portion of effect size (ES)*	$\frac{\bar{X}_T ES}{\bar{X}_T+\bar{X}_C}$	$\frac{\mathrm{RR}^2}{\mathrm{RR}+1}$	$\frac{d}{\mathrm{RR}+1}$	$\bar{X}_T$ sum portion + $\bar{X}_C$ sum portion = ES (either RR or d)
$\bar{X}_C$ sum portion of effect size (ES)*	$\frac{\bar{X}_C ES}{\bar{X}_T+\bar{X}_C}$	$\frac{\mathrm{RR}}{\mathrm{RR}+1}$	$\frac{d}{1+1/\mathrm{RR}}$	
Quadratic mean (QM)	$\sqrt{\frac{\bar{X}_T^2}{2}-\frac{\bar{X}_C^2}{2}}$	$\sqrt{\frac{\mathrm{RR}^2+1}{2}}$	$\frac{d\sqrt{\mathrm{RR}^2+1}}{\sqrt{2}(\mathrm{RR}-1)}$	Also known as root mean square, used as a measure of tree diameter change (see Curtis and Marshall 2000)
Arithmetic mean (AM)	$\frac{\bar{X}_T+\bar{X}_C}{2}$	$\frac{\mathrm{RR}+1}{2}$	$\frac{d(\mathrm{RR}+1)}{2(\mathrm{RR}-1)}$**	See Section 6.3
Logarithmic mean (LM)	$\frac{\bar{X}_T-\bar{X}_C}{\ln(\bar{X}_T/\bar{X}_C)}$	$\frac{\mathrm{RR}-1}{\ln\mathrm{RR}}$	$\frac{d}{\ln\mathrm{RR}}$	An alternative way to interpret ln RR is the MD with respect to LM (Törnqvist et al. 1985)
Geometric mean (GM)	$\sqrt{\bar{X}_T\bar{X}_C}$	$\sqrt{\mathrm{RR}}=e^{1/2\ln\mathrm{RR}}$	$\frac{d\sqrt{\mathrm{RR}}}{\mathrm{RR}-1}$	Used as a measure of biodiversity change (see Buckland et al. 2011)
Harmonic mean (HM)	$\frac{2}{\frac{1}{\bar{X}_T}-\frac{1}{\bar{X}_C}}$	$\frac{2\mathrm{RR}}{\mathrm{RR}+1}$	$\frac{2d(\mathrm{RR}-1)}{1+1/\mathrm{RR}}$	Used as a measure of shifts in animal activities (Dixon and Chapman 1980)

Note: Datatypes are presented as functions (f) of treatment ($\bar{X}_T$) and control means ($\bar{X}_C$) that are either standardised with $f/\bar{X}_C$ for response ratio constructs or with f/s for constructs retrievable from d using RR. Here s is the pooled standard deviation of treatment and control means (Equation 3).
*Also equals: $d/\tanh(0.5\ln\mathrm{RR})$ with tanh being the hyperbolic tangent function.
**Also equals: $d/2\tanh(0.5\ln\mathrm{RR})$.

$$\mathrm{cov}(\ln\mathrm{RR}, d) = \frac{s}{\bar{X}_C}\left[\frac{(\mathrm{SD}_C)^2}{N_C}+\frac{(\mathrm{SD}_T)^2}{N_T\mathrm{RR}}\right]$$

These multiple dependencies are why the bowtie pattern cannot be modelled with a simple bivariate (joint) normal distribution, and why the empirically-derived linear conversion by Murad et al. (2019) imputes poorly despite the strong correlation between metrics (i.e., $d=\ln\mathrm{RR}/0.392$; see Figure 2c). Here, extracting the slope ($\beta=1/0.392$) from a regression between observed response ratios and d (e.g., the linear fit from Figure 3a) will not yield a consistent conversion because it assumes that this slope is common among all effects—which is not realistic given that N, SD and $\bar{X}$ will vary across a diversity of sources extracted from the literature. If study parameters were estimating common population parameters across studies (i.e., all have a common effect, equal SD and large and balanced N), then the expected slope (β) of Murad et al. (2019) no-intercept regression model ($d=\beta\cdot\ln\mathrm{RR}$) equals the standardised logarithmic mean (LM) of the treatment and control groups:

$$\beta = \frac{\mathrm{LM}(\bar{X}_T,\bar{X}_C)}{s} = \left(\frac{1}{s}\right)\frac{\bar{X}_T-\bar{X}_C}{\ln(\bar{X}_T/\bar{X}_C)} = \frac{d}{\ln\mathrm{RR}}$$

while the no-intercept model of unlogged response ratios, $d=\beta(\mathrm{RR}-1)$, will have the scaling factor of the RR-to-d conversion as its $\beta=\bar{X}_C/s$ (see white double arrow in Figure 3c).

6.5 | On Converting d to RR

Unfortunately, converting d to RR may not be possible. At the algebraic level of needing to retrieve $s/\bar{X}_C$ for conversions, the scaling factor ($1/s$) is lost when d computes an effect (Equation 2), and given that d has a non-central t-distribution that needs the standardised mean difference to approximate its variance (see Hedges 1981), $1/s$ is also lost when var(d) is calculated (Equation 12). At the geometric level, d is only constructible in the context of RR—which anchors d to the unit triangle (a defining component of constructability; see Figure 3a and Box 1). Given the model outlined in Figure 3, it may not be possible to retrieve anything from d or var(d) to convert RR, or repurpose the metric into other datatypes of relative change when only d, var(d) and N are known.

7 | Discussion

Extracting effect sizes from published studies and integrating them into high-quality meta-analyses is difficult and time-consuming (Haddaway and Westgate 2019). However, once extracted, these effect sizes have long-term value because they can be converted and reused for new contexts. With the aid of the conversion of the response ratio (RR) to standardised mean difference (Hedges' d or g), existing data available only as RR can now be fully integrated into broader synthesis frameworks (e.g., Yang et al. 2023) and maximally reused to extract additional value from the broad and diverse study outcomes published by ecologists. Rather than repeating guidelines for data reuse and secondary analysis (Bishop and Kuula-Luumi 2017), or best practices for effect size conversions (Lajeunesse 2013b; Polanin and Snilstveit 2016), below I aim to integrate the RR-to-d conversion among the mathematical, interpretive and conceptual models needed to achieve a common currency across ecological studies.

7.1 | Mathematical Models

There are many mathematical models needed to support the conversion of published effect sizes to a common metric—such as d. For example, correlation coefficients (r) can be converted to d's because they are different bivariate expressions of the same general linear model; this translates to an effect size transformation: $r\sqrt{\left(N_T+N_C\right)^2/\left(N_T N_C\right)}=d$ (see Pustejovsky 2014). Odds ratios (lnOR) can be converted to d's because under the generalised linear model, a logistic distribution can have similar shape to a normal distribution; this translates to an effect size approximation: $\ln\mathrm{OR}\sqrt{3}/\pi=d$ (Chinn 2000). Finally, response ratios can be converted to d's because ratios and differences of common study parameters are algebraic expressions of a single, coherent geometric model (Figure 5); this translates to an effect size construction: $(\mathrm{RR}-1)\overline{X}_{\mathrm{C}}/s=d$.

These transformations, approximations and constructions all convert to d and there may be an opportunity to develop a common mathematical framework. Leveraging geometric extensions of the linear model may be one approach (Lee Rodgers and Nicewander 1988; Agresti 2015); although see Herr (1980) for anticipated challenges. A unified framework would help structure the many loose ideas with conversion practices; especially given that research synthesis treats them more as tricks-of-the-trade (Lajeunesse 2013b) rather than a standardised approach aimed at preserving the underlying structure of the effect. A common model would also help define boundaries on what can be achieved—such as when conversions are not commutative. For example, r-to-d and d-to-r are not necessarily interchangeable transformations because r can originate from a broad diversity of bivariate designs with different assumptions (see Pustejovsky 2014). Similarly, the RR-to-d conversion is dependent on what version of the response ratio variance is used (see Best practices for converting response ratios), or that d-to-RR may not be directly possible given that the standardised mean difference does not offer ways to approximate $\overline{X}_{\mathrm{C}}/s$ (see On converting d to RR). A common model is perhaps facilitated by the shared assumptions of conversions: all are non-exact or biased when variances are heterogeneous across groups, when group observations are not independent, or when sample sizes are small or unbalanced.

7.2 | Interpretive Models

Achieving a common currency across studies, or more explicitly total equivalency of scales (sensu Becker and Wu 2007), also requires the consolidation of interpretive models. Starting with RR and d being unitless (Senior et al. 2020). Here 'unitless' is meant to emphasise that the diverse measurement scales of $\overline{X}$ reported across studies (i.e., grams) get cancelled when an effect is calculated. It also means that both metrics are scale-invariant in the sense that the measurement scale/units of both $\overline{X}_{\mathrm{T}}$ and $\overline{X}_{\mathrm{C}}$ will not change the magnitude of effect (e.g., the ratio of $\overline{X}_{\mathrm{T}}/\overline{X}_{\mathrm{C}}$ in grams/grams is equivalent to pounds/pounds; Pustejovsky 2018). However, because lnRR and d achieve this cancellation differently, they operate under different dimensional scales or metric units: with lnRR being in units of the control mean ($1/\overline{X}_{\mathrm{C}}$) on a natural log scale, and d's in units of standard deviations ($1/s$). Effect sizes in units of the control mean ($1/\overline{X}_{\mathrm{C}}$) are intuitive because meaning is not tied to any specific statistical model and can be easily restated as percentages—which are a very direct and numerically meaningful way to describe relative change (e.g., '10% increase'). However, effect sizes in units of standard deviations are more versatile and statistically interpretable because meaning is interconnected with the linear model—which allows for cross-metric interpretation (dimensional consistency) between d and other effect size metrics (i.e., r, OR). This is why, for example, d can be interpreted as event odds without having to convert the effect (i.e., $d=5$ corresponds to odds of once every 4776 years), and why the numerical values of d are meaningful because the odds of high magnitudes of effect exponentially decrease along the tails of the Normal distribution (Hedges and Olkin 1985).

Converting response ratios to d therefore trades intuitive clarity for dimensional consistency. Unfortunately, the RR-to-d conversion itself does not offer new insight into how to cross-interpret these metrics. This is because RR only has partial dimensional consistency with d—with the mean difference ($\overline{X}_{\mathrm{T}}-\overline{X}_{\mathrm{C}}$) being the coherent derived measure between metrics, and not the dimensional scale, which is fully swapped from $1/\overline{X}_{\mathrm{C}}$ to $1/s$ during conversions. This means that although RR will always be proportional (and correlated) with d (see Figures 2b and 3), magnitudes of d are not helpful to interpret values of RR or vice versa.

Having no reference for the dimensional scale of the response ratio results in numerical values never being interpreted directly—which is significant inferential gap with the metric. There is a literal interpretation of lnRR, as the relative change between the treatment and control means with respect to their logarithmic mean, or $\ln\mathrm{RR}=\left(\overline{X}_{\mathrm{T}}-\overline{X}_{\mathrm{C}}\right)/\mathrm{LM}\left(\overline{X}_{\mathrm{T}},\overline{X}_{\mathrm{C}}\right)$

(Vartia 1976; Table 1). But this definition is also not helpful with cross-interpreting effects between metrics, nor does stating mean differences on a $1/\mathrm{LM}(\overline{X}_\mathrm{T}, \overline{X}_\mathrm{C})$ scale offer any intuitive gains with interpreting magnitudes of effect. However, similar to d, lnRR also has an exponential interpretive scale useful to identify outliners. For example, certain magnitudes of lnRR are unlikely and only possible when inappropriate study parameters are used to compute effects—such as when $\overline{X}_\mathrm{C}$ and $\overline{X}_\mathrm{T}$ are on different measurement scales/units, or when imputations or standard errors are used to compute effect size variances rather than standard deviations (see Lajeunesse 2015). Here are a few extreme values: $\ln\mathrm{RR} \approx 4.6$ (treatment group is 100× bigger than the control), $\ln\mathrm{RR} \approx 2.3 (10\times)$ and $\ln\mathrm{RR} \approx 1.6$ (5×) and for reference $\ln\mathrm{RR} \approx 0.05$ (1.05× bigger).

Finally, the geometric constructability framework (Figures 3–5; Box 1) opens new possibilities to interpret and visualise effect sizes, and also perhaps offers a pathway to more generalised, unitless metrics for mean differences. For example, estimating the common angle (φ) between $\mathrm{RR} - 1$, d and $\overline{X}_\mathrm{T} - \overline{X}_\mathrm{C}$ in degrees which equals $\varphi = \mathrm{atan}(\mathrm{RR})180/\pi - 45^{\circ}$ (with atan being the arc tangent function; see Figure 3c), or preferably in radians due to its mathematical versatility ($\varphi = \mathrm{atan}(\mathrm{RR}) - \pi/4$), is an unambiguous way to simplify and standardise all the dimensional scales across $\mathrm{RR} - 1$, d and $\overline{X}_\mathrm{T} - \overline{X}_\mathrm{C}$. Effect sizes as angles (i.e., direction effect sizes) are also intuitive because magnitudes quantify perpendicularity (i.e., effects approaching 90° indicate a strong differences between $\overline{X}_\mathrm{T}$ and $\overline{X}_\mathrm{C}$). Effect sizes as angles also offers unique ways visualise effects because each effect can have a vector to illustrate more information about the underlying study outcome. Similar to bubble plots used in meta-regressions, vector length can be proportional to effect size weights (i.e., $1/\mathrm{var}[\mathrm{RR}]$), or proportional to normalised to relative weights (i.e., $\max[\mathrm{var}(\mathrm{RR})]/\mathrm{var}[\mathrm{RR}]$) with the unit circle defining the maximum relative weight. Pooling direction effect sizes may also be possible by adapting meta-analysis models to circular statistics (Fisher 1993).

7.3 | Conceptual Models

Converting effect sizes to a common currency will homogenise the core conceptual models of these data. In particular, converting RR-to-d will restate multiplicative effects (modelled by RR) to additive ones (modelled by d), and this can be a significant recontextualization of ecological scale (Spake et al. 2023). In fact, one reason why response ratios were developed and introduced into ecology was because of criticisms that d poorly models non-linear (multiplicative) effects (see Osenberg et al. 1997, 1999). However, it is an open question on how many published meta-analyses justify choosing RR over d explicitly to model multiplicative biological processes over additive ones (as advocated by Osenberg et al. 1999; Spake and Doncaster 2017); as opposed to choosing RR due to equation simplicity, inclusivity of poorly reported study outcomes (Lajeunesse and Forbes 2003), or feature prominence bias in ecological publications and software (Osenberg and St. Mary 1998). Nonetheless, although achieving total equivalency of metric scales across effect sizes is possible (i.e., conversions can now generate a common currency), it will not be total equivalency of ecological scales because the data will form a mixture of additive and incorrectly modelled multiplicative effects (or vice versa). This mixture is either due to incorrect applications of the RR-to-d conversion, or due to the original (unconverted) effect size not being quantified with the appropriate metric (Osenberg et al. 1997). Therefore, without careful attention to distinguish linear and non-linear mechanisms, the values and variances of effects in a unified dataset will likely not offer clean insight on the underlying ecological processes. However, in the context of broad secondary analyses across ecological studies, the RR-to-d conversion is an opportunity to test whether conceptual models were justified given that non-linear ecological processes or relationships modelled incorrectly would presumably be more variable or biased.

Finally, the ability to restate underlying effects without using new information (Figure 4) also emphasises that additive and multiplicative effects are not necessarily conceptually isolated, or even independent ways to model ecological phenomena, but expressions of a single, coherent framework (Figure 5). With simple algebraic rules (Box 1), there is fluid movement from one effect to another—emphasising that ecological outcomes can be reframed to shift interpretive scales without loss of the core underlying effect. The constructible framework further enables many new and unconventional ways to recontextualize study outcomes (Table 1); such as restating RR or d as homogenised effects (arithmetic or pooled effects; see Section 6.3), cumulative effects (sum or total effects), geometric effects (central tendency of multiplicative effects), or even reciprocal effects (inverse or harmonic effects). A proportional response of two pesticides (multiplicative effects) for instance can be restated as total pesticide impact (cumulative effects). Restating underlying effects may be an opportunity to better match study outcomes to the underlying ecological model without having to derive new effect size metrics, extract new data, or adjust existing metrics with known confounders (sensu Osenberg et al. 1997).

Author Contributions

M.J.L. conceptualised, derived equations, analysed/visualised data and wrote the study.

Acknowledgements

I took an abstract algebra course in college, nearly failed it, thought it was all nonsense. With that said, I thank three anonymous reviewers for improving clarity and validating that perhaps it was not all nonsense in the end.

Funding

This work was supported by the University of South Florida.

Data Availability Statement

R scripts and data for replicating Figures 1–5 are available on Figshare: https://doi.org/10.6084/m9.figshare.30782225.

Peer Review

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/ele.70335.

References

Adams, D. C., J. Gurevitch, and M. S. Rosenberg. 1997. "Resampling Tests for Meta-Analysis of Ecological Data." *Ecology* 78: 1277–1283.

Agresti, A. 2015. *Foundations of Linear and Generalized Linear Models*. John Wiley & Sons.

Aoki, S. 2020. "Effect Sizes of the Differences Between Means Without Assuming Variance Equality and Between a Mean and a Constant." *Heliyon* 6: e03306.

Arnqvist, G., and D. Wooster. 1995. "Meta-Analysis: Synthesizing Research Findings in Ecology and Evolution." *Trends in Ecology & Evolution* 10: 236–240.

Barto, E. K., and M. C. Rillig. 2012. "Dissemination Biases in Ecology: Effect Sizes Matter More Than Quality." *Oikos* 121: 228–235.

Bebout, J., and J. W. Fox. 2024. "Coordinated Distributed Experiments in Ecology Do Not Consistently Reduce Heterogeneity in Effect Size." *Oikos* 2024: e10722.

Becker, B. J., and M.-J. Wu. 2007. "The Synthesis of Regression Slopes in Meta-Analysis." *Statistical Science* 22: 414–429.

Berger, M. 2009. *Geometry I*. Springer Science & Business Media.

Bergquist, M., M. Thiel, M. H. Goldberg, and S. van der Linden. 2023. "Field Interventions for Climate Change Mitigation Behaviors: A Second-Order Meta-Analysis." *Proceedings of the National Academy of Sciences* 120: e2214851120.

Bishop, L., and A. Kuula-Luumi. 2017. "Revisiting Qualitative Data Reuse: A Decade on." *SAGE Open* 7: 1–15.

Brachetta, V., C. E. Schleich, and R. R. Zenuto. 2016. "Source Odor, Intensity, and Exposure Pattern Affect Antipredatory Responses in the Subterranean Rodent *Ctenomys talarum*." *Ethology* 122: 923–936.

Buckland, S. T., A. C. Studeny, A. E. Magurran, J. B. Illian, and S. E. Newson. 2011. "The Geometric Mean of Relative Abundance Indices: A Biodiversity Measure With a Difference." *Ecosphere* 2: 1–15.

Bullen, P. S. 2003. "The Arithmetic, Geometric and Harmonic Means." In *Handbook of Means and Their Inequalities*, 60–174. Springer Netherlands.

Cardinale, B., D. Srivastava, E. J. Duffy, et al. 2006. "Effects of Biodiversity on the Functioning of Trophic Groups and Ecosystems." *Nature* 443: 989–992.

Chinn, S. 2000. "A Simple Method for Converting an Odds Ratio to Effect Size for Use in Meta-Analysis." *Statistics in Medicine* 19: 3127–3131.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. Routledge Academic.

Costello, L., and J. W. Fox. 2022. "Decline Effects Are Rare in Ecology." *Ecology* 103: e3680.

Curtis, R. O., and D. D. Marshall. 2000. "Why Quadratic Mean Diameter?" *Western Journal of Applied Forestry* 15: 137–139.

Delacre, M., D. Lakens, C. Ley, L. Liu, and C. Leys. 2021. "Why Hedges' g Based on the Non-Pooled Standard Deviation Should Be Reported With Welch's *t*-Test." *PsyArXiv Preprint*. https://doi.org/10.31234/osf.io/tu6mp. (Last Accessed 23 August 2025).

Dixon, K. R., and J. A. Chapman. 1980. "Harmonic Mean Measure of Animal Activity Areas." *Ecology* 61: 1040–1044.

Elser, J. J., M. E. S. Bracken, E. E. Cleland, et al. 2007. "Global Analysis of Nitrogen and Phosphorus Limitation of Primary Producers in Freshwater, Marine and Terrestrial Ecosystems." *Ecology Letters* 10: 1135–1142.

Fisher, N. I. 1993. *Statistical Analysis of Circular Data*. Cambridge University Press.

Fox, J. W. 2022. "How Much Does the Typical Ecological Meta-Analysis Overestimate the True Mean Effect Size?" *Ecology and Evolution* 12: e9521.

Friedrich, J. O., N. K. J. Adhikari, and J. Beyene. 2008. "The Ratio of Means Method as an Alternative to Mean Differences for Analyzing Continuous Outcome Variables in Meta-Analysis: A Simulation Study." *BMC Medical Research Methodology* 8: 32.

Friedrich, J. O., N. K. J. Adhikari, and J. Beyene. 2011. "Ratio of Means for Analyzing Continuous Outcomes in Meta-Analysis Performed as Well as Mean Difference Methods." *Journal of Clinical Epidemiology* 64: 556–564.

Grissom, R. J., and J. J. Kim. 2001. "Review of Assumptions and Problems in the Appropriate Conceptualization of Effect Size." *Psychological Methods* 6: 135–146.

Haddaway, N. R., and A. S. Pullin. 2014. "The Policy Role of Systematic Reviews: Past, Present and Future." *Springer Science Reviews* 2: 179–183.

Haddaway, N. R., and M. J. Westgate. 2019. "Predicting the Time Needed for Environmental Systematic Reviews and Systematic Maps." *Conservation Biology* 33: 434–443.

Hamman, E. A., P. Pappalardo, J. R. Bence, S. D. Peacor, and C. W. Osenberg. 2018. "Bias in Meta-Analyses Using Hedges' d." *Ecosphere* 9: e02419.

Hargreaves, A. L., R. M. Germain, M. Bontrager, J. Persi, and A. L. Angert. 2020. "Local Adaptation to Biotic Interactions: A Meta-Analysis Across Latitudes." *American Naturalist* 195: 395–411.

Hedges, L. V. 1981. "Distribution Theory for Glass's Estimator of Effect Size and Related Estimators." *Journal of Educational Statistics* 6: 107–128.

Hedges, L. V. 1982. "Estimation of Effect Size From a Series of Independent Experiments." *Psychological Bulletin* 92: 490–499.

Hedges, L. V., J. Gurevitch, and P. S. Curtis. 1999. "The Meta-Analysis of Response Ratios in Experimental Ecology." *Ecology* 80: 1150–1156.

Hedges, L. V., and I. Olkin. 1985. *Statistical Methods for Meta-Analysis*. Academic Press.

Herr, D. G. 1980. "On the History of the Use of Geometry in the General Linear Model." *American Statistician* 34: 43–47.

Herstein, I. N. 1996. *Abstract Algebra*. 3rd ed. Wiley.

Hillebrand, H., and J. Gurevitch. 2014. "Meta-Analysis Results Are Unlikely to Be Biased by Differences in Variance and Replication Between Ecological Lab and Field Studies." *Oikos* 123: 794–799.

Hoeksema, J. D., V. B. Chaudhary, C. A. Gehring, et al. 2010. "A Meta-Analysis of Context-Dependency in Plant Response to Inoculation With Mycorrhizal Fungi." *Ecology Letters* 13: 394–407.

Hungerford, T. W. 1997. *Abstract Algebra: An Introduction*. 2nd ed. Saunders College Publishing.

Jennions, M. D., and A. P. Møller. 2003. "A Survey of the Statistical Power of Research in Behavioral Ecology and Animal Behavior." *Behavioral Ecology* 14: 438–445.

Lajeunesse, M. J. 2010. "Achieving Synthesis With Meta-Analysis by Combining and Comparing All Available Studies." *Ecology* 91: 2561–2564.

Lajeunesse, M. J. 2011. "On the Meta-Analysis of Response Ratios for Studies With Correlated and Multi-Group Designs." *Ecology* 92: 2049–2055.

Lajeunesse, M. J. 2013a. "Power Statistics for Meta-Analysis: Tests for Mean Effects and Homogeneity." In *Handbook of Meta-Analysis in Ecology and Evolution*, edited by J. Koricheva, J. Gurevitch, and K. Mengersen, 348–363. Princeton University Press.

Lajeunesse, M. J. 2013b. "Recovering Missing or Partial Data From Studies: A Survey of Conversions and Imputations for Meta-Analysis." In *Handbook of Meta-Analysis in Ecology and Evolution*, edited by J. Koricheva, J. Gurevitch, and K. Mengersen, 195–206. Princeton University Press.

Lajeunesse, M. J. 2015. "Bias and Correction for the Log Response Ratio in Ecological Meta-Analysis." *Ecology* 96: 2056–2063.

Lajeunesse, M. J., and M. R. Forbes. 2003. "Variable Reporting and Quantitative Reviews: A Comparison of Three Meta-Analytical Techniques." *Ecology Letters* 6: 448–454.

Lee Rodgers, J., and W. A. Nicewander. 1988. "Thirteen Ways to Look at the Correlation Coefficient." *American Statistician* 42: 59–66.

Lipsey, M. W., and D. B. Wilson. 2001. *Practical Meta-Analysis*. Sage Publications, Inc.

Macauley, M., and N. Youngs. 2020. "The Case for Algebraic Biology: From Research to Education." *Bulletin of Mathematical Biology* 82: 115.

McGrath, R. E., and G. J. Meyer. 2006. "When Effect Sizes Disagree: The Case of r and d." *Psychological Methods* 11: 386–401.

Møller, A., and M. D. Jennions. 2002. "How Much Variance Can Be Explained by Ecologists and Evolutionary Biologists?" *Oecologia* 132: 492–500.

Murad, M. H., Z. Wang, H. Chu, and L. Lin. 2019. "When Continuous Outcomes Are Measured Using Different Scales: Guide for Meta-Analysis and Interpretation." *BMJ (Clinical Research Ed.)* 364: k4817.

Nakagawa, S., D. W. Noble, M. Lagisz, R. Spake, W. Viechtbauer, and A. M. Senior. 2023. "A Robust and Readily Implementable Method for the Meta-Analysis of Response Ratios With and Without Missing Standard Deviations." *Ecology Letters* 26: 232–244.

Nakagawa, S., and E. S. A. Santos. 2012. "Methodological Issues and Advances in Biological Meta-Analysis." *Evolutionary Ecology* 26: 1253–1274.

Osenberg, C. W., O. Sarnelle, and S. D. Cooper. 1997. "Effect Size in Ecological Experiments: The Application of Biological Models to Meta-Analysis." *American Naturalist* 150: 798–812.

Osenberg, C. W., O. Sarnelle, S. D. Cooper, and R. D. Holt. 1999. "Resolving Ecological Questions Through Meta-Analysis: Goals, Metrics, and Models." *Ecology* 80: 1105–1117.

Osenberg, C. W., and C. M. St. Mary. 1998. "Meta-Analysis: Synthesis or Statistical Subjugation?" *Integrative Biology: Issues, News and Views* 1: 43–48.

Parker, J. D., D. E. Burkepile, and M. E. Hay. 2006. "Opposing Effects of Native and Exotic Herbivores on Plant Invasions." *Science* 311: 1459–1461.

Polanin, J. R., and B. Snilstveit. 2016. *Campbell Methods Policy Note on Converting Between Effect Sizes (Version 1.1, Updated December 2016)*. Campbell Collaboration.

Poom, L., and A. af Wåhlberg. 2022. "Accuracy of Conversion Formula for Effect Sizes: A Monte Carlo Simulation." *Research Synthesis Methods* 13: 508–519.

Pustejovsky, J. E. 2014. "Converting From d to r to z When the Design Uses Extreme Groups, Dichotomization, or Experimental Control." *Psychological Methods* 19: 92–112.

Pustejovsky, J. E. 2015. "Measurement-Comparable Effect Sizes for Single-Case Studies of Free-Operant Behavior." *Psychological Methods* 20: 342–359.

Pustejovsky, J. E. 2018. "Using Response Ratios for Meta-Analyzing Single-Case Designs With Behavioral Outcomes." *Journal of School Psychology* 68: 99–112.

R Core Team. 2024. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/.

Rao, C. R. 1973. *Linear Statistical Inference*. John Wiley & Sons.

Rinehart, S., and D. Hawlena. 2020. "The Effects of Predation Risk on Prey Stoichiometry: A Meta-Analysis." *Ecology* 101: e03037.

Rosenthal, R. 1994. "Parametric Measures of Effect Size." In *Handbook of Research Synthesis*, edited by H. Cooper and L. Hedges, 231–244. Russel Sage Foundation.

Rubin, D. B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66: 688–701.

Senior, A. M., C. E. Grueber, T. Kamiya, et al. 2016. "Heterogeneity in Ecological and Evolutionary Meta-Analyses: Its Magnitude and Implications." *Ecology* 97: 3293–3299.

Senior, A. M., W. Viechtbauer, and S. Nakagawa. 2020. "Revisiting and Expanding the Meta-Analysis of Variation: The Log Coefficient of Variation Ratio." *Research Synthesis Methods* 11: 553–567.

Spake, R., D. E. Bowler, C. T. Callaghan, et al. 2023. "Understanding 'It Depends' in Ecology: A Guide to Hypothesising, Visualising and Interpreting Statistical Interactions." *Biological Reviews* 98: 983–1002.

Spake, R., and C. P. Doncaster. 2017. "Use of Meta-Analysis in Forest Biodiversity Research: Key Challenges and Considerations." *Forest Ecology and Management* 40: 429–437.

Spake, R., A. S. Mori, M. Beckmann, et al. 2021. "Implications of Scale Dependence for Cross-Study Syntheses of Biodiversity Differences." *Ecology Letters* 24: 374–390.

Spake, R., R. E. O'Dea, S. Nakagawa, et al. 2022. "Improving Quantitative Synthesis to Achieve Generality in Ecology." *Nature Ecology & Evolution* 6: 1818–1828.

Törnqvist, L., P. Vartia, and Y. O. Vartia. 1985. "How Should Relative Changes Be Measured?" *American Statistician* 39: 43–46.

Vartia, Y. O. 1976. "Ideal Log-Change Index Numbers." *Scandinavian Journal of Statistics* 3: 121–126.

Viechtbauer, W. 2010. "Conducting Meta-Analyses in R With the Metafor Package." *Journal of Statistical Software* 36: 1–48.

Wantzel, L. 1837. "Recherches Sur les Moyens de Reconnaître si un Problème de Géométrie Peut se Résoudre Avec la Règle et le Compas." *Journal de Mathématiques Pures et Appliquées* 2: 366–372.

Yang, Y., H. Hillebrand, M. Lagisz, I. Cleasby, and S. Nakagawa. 2022. "Low Statistical Power and Overestimated Anthropogenic Impacts, Exacerbated by Publication Bias, Dominate Field Studies in Global Change Biology." *Global Change Biologyw* 28: 969–989.

Yang, Y., M. Lagisz, and S. Nakagawa. 2023. "Decline Effects Are Rare in Ecology: Comment." *Ecology* 104: e4069.

Supporting Information

Additional supporting information can be found online in the Supporting Information section. **Data S1:** ele70335-sup-0001-zip.