

CONCEPTS & SYNTHESIS

From metrics to meaning: diversity as an essentially contested concept

Jonathan A. Newman¹ | Pedro R. Peres-Neto²¹Department of Biology, Wilfrid Laurier University, Waterloo, Ontario, Canada²Department of Biology, Concordia University, Montreal, Quebec, Canada**Correspondence**Corresponding author Jonathan Newman,
Email: jonathannewman@wlu.ca**Abstract**

Biodiversity is among ecology's most widely invoked but least consistently defined concepts. Despite decades of theoretical and methodological development, ecologists continue to disagree about what "diversity" measures and how it should be quantified. These disagreements have direct consequences for conservation science, policymaking, and public engagement. We argue that this enduring debate is not simply a sign of conceptual or methodological failure, but evidence that diversity is an *essentially contested concept* (ECC; *sensu* Gallie 1955). Diversity is internally complex, appraisive, and historically dynamic: it comprises multiple dimensions whose relative importance depends on context and purpose. Using examples from taxonomic diversity, we show that choices among richness, Shannon entropy, Simpson diversity, Hill numbers, and alternative partitioning frameworks involve both substantive commitments about which properties of ecological communities matter and representational commitments about how those properties should be quantified. These choices should be distinguished from technical and statistical consequences, such as sampling sensitivity, estimator bias, and precision, which can often be evaluated empirically once the quantity of interest has been specified. Beyond metrics, contestation also concerns which dimensions of biodiversity should be prioritized and where the boundaries of biodiversity assessment should be drawn. Treating diversity as an ECC helps explain why increasing empirical knowledge and methodological sophistication can clarify these disagreements without necessarily resolving them. It also provides a practical framework for distinguishing disagreements that can be settled through better evidence from those that require explicit normative judgment. Recognizing this distinction calls for greater transparency about the commitments underlying biodiversity assessment, sensitivity to reasonable alternatives when they affect conclusions, and greater clarity in linking ecological measurements to conservation and policy objectives.

KEYWORDS

biodiversity; diversity metrics; essentially contested concepts; normative commitments; Shannon entropy; Simpson index; Hill numbers; alpha-beta-gamma diversity; philosophy of ecology; conservation policy

1 | INTRODUCTION

Globally, we spend about \$140 billion USD per year on biodiversity conservation (Deutz et al. 2020). The Kunming–Montreal Global Biodiversity Framework (KMGBF; 2022) calls for raising this by an additional \$700 billion annually by 2050. The scale of biodiversity investment dwarfs the cost of ending global hunger (Chichaibelu et al. 2023), yet we still struggle to define what we are paying for. Such figures underscore the substantial value society places on biodiversity, but also raise a fundamental question: what, *exactly*, are we conserving and managing?

Definitions such as those offered by the World Health Organization (2025) describe biodiversity as "the variability among living organisms from all sources," encompassing genetic, species (taxonomic), and ecosystem levels, among others—closely following the definition adopted at the Rio Earth Summit (United Nations Conference on Environment and Development 1992). While sufficient as general descriptions, such definitions have

Abbreviations: ANA, anti-nuclear antibodies; APC, antigen-presenting cells; IRF, interferon regulatory factor.

limited use for guiding specific decision-making at any spatial scale. Instead, conservation planning requires biodiversity to be described in ways that are quantifiable, comparable, and sensitive to change over time and space. Nevertheless, ecologists have long pursued empirical approaches to studying biodiversity, developing a wide array of metrics and concepts to quantify its variation across communities and ecosystems. Although these were originally motivated by scientific inquiry, they have become central to conservation planning and policy. Critically, they reflect two key considerations: the *dimensions* of variation being considered (e.g., taxonomic, functional, phylogenetic, genomic) and the *mathematical approaches* used to quantify those dimensions, including richness counts and abundance-sensitive or probability-based indices (e.g., Shannon, Simpson, Hill numbers). Yet these are not merely technical alternatives; they embody *normative commitments* about which aspects of life matter most, how those aspects should be quantified, and how the resulting quantities should be interpreted and valued (Nobles 2025, Pyron and Mooers 2022, Frank 2022). These commitments arise even before particular biodiversity metrics are chosen.

As Magurran (2004) noted, biodiversity measurement typically assumes that all species count equally and, when abundances are considered, that all individuals count equally. These assumptions are not value-neutral: they shape how biodiversity is quantified and interpreted, with consequences for conservation and management priorities. Ecologists provide the conceptual and technical means to measure diversity, but the choice and interpretation of those measures inevitably reflect judgments about which aspects of biodiversity matter most.

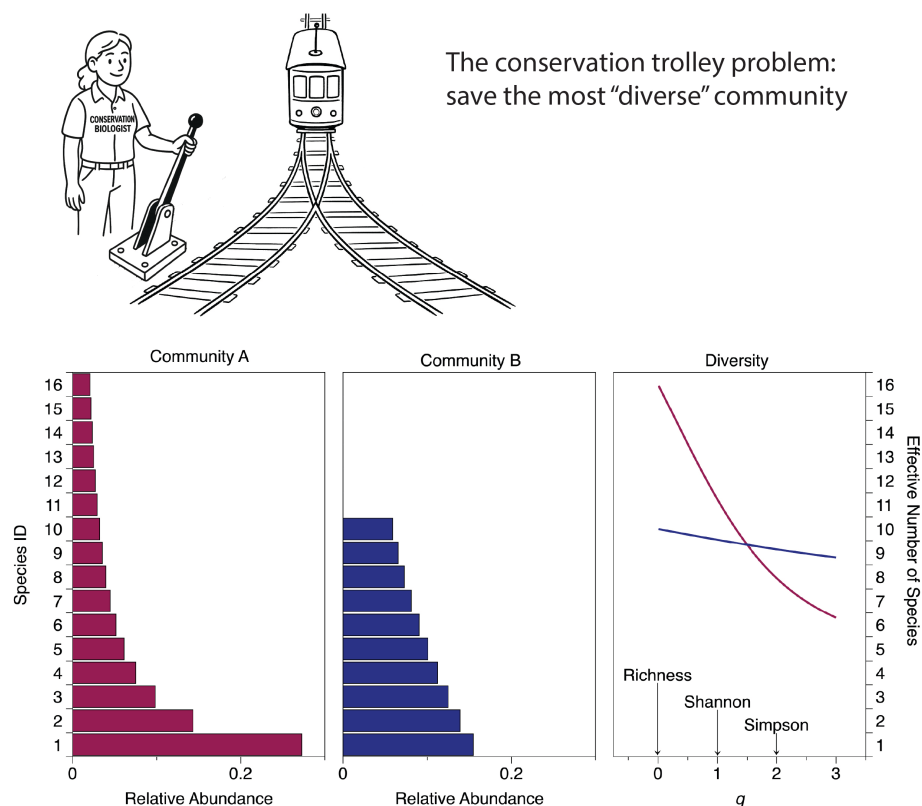


FIGURE 1 A conservation trolley problem (*sensu* Foot 1967, Edmonds 2013). The biologist must recommend which of the two sites should be protected from development. Diversity at each site is shown as the effective number of species, also known as the Hill number (see Section 2.2). The parameter q comes from the Hill-number formula (see Box 2). The choice of q is not determined by the data alone. Different values prioritize different aspects of diversity, so choosing q entails a normative commitment concerning species weighting. When $q = 0$, the index reflects species richness, i.e., the number of different species present in the community. When $q = 1$, Hill diversity equals the exponential of Shannon entropy (Equation (1)), and when $q = 2$, it equals the inverse of Simpson concentration (Equation (3); see Box 2). We used ChatGPT 5.5 to generate the relative abundances and the trolley illustration.

If we are going to invest substantial resources to conserve biodiversity, should we not be able to clearly articulate what we are conserving? The practical implications of this question are anything but abstract. Consider a conservation trolley problem (Figure 1): this schematic illustrates how the evaluation of “diversity” depends on the metric applied. Given constraints in funding, time, and space, a conservation biologist must decide which of two communities (A or B) to protect. The communities differ in their rank–abundance structures (left and centre panels), and their diversity

profiles (right panel) show that Community A has higher species richness but lower evenness, while Community B has fewer species but a more even distribution of abundances. As will be reviewed in Section 2.2 and Box 2, diversity can be expressed using Hill numbers, which vary across a parameter q to reflect different weightings of species abundances. Using this framework, Community A appears more diverse at low q values (where richness dominates), whereas Community B appears more diverse at higher q values (where common species are weighted more heavily). The figure highlights that judgments about which community is “more diverse” are contingent on normative choices about what aspect of diversity matters most.

This ambiguity is not merely hypothetical. KMGBF Target #1 for 2030 calls for “all areas of high biodiversity importance” to be managed to reduce biodiversity loss. Yet the framework itself acknowledges that “there is no single definition” of such areas, leaving it to countries (or other regional authorities) to decide based on their specific contexts and priorities (Convention on Biological Diversity 2022). This admission reveals a central challenge: despite its prominence in global science and policy, biodiversity lacks a universally agreed definition or mode of quantification. Yet conservation decisions are often made without a full understanding of why one metric is chosen over another, or of the conceptual and normative commitments each metric or dimension entails. These underlying assumptions ultimately shape how biodiversity is quantified, valued, and/or prioritized.

In this paper, we argue that biodiversity, and ecological *diversity* more broadly, is best understood as an *essentially contested concept* (ECC), in the sense articulated by Walter Bryce Gallie (1955). Unlike ordinary contested concepts, which may be resolved through clearer definitions or sharper measurements, ECCs remain inherently plural: they admit multiple, mutually incompatible yet reasonable interpretations. The disagreements are not merely semantic but reflect divergent normative commitments about what is valuable. Familiar examples include “fairness” in legal sentencing, “democratic” representation, and the “aesthetic” value of works of art. *Diversity*, we contend, belongs to this same class of enduring conceptual disputes in which scientific definitions are inseparable from underlying ethical and epistemological commitments.

ECCs are concepts whose persistent disagreement arises from internal complexity, evaluative significance, and the availability of multiple reasonable interpretations (Gallie 1955, Collier et al. 2006). We argue that biodiversity exhibits this structure. It is both descriptive, referring to measurable biological variation, and appraisive, because judgments about what aspects of that variation matter can influence conservation and policy decisions. We use the term *normative commitments* rather than the more general term *assumptions* to emphasize that some choices involved in qualifying and quantifying biodiversity are value-laden, reflecting judgments about what should count as relevant variation and how it should be represented and interpreted. We return in Section 3 to the fuller set of features that support treating biodiversity as an essentially contested concept.

We propose that understanding biodiversity as an essentially contested concept clarifies persistent disagreements over its definition and measurement and provides a framework for evaluating alternative biodiversity concepts and metrics. While some points will be familiar to experienced ecologists and conservation practitioners, we have found through research, teaching, and policy work that these conceptual and normative tensions are often under-acknowledged, leading to confusion or misplaced certainty.

Our analysis unfolds in four stages, each building toward the claim that the framework of the essentially contested concept (ECC) brings clarity to enduring debates about biodiversity. First, we examine how ecologists quantify *diversity*, focusing on the substantive and representational commitments embedded in alternative metrics, diversity orders, and partitioning frameworks. Second, we ask whether the persistence and structure of these disagreements support treating *diversity* as an essentially contested rather than an ordinary concept, and extend the analysis beyond metrics to biodiversity dimensions and the boundaries of biodiversity assessment. Third, we consider what follows from this diagnosis for ecological research, conservation, and policy, emphasizing transparency, pluralism, and explicit justification of consequential choices. Fourth, we synthesize these forms of contestation into a framework for distinguishing disagreements that can be resolved empirically from those that require substantive or representational judgment. Box 4 translates this framework into a practical guide for making consequential commitments explicit and reporting the choices through which they enter biodiversity research, conservation assessment, and decision-making.

We wish to emphasize that the claim that biodiversity assessment involves normative judgments is not new (e.g., Noss 1990, Sarkar 2005). A substantial body of work in ecology and conservation science has shown that decisions about what to measure, which data to prioritize, and how to interpret diversity metrics are shaped by underlying values and purposes. Our contribution builds on this foundation by introducing the framework of essentially contested concepts (ECCs) to explain why such disagreements persist.

2 | QUANTIFYING DIVERSITY: UNITY, DIVISION, AND DEBATE

For clarity and focus, we limit our discussion in this section to taxonomic diversity, defined by the relative abundances and identities of species within ecological communities. We acknowledge, of course, that diversity is also commonly conceptualized across other important dimensions, including functional, phylogenetic, and genetic, all of which have been recognized as important for conservation (Pollock et al. 2017). Each of these

adds (potentially) distinct dimensions and often introduces competing criteria for what it means for a system to be *diverse*. We return to some of these additional facets of biodiversity in Section 3.2.

TABLE 1 How a single equation conceals plural meanings. In information theory, *surprise* (or *self-information*) is a measure of the information content of an event. The less probable a species is, the more surprising and informative it is when it occurs in a community. It is defined as $I(x) = -\log_b p(x)$, where $I(x)$ is the self-information of an event, $p(x)$ is the probability of event x , and b is the logarithmic base. Shannon entropy is then given by $H = \sum_{i=1}^S p(x_i) [-\log_b p(x_i)]$, where S is the number of species in a community.

Component	Mathematical Form	Conceptual Role	Ecological Interpretation
Species probability	$p_i = \frac{n_i}{N}$	Relative frequency of species i	Proportion of individuals belonging to species i ; probability of drawing species i .
Negative logarithm of probability	$-\log_b(p_i)$	Surprise or information content	Rare species have high surprise; common species have low surprise.
Product	$p_i \cdot (-\log_b p_i)$	Weighted surprise	Combines likelihood with surprise.
Summation	$\sum_{i=1}^S$	Weighted average of surprises across species	Expected information from a randomly drawn individual.

How we quantify biodiversity shapes both our understanding of ecosystems and the decisions we make about their conservation and management. Yet quantifying diversity is far from straightforward because it inevitably involves normative commitments and judgments about what aspects of biodiversity matter and why. Ecologists have suggested various metrics and frameworks to quantify and compare diversity across different spatial scales, each emphasizing different properties of ecological communities. While these approaches provide mathematical clarity, their differences reflect, at least in part, these underlying judgments about what should count as diversity and how it should be valued or compared. The proliferation of approaches can therefore make it difficult to determine which measure is most appropriate for a particular question, scale, or decision context.

Because some readers may already be familiar with the mathematical foundations of common diversity indices, whereas others may not, we present the more technical exposition in dedicated boxes throughout this section. These boxes provide optional mathematical and pedagogical elaboration while also making explicit how different formulations embody particular normative commitments about what aspects of diversity are emphasized, weighted, or treated as equivalent. Our goal is not to treat the mathematics as ancillary, since many of the normative commitments discussed here emerge directly from the mathematical formulation of the indices themselves, but rather to allow readers to engage with those details at different levels of depth.

To examine the sources of disagreement over taxonomic diversity, we consider four interrelated aspects of how ecologists think about and measure *diversity*. First (§2.1), we consider commonly used metrics and perspectives that reflect not merely methodological variation, but different views of what *diversity* should represent and the normative commitments they embody. Second (§2.2), we explore “effective species counts” (Hill numbers), which offer a common quantitative framework while retaining many of the normative commitments embedded in traditional diversity metrics. Third (§2.3), we revisit alpha (α), beta (β), and gamma (γ) diversity as scale-dependent constructs that embed judgments about meaningful ecological variation. Finally (§2.4), we distinguish disagreements rooted in technical and statistical consequences from those involving substantive or representational commitments, such as how rarity should be weighted or how diversity should scale across communities. Together, these examples reveal a pattern of persistent, structured disagreement whose significance we consider explicitly in Section 3.

2.1 | Our choice of metrics entails normative commitments

Diversity metrics are not value-neutral; they embed normative commitments about how biodiversity should be represented. While often unacknowledged, choosing between richness, Shannon entropy, or Simpson’s index entails implicit conceptual commitments about which species are most relevant; commitments that should be made explicit when interpreting and applying diversity measures.

Perhaps one broadly uncontentious definition of *diversity* is as a form of uncertainty. However, what remains contested is how that uncertainty should be defined and measured, particularly in terms of the weight (value) assigned to each species within a local community or area, a point we clarify further in this section. Even species richness reflects a limited form of uncertainty: it captures how many distinct events (outcomes) are possible (i.e., the number of species), but not how likely each outcome is to occur.

In this section, we examine Shannon, Simpson, and species richness as conceptual exemplars, illustrating how diversity metrics encode distinct normative commitments. We begin with the classic Shannon (entropy-based) diversity index:

$$H = \sum_{i=1}^S p_i (-\log p_i), \quad (1)$$

where p_i is the relative abundance of species i and S is the number of species in an ecological community.

Although ecologists generally recognize that different diversity indices emphasize different aspects of a community, the weighting structures and assumptions underlying these differences are often left implicit. We use Shannon diversity as an illustrative case because it is both widely used in ecology and particularly useful for revealing how such assumptions enter the quantification of diversity. Box 1 briefly unpacks how Shannon entropy links uncertainty, rarity, and species contributions to diversity, while Table 1 decomposes Equation (1) into its constituent operations and makes explicit the assumptions each introduces. Together, they illustrate how normative commitments can be embedded within the mathematical structure of biodiversity metrics.

Box 1. What Shannon entropy assumes about diversity

Normative commitments can be embedded within the mathematical structure of diversity indices. Shannon entropy provides a useful example because its single equation combines several operations that determine how species abundances contribute to diversity.

Shannon entropy originates in information theory, where entropy measures uncertainty about a system's state (Shannon 1948). The *surprise* associated with an event is defined as the negative logarithm of its probability, $-\log p_i$, so rare events are more surprising than common ones. For a fair coin, for example, each outcome has probability $\frac{1}{2}$ and surprise

$$-\log_2 \left(\frac{1}{2} \right) = 1 \text{ bit.}$$

A rarer outcome from a biased coin produces greater surprise.

In ecological applications, species are treated as possible outcomes of a random draw and their relative abundances as probabilities. Shannon entropy is therefore the expected surprise associated with predicting the species identity of a randomly selected individual:

$$H_\alpha = \sum_{i=1}^S \frac{n_i}{N} \left[-\log \left(\frac{n_i}{N} \right) \right], \quad (2)$$

where S is species richness, n_i is the abundance of species i , and

$$N = \sum_{i=1}^S n_i$$

is total abundance. More even communities have greater uncertainty and therefore higher entropy; communities dominated by a few species are more predictable and have lower entropy.

The relevant commitment lies in representing diversity through logarithmic uncertainty. The logarithmic base changes only the units of entropy—for example, bits or nats—and not the ranking of communities, because changing bases multiplies all values by the same constant.

Another classic abundance-based diversity index is the Simpson (1949) index. Simpson's paper, *Measurement of Diversity*, aimed to quantify ecological diversity and, unlike Shannon's work, was not grounded in information theory. Simpson criticized earlier indices of diversity like Fisher's *alpha* (Fisher et al. 1943) for relying on specific assumptions about species abundance distributions, particularly the logarithmic series, which made them sensitive to violations and less generalizable to population-level properties. Note that Shannon entropy, while logarithmic, is non-parametric and assumes no underlying distribution; lower values under a log-series simply reflect the observed data, not a flaw in the index. Simpson's index emphasizes dominance (Whittaker 1965) by measuring the likelihood of repeated identity, i.e., the probability that two randomly drawn individuals belong to the same species. It captures a distinct dimension of diversity: one not based on surprise, as in Shannon's index (Box 1), but still reflecting uncertainty through the prevalence of common species. Lower values indicate greater unpredictability and serve as an inverse indicator of dominance. The Simpson index quantifies diversity using squared relative abundances:

$$\lambda = \sum_{i=1}^S p_i^2, \quad (3)$$

where $p_i = \frac{n_i}{N}$ is the relative abundance of species i (i.e., probability to be found in the community), n_i is the count of individuals of species i , $N = \sum_{i=1}^S n_i$ is the total number of individuals, and S is the total number of species. Since the original Simpson index (Equation (3)) increases as abundance becomes more concentrated among fewer species, ecologists often use its complement, known as the Gini–Simpson index of diversity: $D = 1 - \lambda$. Note that, because Simpson's index is based on a quadratic form (squared probabilities), it places more weight on common species and is less influenced by rare ones. Both Shannon and Simpson diversity therefore entail normative commitments about how species abundances should influence measured diversity, but they differ in the weighting structures through which those commitments are expressed.

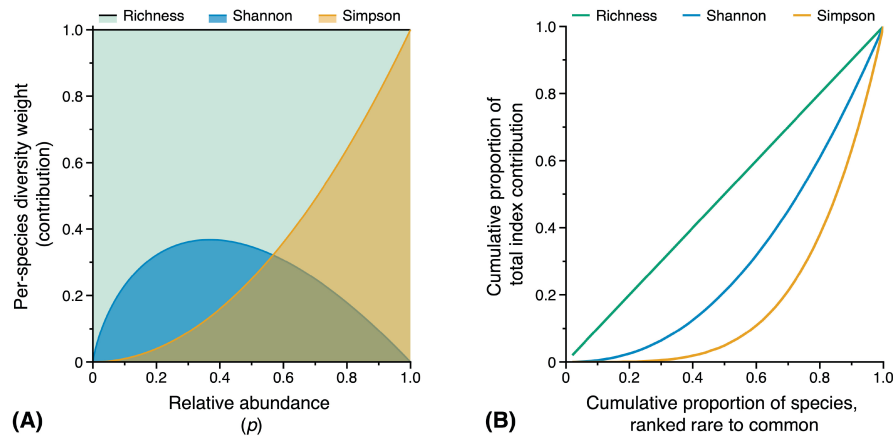


FIGURE 2 Per-species weighting functions underlying richness, Shannon entropy, and Simpson concentration. (A) Per-species contribution as a function of relative abundance (p). Richness assigns equal weight to all species regardless of abundance, Shannon assigns the greatest weight to species of intermediate abundance, and Simpson increasingly emphasizes dominant species. (B) Cumulative proportion of the summed index contributions as species are added from rarest to most common. Richness accumulates evenly across species, Shannon accumulates most strongly through species of intermediate abundance, and Simpson accumulates predominantly through the most abundant species.

While qualitative descriptions help distinguish the normative commitments embedded in diversity indices, quantitative contrasts reveal precisely how those commitments are encoded, showing which species, and which abundance classes, are privileged in the calculation of diversity. As shown in Panel A of Figure 2, both Shannon and Simpson indices can be expressed as per-species contribution functions based on relative abundance $p \in (0, 1]$. Shannon assigns each species a contribution $c(p) = -p \cdot \ln(p)$, which peaks at $p = 1/e \approx 0.37$, reflecting maximal contribution from species of intermediate abundance. Simpson assigns $c(p) = p^2$, a monotonically increasing function that emphasizes dominant species while assigning little weight to rare ones. Although their peaks are not dramatically different in magnitude, the shapes of the contribution curves are fundamentally distinct. Shannon's curve is concave and roughly symmetric, spreading weight across a broad range of intermediate abundances, while Simpson's is convex and steeply increasing, concentrating weight toward dominant species. These differences reveal how deeply each index embeds normative commitments about which species “count” most when measuring diversity.

If conservation priorities were interpreted in terms of the abundance weighting embodied in these indices, Shannon would give greatest per-species influence to species of intermediate abundance ($p \approx 0.37$), whereas Simpson concentration would increasingly emphasize dominant species. These differences become even more apparent when the contributions to the underlying index calculations are accumulated across species ranked from rarest to most common (Panel B of Figure 2). Under richness, biodiversity value accumulates linearly because all species contribute equally regardless of abundance. Simpson, by contrast, concentrates most biodiversity value among dominant species, such that the loss of many rare species may have little influence on the index. Shannon occupies an intermediate position, assigning greater weight to common species while still recognizing contributions from rare species. Consequently, richness, Shannon, and Simpson differ not only in their mathematical formulation but also in the conservation priorities they implicitly support: richness values all species equally, Simpson prioritizes dominant species, whereas Shannon is sensitive to variation across a broader range of species abundances. Although these differences are often described qualitatively in terms of “rare,” “common,” or “dominant” species, they arise directly from the mathematical weighting structures of the indices themselves, making their normative commitments explicit and quantifiable.

2.2 | Effective numbers of species: a common currency for divergent normative commitments

Hill numbers transform Shannon and Simpson indices into effective numbers of species on a common scale. Although this may facilitate comparison across measures, it also assumes that diversity should behave as a countable, scalable quantity, illustrating that no metric is normatively neutral.

In *Diversity and Evenness: A Unifying Notation and Its Consequences*, Hill (1973) introduced a mathematical framework that attempts to unify several (not all) common diversity indices, including richness, Shannon, and Simpson, under a single formulation. The mathematical foundations of Hill numbers, including their derivation from Shannon entropy, their interpretation as effective numbers of equally abundant species, and the

general formulation across diversity orders, are presented in Box 2. Here we focus primarily on their conceptual interpretation and normative implications.

The core innovation is that Hill numbers convert traditional diversity indices into the number of equally abundant species required to produce the same diversity. Although the “effective number of equally abundant species” can yield non-integer values, it offers a comparable measure of diversity that accounts for both richness and evenness. It tells us how many equally abundant species would be needed to match the *diversity* of an uneven community.

However, note that what counts as diversity in this statement is determined by the choice of the parameter, q , in the Hill-number framework (Box 2). The choice of this parameter reflects particular normative commitments. For instance, a community with four species of unequal abundance may have the same Shannon-based Hill number as a perfectly even community of fewer species. This does not mean that the community literally contains fewer species; rather, its diversity behaves as if it did under a particular weighting scheme. The more uneven the distribution, the fewer equally abundant species would be needed to produce the same level of uncertainty or diversity.

It follows that species richness corresponds to Hill number ${}^0D$, as it requires no transformation: its value is simply the count of species, regardless of their relative abundances. In a perfectly even community, richness, the exponential of Shannon entropy (${}^1D$), and inverse Simpson diversity (${}^2D$) are all equal to the number of species. As unevenness increases, however, ${}^1D$ and ${}^2D$ fall below richness, with ${}^2D$ declining more strongly because it places greater weight on common species. This divergence reveals how increasing values of q progressively reduce the contribution of rare species to effective diversity.

For a conservationist normatively committed to ${}^1D$, this means that although a site may contain more species, it is not necessarily more diverse than a site containing fewer but more evenly represented species. Whether this conclusion is desirable depends on the conservation objective. If the goal is to preserve diversity understood as uncertainty in species identity, then evenness becomes part of what is valued. If the goal is to preserve species regardless of abundance, then species richness (${}^0D$) may remain the preferred metric. Hill numbers do not resolve this tension; rather, they provide a common framework for expressing and comparing these alternative abundance-weighting commitments.

Jost (2007) illustrates the interpretive advantage of Hill numbers with a simple example: if a continent loses half of its 30 million equally common species, species richness drops by 50%, while the raw Shannon and Simpson indices change only slightly (17.2 to 16.5 for Shannon and 0.99999997 to 0.99999993 for Simpson). Expressed as Hill numbers, however, the magnitude of the loss becomes immediately apparent. In this particular example, because all species are equally abundant, Hill diversity at both $q = 1$ (Shannon) and $q = 2$ (Simpson) coincides numerically with species richness: each declines from approximately 30 million to 15 million effective species, exactly reflecting the 50% loss. The Hill-number transformation therefore does not change the information contained in the underlying indices, but places it on a scale that allows diversity to be compared in terms of proportional differences. For instance, a community with a Hill diversity of 20 can be interpreted as twice as diverse as one with a Hill diversity of 10 at the same diversity order q .

Each choice of q retains a different weighting of species abundances, reflecting different assumptions about the relative importance of rarity, dominance, and intermediate abundance. The replication principle and its interpretive implications are considered in §2.4. The transformation to Hill numbers rescales diversity onto an effective-number, count-like scale but does not alter the underlying weighting structure of the corresponding indices. For the underlying indices, species contributions to Shannon entropy ($q = 1$) follow $c(p_i) = -p_i \ln p_i$, which peaks at $p = 1/e \approx 0.37$, whereas contributions to Simpson concentration ($q = 2$) follow $c(p_i) = p_i^2$, increasing monotonically with abundance and thus placing greater emphasis on dominant species. Because Hill numbers are strictly increasing transformations of these indices at their corresponding orders q , these contribution functions (Panel A in Figure 2) remain informative about which abundance classes exert greater influence under the Hill-number framework. Thus, expressing diversity as effective numbers changes its interpretation and scale, but not the underlying normative commitments associated with species weighting. The transformations themselves, together with other values of q and the construction of diversity profiles across a continuum of diversity orders, are described in Box 2.

While the species weights used in calculating Shannon entropy and Simpson's index are preserved under the Hill-number transformation, the sensitivity of Hill diversity to changes in species abundances is not identical to that of the original indices. Because the transformation rescales diversity, the magnitude of a change depends partly on the diversity of the community in which it occurs. For example, the Hill version of Shannon diversity (${}^1D$) becomes more sensitive to changes in rare species in already diverse communities because their effects are scaled by total diversity. Thus, transforming an index into an effective number of species preserves its underlying ordering and weighting structure while introducing additional representational consequences for how changes in diversity are expressed and interpreted. Hill numbers can also be extended across spatial scales to quantify α , β , and γ diversity. Doing so introduces a further set of choices concerning how diversity at different scales should be related and interpreted, to which we turn next.

Box 2. Effective numbers of species: the mathematical foundations of Hill diversity

Hill numbers transform traditional diversity indices into the effective number of equally abundant species, placing them on a common, count-like scale. Consider a community with relative abundances 0.7, 0.1, 0.1, 0.1. Its Shannon entropy is

$$H = - \sum_i p_i \ln p_i = - (0.7 \ln 0.7 + 3 \times 0.1 \ln 0.1) \approx 1.03.$$

Because Shannon entropy is expressed in logarithmic units, its value is not directly interpretable as a species count. Exponentiating gives

$$^1D = e^H \approx e^{1.03} \approx 2.8.$$

Thus, this uneven four-species community has the same Shannon diversity as a perfectly even community of approximately 2.8 species. Unevenness reduces effective diversity relative to species richness, which weights all species equally.

The general Hill-number formulation is

$$^qD = \begin{cases} \left(\sum_{i=1}^S p_i^q \right)^{1/(1-q)} & \text{for } q \neq 1, \\ \exp \left(- \sum_{i=1}^S p_i \ln p_i \right) & \text{for } q = 1. \end{cases}$$

Here, p_i is the relative abundance of species i , S is species richness, and q is the diversity order. The value of q determines the influence of relative abundance. At $q = 0$, Hill diversity equals species richness; at $q = 1$, it equals the exponential of Shannon entropy; and at $q = 2$, it is the inverse of Simpson concentration,

$$^2D = \frac{1}{\sum_{i=1}^S p_i^2}.$$

As q increases, common species receive progressively greater influence; values between 0 and 1 give relatively more influence to rare species. Negative values are mathematically possible, but they strongly weight rare species and are highly sensitive to sampling error and incomplete detection of uncommon species.

Plotting ^qD across values of q produces a diversity profile (Figure 1), showing how comparisons among communities change as the weighting of rare and common species changes. Hill numbers therefore provide a common scale without eliminating the weighting commitments of the underlying diversity measures. The framework also generalizes to functional, genomic, and phylogenetic diversity (Chao et al. 2014).

2.3 | Diversity at different spatial scales: partitioning frameworks and their normative foundations

Partitioning biodiversity into α , β , and γ components reflects core commitments about what diversity is and where it matters. Additive frameworks treat diversity as a sum of local contributions, emphasizing site-level diversity, while multiplicative frameworks view diversity as a scaling relationship, highlighting differentiation across space (Legendre and De Cáceres 2013). Choosing between them means deciding whether to emphasize the accumulation of species or the maintenance of ecological distinctiveness across spatial scales.

Ecologists have long recognized that biodiversity can be characterized at different spatial scales, with Whittaker's influential framework distinguishing local alpha (α), among-site beta (β), and regional gamma (γ) diversity components (Whittaker 1972). In formal quantification, α , β , and γ diversity are related within a nested spatial framework, with γ representing regional diversity, α local diversity, and β the differentiation among communities. The preceding discussion has largely concerned diversity within communities. Similar choices arise when diversity is compared across spatial scales, where they become particularly visible in debates over the relationships among α , β , and γ diversity.

A lasting contestation in the diversity literature concerns whether α , β , and γ diversity should be combined additively or multiplicatively (Box 3; Jost 2007, Veech and Crist 2010, Anderson et al. 2011, Pavoine and Ricotta 2024). However, some have argued that the distinction between additive and multiplicative partitioning is a false dichotomy, given the substantial conceptual and mathematical similarities between the two approaches (Veech and Crist 2010). Additive and multiplicative partitioning differ primarily in how they represent relationships among α , β , and γ diversity, rather than in how they weight rare versus common species. Species weighting depends on the diversity measure used within either framework. Partitioning and weighting should therefore be treated as conceptually distinct choices. Partitioning determines whether regional diversity is represented as an additive contribution or a multiplicative scaling relationship, whereas weighting determines how differences in species abundances influence estimates of α , β , and γ diversity. These distinctions are not merely academic. The choice between additive and multiplicative partitioning can affect empirical assessments of spatial diversity and, consequently, conservation priorities. Partitioning therefore represents a further point at which an apparently technical choice can carry normative implications (Baselga 2010, see also below).

The normative commitments underlying diversity partitioning concern assumptions about which ecological processes and spatial scales matter most, and which conservation or management goals should be prioritized. Additive frameworks may be preferred when the aim is to quantify absolute contributions to regional diversity, whereas multiplicative frameworks emphasize proportional differentiation among communities. For example, two sites that each add 10 species make equal additive contributions, but a change from 10 to 20 species represents a doubling whereas

a change from 20 to 30 represents only a 1.5-fold increase. The two approaches therefore foreground different priorities: cumulative contribution versus relative distinctiveness.

Neither framework is universally superior. Additive partitioning emphasizes contributions of particular sites or habitats, whereas multiplicative partitioning emphasizes turnover, heterogeneity, and scaling relationships among communities. The choice between them can therefore influence how biodiversity change is interpreted and how conservation priorities are justified. See [section 3](#) for a broader consideration of these commitments.

Beyond the additive versus multiplicative debate, Whittaker's widely adopted tripartite diversity framework, partitioning diversity into α , β , and γ , has itself faced critique. Among the components of Whittaker's partitioning scheme, β diversity remains especially contentious, despite being the most developed in terms of available metrics. While some view it as a foundational concept for understanding ecological turnover and spatial structure, others criticize it as ambiguous and inconsistently defined across studies. One proposal ([Jurasinski et al. 2009](#)) is to distinguish between two conceptually distinct forms: *differentiation diversity*, based on compositional dissimilarity between samples, and *proportional diversity*, based on how diversity scales across space. This distinction suggests that β diversity may not represent a singular concept, but rather a cluster of interpretations, each shaped by distinct ecological contexts and normative assumptions. This ambiguity is not merely academic: it can directly affect how landscapes are prioritized for conservation. For example, some planning approaches use β diversity to identify complementary sites that together maximize regional coverage of species, while others interpret it as a signal of ecological distinctiveness or turnover that should be protected in its own right ([Tuomisto 2010ab](#), [Anderson et al. 2011](#)).

Finally, it has been further argued that disagreements about biodiversity across scales arise not only from alternative interpretations of α , β , and γ diversity, but also from different conceptions of scale itself. Beyond the distinction between local and regional diversity, scale can be decomposed into grain (the size of the sampling unit), extent (the total area encompassed by an analysis), and focus (the scale represented by each observation or inference; [Scheiner et al. 2000](#), [Whittaker et al. 2001](#), [Scheiner 2003](#)). Different combinations of these components can yield different patterns of diversity and turnover, even when the same organisms and locations are considered. For example, biodiversity measured using nested sampling units may emphasize accumulation within habitats, whereas dispersed sampling units may emphasize differentiation among habitats. These distinctions further illustrate that biodiversity is not a single property waiting to be measured, but a family of related phenomena whose interpretation depends on choices about scale and spatial organization.

Box 3. Diversity partitioning across scales: alternative interpretations of biodiversity

Whittaker's framework partitions biodiversity into local (α), among-community (β), and regional (γ) components. These components can be related additively or multiplicatively, yielding different interpretations of how diversity is distributed across space.

In the additive formulation ([Lande 1996](#)),

$$\gamma = \alpha + \beta, \quad \beta = \gamma - \alpha.$$

Here, β represents the additional diversity obtained when communities are considered together. The framework therefore treats regional diversity as a cumulative quantity composed of within-community diversity and differentiation among communities.

The multiplicative formulation, originally proposed by Whittaker ([1972](#)), instead defines

$$\gamma = \alpha \times \beta, \quad \beta = \frac{\gamma}{\alpha}.$$

Here, β can be interpreted as the effective number of distinct communities required to scale α diversity to regional γ diversity. This formulation follows naturally from Hill numbers, in which diversity is expressed as effective numbers of species (see [§2.2](#); [Hill 1973](#), [Jost 2007](#), [Gaggiotti et al. 2018](#), [Chao et al. 2024](#)). Multiplicative partitioning therefore emphasizes proportional scaling and differentiation among communities rather than additive accumulation.

In both formulations, β is typically derived from estimates of α and γ rather than measured directly. The choice between them consequently affects interpretation as well as calculation: additive partitioning emphasizes absolute contributions to regional diversity, whereas multiplicative partitioning emphasizes proportional distinctiveness and turnover among communities.

2.4 | Distinguishing technical consequences from conceptual commitments

Not every mathematical or statistical property of a diversity measure constitutes a normative commitment, but neither are all apparently technical criteria normatively neutral. Some properties are consequences of an adopted metric, whereas others express substantive judgments about how biodiversity ought to be represented and interpreted.

Distinguishing technical from normative considerations is not always straightforward. Some properties of biodiversity metrics are genuinely technical in the sense that they arise as mathematical or statistical consequences once the dimension (e.g., taxonomic, functional, or phylogenetic diversity) and the metric have been specified. For example, once Shannon diversity is adopted, its sensitivity to incomplete sampling can be evaluated as a technical property of the metric. Others, although formulated mathematically, embody prior judgments about what an adequate measure

of biodiversity should represent; for example, the replication principle assumes that combining two equally diverse but completely distinct communities should double diversity. Distinguishing technical properties from normative commitments is therefore important for determining whether disagreements about biodiversity assessment can be resolved technically or instead reflect differences in how biodiversity should be represented.

Importantly, such commitments need not be consciously adopted. Many enter biodiversity assessment through disciplinary convention: a researcher may use a familiar metric, diversity order, or spatial framework because it is “standard practice” without explicitly endorsing the weighting or representational commitments it entails. Convention can therefore obscure the normative content of a choice without eliminating it. The relevant distinction is not between choices that are consciously value-laden and those that are not, but between choices whose justification follows from the stated purpose of the analysis and those retained primarily as inherited defaults. Making that distinction explicit is part of the transparency we advocate.

This point is particularly relevant to criteria that are often presented as desirable mathematical properties of diversity measures. Jost (2007), for example, proposed several characteristics of an informative diversity metric:

1. It should allow changes in diversity to be interpreted meaningfully, rather than relying solely on statistical significance;
2. It should be expressed as a “numbers equivalent” or effective number of species and satisfy the replication principle, whereby combining two equally diverse but completely distinct communities doubles diversity;
3. It should yield intuitive quantities, such as the number of equally common species or equally distinct communities;
4. It should remain interpretable and consistent across scales (α , β , γ).

These criteria are mathematical in form, but they are not all merely technical requirements that apply after diversity has already been defined. The preference for numbers equivalents and the replication principle, in particular, expresses a substantive view about how diversity should behave as a quantity. Shannon entropy and its corresponding Hill number at $q = 1$ contain the same information and produce the same ordering of communities, but they express that information on different scales. Transforming Shannon entropy into an effective number of species makes ratios directly interpretable: a community with twice the Hill diversity can be understood as having twice the effective number of equally abundant species. The choice of scale therefore determines which quantitative relationships are treated as meaningful. On the entropy scale, additive differences are natural; on the Hill-number scale, multiplicative comparisons and proportional changes become directly interpretable. Preferring one representation over the other is thus not simply a matter of mathematical convenience. It reflects a judgment (normative commitment) about whether diversity should be understood primarily in terms of differences in uncertainty or in terms of effective numbers of species, and about which kinds of comparisons should carry ecological significance. In this sense, the choice of scale itself constitutes a normative commitment about how differences in biodiversity ought to be represented and interpreted.

The same point applies to the preference for multiplicative rather than additive partitioning. Consider two equally diverse and completely distinct communities. Under the replication principle, pooling them doubles diversity. Multiplicative partitioning expresses this doubling as $\beta = 2$: the pooled community contains twice the diversity of an individual community. Additive partitioning instead expresses the same relationship as a difference, $\beta = \gamma - \alpha$, whose value depends on the magnitude of α . Thus, the two approaches represent the same change differently: multiplicative β expresses proportional differentiation, whereas additive β expresses the additional diversity obtained by considering communities together. Choosing between them therefore involves a judgment about which representation of biodiversity is most meaningful. The choice between these formulations therefore concerns more than mathematical convenience. It depends partly on what relationships among α , β , and γ diversity are regarded as conceptually appropriate and ecologically informative. Criteria such as replication, numbers equivalence, and scale consistency are consequently better understood as *representational criteria*: they specify desirable properties of a biodiversity quantity and thereby embody a particular interpretation of what that quantity should mean.

This does not imply that every mathematical or statistical feature of a diversity measure is normatively loaded. A useful distinction remains between substantive choices about the object or representation of measurement and the technical consequences that follow from those choices. For example, Shannon diversity gives relatively greater weight to rare species than Simpson diversity and is consequently more sensitive to incomplete sampling, whereas Simpson diversity downweights rare species and is generally more robust to sampling variation. The decision about how strongly rarity should influence measured diversity concerns what differences among communities are to count most heavily. Once that weighting has been adopted, however, its consequences for sampling sensitivity are technical properties of the metric rather than additional normative commitments.

The same distinction applies to methods for standardizing sampling effort. Equal effort, rarefaction, and coverage-based approaches can produce different statistical behavior and may be more or less appropriate under different sampling conditions (Roswell et al. 2021). Equal-effort sampling may underestimate richness in highly diverse communities (Gotelli and Colwell 2001), while rarefaction can alter abundance-based diversity estimates when many rare species are present. Diversity indices also differ in their sensitivity to which species are detected or missed (Alroy 2025). Questions about estimator bias, sampling variance, precision, and sensitivity are more clearly technical: once the quantity to be estimated

has been specified, these properties can often be evaluated empirically or statistically. They concern how well a chosen measure is estimated rather than, by themselves, what biodiversity ought to mean.

It is therefore useful to distinguish among at least three kinds of considerations. First are explicitly *substantive commitments* concerning which properties of ecological communities should contribute to biodiversity, such as richness, rarity, dominance, functional difference, or compositional turnover. Second are *representational criteria* concerning how biodiversity should behave as a quantitative concept, including numbers equivalence, replication, and additive or multiplicative relationships across scales. Third are *technical and statistical consequences* that arise once a particular quantity and representation have been adopted, including sampling sensitivity, estimator bias, and precision. The boundaries among these categories are not always sharp, because criteria that appear purely mathematical can still reflect judgments about how biodiversity ought to be represented. The relevance of these choices is also purpose-dependent: scientific description, hypothesis testing, conservation prioritization, restoration, and policy may provide different reasons for preferring one otherwise defensible representation over another.

TABLE 2: Normative commitments underlying three commonly used measures of ecological diversity when applied to conservation decisions or policy.

Measure	Embedded Normative Commitments
Species Richness (S)	S-1 Assigns equal contribution to each species in the diversity calculation: presence, rather than relative abundance, determines contribution. S-2 Treats abundance differences as irrelevant to measured diversity: a species represented by one individual contributes as much as one represented by thousands. S-3 Treats species identity as the relevant unit of taxonomic differentiation, leaving differences in abundance, ecological role, and evolutionary distinctiveness outside the measure. S-4 Conservation implication: when used for prioritization, favours areas containing more species, irrespective of their population sizes or ecological roles.
Shannon Entropy (H')	H'-1 Represents diversity as uncertainty in species identity, so both species richness and the distribution of individuals among species contribute to the quantity. H'-2 Assigns little contribution to very rare species, greatest per-species contribution to species of intermediate abundance, and lower contribution again as a species approaches complete dominance, thereby distributing influence across a broader range of abundances than Simpson diversity. H'-3 Represents abundance differences through an information-theoretic framework in which diversity is the expected surprise associated with the species identity of a randomly drawn individual. H'-4 Conservation implication: when used for prioritization, may favour communities combining relatively high richness with relatively even abundance distributions, while allowing rare species to contribute less than more abundant species.
Simpson concentration (λ) / Gini-Simpson diversity ($1 - \lambda$)	λ-1 Gives greater influence to common species because contributions depend on squared relative abundances, making dominant taxa disproportionately influential in the resulting quantity. λ-2 Gives relatively little influence to rare species, so changes involving uncommon taxa may have little effect on measured diversity unless their relative abundances increase. λ-3 Represents diversity in terms of the probability that randomly drawn individuals share, or do not share, species identity, thereby emphasizing dominance and the distribution of abundance among common species. λ-4 Conservation implication: when used for prioritization, may favour communities in which dominance is less concentrated among a small number of species, while giving comparatively little influence to rare species.

This refinement helps locate the sources of contestation more precisely. Some disagreements can be settled by better statistical methods or by demonstrating that one estimator has preferable technical properties for a specified purpose. Others concern which properties a diversity measure should possess or what ecological differences it should represent. Mathematical coherence constrains the available choices, but it does not uniquely determine them. The persistence of multiple defensible ways of representing biodiversity, including different species weightings, diversity transformations, and partitioning frameworks, therefore reveals a pattern of structured disagreement whose significance we consider in the following section.

3 | DIVERSITY AS AN ESSENTIALLY CONTESTED CONCEPT

Diversity is contested not because ecologists have failed to define it, but because they have good reasons to define it differently.

The preceding section demonstrated one level at which biodiversity is contested: even when ecologists agree that taxonomic diversity is the object of interest, alternative approaches to species weighting, quantitative representation, and spatial partitioning embody different substantive

and representational commitments (Tables 2 and 3). Increasing mathematical sophistication can clarify these differences without uniquely determining which representation should be preferred. We argue that this pattern of persistent, structured disagreement is best understood by treating diversity as an *essentially contested concept* (ECC) in the sense articulated by Gallie (1955). We first consider why biodiversity fits the ECC framework, then show that contestation extends beyond the choice among diversity metrics to what properties constitute biodiversity, which forms of biological variation should be prioritized, and where the boundaries of the object being assessed should be drawn.

TABLE 3: Normative commitments underlying different diversity partitioning frameworks. Additive and multiplicative formulations embody distinct views of what diversity represents, how it scales, and what should be prioritized in ecological interpretation and conservation. This view reflects a widely held concern, though others have argued that these frameworks may be more of a conceptual continuum than a true dichotomy.

Partitioning	Normative commitments and assumptions
General α - β - γ partitioning	Diversity is treated as hierarchically structured across spatial scales, and its partitioning reflects assumptions about where diversity resides and which levels are most informative for explanation or conservation. Whittaker's framework, in particular, introduces commitments about how local, among-community, and regional diversity should be distinguished and related. <i>Conservation implication:</i> Decisions may differ depending on whether sites are prioritized because they are locally diverse, compositionally distinct from others, complementary in representing regional diversity, or important for some other stated conservation objective.
Additive framework	Regional diversity (γ) is represented as the sum of average local diversity (α) and additional among-community diversity (β). In this representation, β expresses the absolute amount of diversity added when communities are considered together, thereby foregrounding cumulative contributions to regional diversity. Because additive β is expressed as a difference, its magnitude depends on the scale of α and γ and is less directly comparable across systems that differ substantially in local diversity. <i>Conservation implication:</i> This framework is particularly suited to questions concerned with absolute contributions to regional biodiversity. Its sensitivity to rare or common species is analytically separate and depends on the diversity measure used within the partitioning framework.
Multiplicative framework	Regional diversity (γ) is represented as the product of local diversity (α) and among-community differentiation (β). Under a Hill-number formulation, β can be interpreted as the effective number of distinct communities required to scale average local diversity to regional diversity. This representation foregrounds proportional differentiation and scaling rather than absolute accumulation. <i>Conservation implication:</i> This framework is particularly suited to questions concerned with proportional differentiation, distinctiveness, or heterogeneity among communities. Its sensitivity to rare or common species depends on the diversity measure or diversity order used; within a Hill-number formulation, this sensitivity is controlled by q.
Choice of partitioning scheme	Reflects a representational commitment about whether relationships among α, β, and γ diversity are most meaningfully expressed as absolute differences or proportional scaling. Additive and multiplicative formulations can describe the same underlying diversity pattern while making different numerical relationships salient. The choice therefore concerns how biodiversity differences across scales are represented and interpreted rather than how rare and common species are weighted. <i>Conservation implication:</i> Different representations may support different interpretations of contribution, distinctiveness, or change across spatial scales. This choice is analytically distinct from the choice of species weighting, and the two should be justified separately.

3.1 | Why biodiversity fits the ECC framework

Gallie's account was intended to characterize concepts for which persistent disagreement is not simply the result of confusion, insufficient evidence, or defective definition. Such disagreement has a recognizable structure. Biodiversity aligns closely with Gallie's diagnostic features of an ECC. It is (i) *appraisive*, since assessments of biodiversity can carry judgments about ecological value or significance; (ii) *internally complex*, comprising multiple interacting properties including richness, evenness, composition, scale, and different dimensions of biological variation; (iii) admits *multiple descriptibility*, because no single metric or dimension exhausts its meaning and multiple legitimate formulations coexist; (iv) exhibits *reciprocal recognition*, as proponents of different approaches acknowledge alternatives while defending their own; (v) displays *historicity*, evolving alongside shifts in ecological theory, methodology, and conservation priorities; (vi) encourages *progressive competition*, whereby disagreement generates conceptual and methodological innovations such as Hill numbers and multidimensional diversity frameworks; and (vii) remains *open to modification*, adapting to new dimensions, conceptual frameworks, and conservation objectives (Gallie 1955). Together, these features characterize not simply the existence of disagreement, but its particular structure.

Several aspects of this diagnosis are especially important in the present context. Biodiversity is not merely descriptive. In conservation and environmental policy, assessments of what constitutes “high,” “important,” or “priority” biodiversity can inform decisions about protection, restoration, and resource allocation. This does not mean that every calculation of a diversity index is itself a normative judgment, but differences in how biodiversity is constituted and represented can become evaluatively consequential.

Biodiversity’s internal complexity also permits multiple defensible representations. As Section 2 showed, commonly used metrics and partitioning frameworks organize biological variation differently rather than simply providing more or less precise estimates of a uniquely defined quantity. The resulting disagreements have also been productive: frameworks such as Hill numbers and diversity profiles clarify relationships among alternatives and expose assumptions that might otherwise remain implicit, without eliminating the plurality they reveal.

The historical development of the concept reinforces this point. What ecologists and conservationists include under biodiversity has expanded from a predominant focus on species richness to increasing attention to functional, phylogenetic, genetic, and other dimensions of biological variation. The concept has therefore not simply converged on an increasingly precise definition; its scope and interpretation have evolved as new biological properties, methods, and conservation purposes have become salient.

These features help explain why the disagreements identified in Section 2 persist even after their technical consequences are clarified. The ECC claim is not that every question about diversity is irresolvable, but that reasonable disagreement can remain over which properties, dimensions, boundaries, and representations should constitute biodiversity for a given purpose. The following subsections show that this contestation extends beyond alternative measurements of an agreed object to the constitution and scope of the object itself.

3.2 | Contestation over biodiversity dimensions

The question is not only how diversity should be measured, but what kind of diversity matters.

One indication of biodiversity’s internal complexity is that disagreement extends beyond how a particular dimension should be measured to what forms of biodiversity matter in the first place. Although taxonomic diversity has historically dominated ecological research and conservation practice, biodiversity can also be understood in terms of functional traits, phylogenetic history, genetic variation, ecological interactions, and other dimensions of biological organization (Faith 1992, Díaz et al. 2007, Mace et al. 2012). These are not simply alternative ways of estimating the same underlying property. They characterize different forms of biological variation and can reflect different conceptions of what biodiversity is and why it is valuable (Mace et al. 2012, Scheiner et al. 2025).

These dimensions can therefore produce contrasting assessments of ecological importance. Taxonomic diversity emphasizes the variety of species and is commonly associated with species persistence and extinction prevention. Functional diversity focuses on differences in ecological roles and ecosystem functioning (Díaz et al. 2007), phylogenetic diversity emphasizes evolutionary history (Faith 1992), and genetic diversity highlights adaptive potential and long-term resilience (Frankham 2005). A community that appears highly diverse taxonomically may exhibit relatively little functional diversity, while a species-poor community may contain unique functional strategies or disproportionately large amounts of evolutionary history. Correspondingly, a conservation strategy focused on species richness may prioritize different places from one focused on functional or phylogenetic diversity. The disagreement is therefore not necessarily over which measure most accurately captures a common quantity, but over which biological properties should be regarded as relevant to biodiversity in the first place.

Contestation can also occur at several nested levels. Ecologists may disagree *between dimensions* about whether taxonomic, functional, phylogenetic, genetic, or other forms of variation should be prioritized. Even when they agree on the dimension of interest, they may disagree about how it should be quantified. Functional, phylogenetic, and genetic diversity can themselves be represented using richness-based, Shannon-type, Simpson-type, or Hill-number formulations (Chao et al. 2014), reproducing many of the weighting and representational choices discussed in Section 2.

Some dimensions introduce an additional choice before a diversity metric is even applied. Functional diversity, for example, depends on which traits are taken to represent ecologically relevant differences among organisms. Selecting body size, trophic position, dispersal capacity, phenology, or other traits can foreground different aspects of ecological function and thereby change the biological variation represented by the analysis. Researchers may therefore disagree about which dimension matters, how that dimension should be biologically instantiated, and how the resulting variation should then be weighted and represented quantitatively.

Attempts to integrate multiple dimensions do not remove this plurality. Combining taxonomic, functional, phylogenetic, and genetic information requires further decisions about how different dimensions should be weighted and interpreted, and recent syntheses suggest that these dimensions often quantify fundamentally different properties rather than interchangeable expressions of a single underlying phenomenon (Scheiner et al. 2025, Cadotte et al. 2013). Increasing dimensionality therefore does not necessarily move biodiversity science toward a single definitive measure; it can instead multiply the defensible perspectives from which biodiversity is characterized.

This nested structure of disagreement is important to the ECC diagnosis. If disagreements concerned only competing estimators of a clearly specified underlying quantity, improved theory and data might reasonably be expected eventually to identify a preferred approach. But the choice among biodiversity dimensions concerns what properties are constitutive of the object being evaluated. Greater empirical knowledge can reveal the consequences of prioritizing taxonomic, functional, phylogenetic, or genetic variation, but it cannot by itself establish which of those forms of variation should take precedence. The multiplicity of dimensions therefore exemplifies both biodiversity's *internal complexity* and its *multiple descriptibility*: biodiversity is better understood as a family of related properties whose relevance depends partly on the ecological questions and conservation purposes at issue.

3.3 | Contestation over the boundaries of biodiversity

To count only what is observed is already to make a choice about what counts.

Essential contestability is also evident in disagreements over the boundaries of the entities and possibilities included in biodiversity assessment. What constitutes an ecological community, for example, remains unsettled (see e.g., [Garcia and Newman 2016](#), [Newman 2025](#), [Ricklefs 2008](#)). Communities may be defined taxonomically, functionally, spatially, or interactionally; sampling units differ in size and individuation; and communities themselves change through time. These are not merely technical decisions about how biodiversity is measured. They help constitute the object whose diversity is being assessed.

A spatial boundary determines whether variation among nearby assemblages contributes to the biodiversity under consideration; a temporal boundary determines whether turnover through time is treated as biodiversity or as change between biodiversity states; and a taxonomic or functional boundary determines which organisms or forms of variation are regarded as relevant in the first place. The choice of temporal baseline entails a related commitment, as it determines which states serve as reference points for interpreting loss, recovery, or change. Empirical information can describe these systems in increasing detail, but it cannot by itself determine which boundaries define the appropriate object of biodiversity assessment.

The distinction between empirical uncertainty and conceptual scope becomes particularly important because ecological knowledge is systematically incomplete. Biodiversity assessment is constrained by the *Linnean shortfall* (undescribed species), the *Wallacean shortfall* (incomplete knowledge of species distributions), the *Darwinian shortfall* (limited understanding of traits, interactions, and evolutionary relationships), and related gaps in functional, phylogenetic, and genetic knowledge ([Hortal et al. 2015](#), [Diniz-Filho et al. 2013](#)). Discussions of "shadow diversity" likewise emphasize that unknown, overlooked, and poorly studied organisms may remain effectively invisible to conservation assessment despite their ecological significance ([Turton-Hughes et al. 2024](#)). These shortfalls are fundamentally epistemic: they concern what we know about biodiversity. Better data can therefore reduce at least some of the resulting uncertainty.

Yet epistemic limitations acquire normative significance when what is known or observable determines what is allowed to count. Restricting assessment to well-described species, documented distributions, measured traits, or readily detectable organisms implicitly privileges biological variation that is accessible to current knowledge and methods. The empirical fact that some components of biodiversity are better known than others does not itself establish that they should matter more. Here, the epistemic and conceptual problems interact: incomplete knowledge affects what can practically be represented, while judgments about scope determine whether poorly known components are treated as relevant to the object of assessment.

The concept of *dark diversity* makes the distinction especially clear. Species that are present in a community but remain undetected because of imperfect sampling create an epistemic problem: better observation can, at least in principle, improve our estimate of biodiversity that is actually present. Species that are absent locally but belong to the broader species pool and could occur under suitable ecological conditions raise a different question. Their inclusion concerns the *scope* of biodiversity assessment itself: should biodiversity refer only to realized biological variation, or should ecologically plausible but unrealized possibilities also count?

This latter question cannot be settled simply by improving sampling. Including potential members of a community treats unrealized ecological possibilities as relevant to biodiversity evaluation, whereas excluding them restricts biodiversity to realized states. In this respect, debates about dark diversity parallel discussions of future generations in environmental ethics: in both cases, entities absent from the present object of evaluation may nonetheless be regarded as relevant to present judgments. The analogy need not imply that potential species and future persons have equivalent moral status. Its significance is structural: both cases require a judgment about whether present evaluation should extend beyond what presently exists.

Knowledge shortfalls and concepts such as shadow and dark diversity therefore reveal two importantly different limits to biodiversity assessment. Some concern how accurately we know the biodiversity that exists and may be reduced through better data; others concern what the scope of biodiversity should include. The latter illustrate why disagreement over biodiversity can persist even when empirical uncertainty is reduced.

Disagreements over spatial, temporal, biological, and counterfactual boundaries therefore provide further evidence that biodiversity admits multiple defensible specifications rather than awaiting discovery of a uniquely correct boundary.

3.4 | Clarifying the claim and anticipating objections

If a concept remains contested after its objections are faced, the contest may be part of its nature rather than a defect in its use.

The preceding analysis strengthens the case for treating biodiversity as essentially contested, but it also raises several objections to that diagnosis. Critics of Gallie's framework have questioned whether essentially contested concepts are genuinely distinctive, whether the framework encourages relativism, and whether persistent disagreement might instead reflect vagueness or conceptual immaturity (cf. Douglas 2009). These concerns are particularly important in a scientific context, where many disagreements are appropriately expected to diminish as evidence and methods improve.

One concern is that labeling a concept as "essentially contested" invites a form of relativism in which any interpretation is as good as any other (e.g., Gellner 1974, MacIntyre 1973). If no single definition can be decisive, then the concept may appear to lose its analytical usefulness. However, this interpretation mischaracterizes the ECC framework. Contestation does not imply that all interpretations are equally defensible. Empirical evidence, mathematical consistency, conceptual coherence, and the purposes of an inquiry can all constrain the range of reasonable interpretations. The claim is instead that more than one interpretation may remain defensible within those constraints because competing approaches embody different judgments about what features of the concept matter. As shown in Section 2 and in the preceding subsections, alternative diversity metrics, dimensions, and boundaries can each be coherent while foregrounding different aspects of biological variation. The task is therefore not to accept all interpretations indiscriminately, but to make the grounds for preferring among them explicit and open to critical evaluation.

A second, and more substantive, criticism is that many so-called ECCs are not inherently contested, but simply vague, underspecified, or awaiting better definition (e.g., Gray 1977, Connolly 1993, Waldron 1994). On this view, disagreements about diversity might eventually be resolved through improved data, clearer terminology, or more refined metrics. Some disputes about biodiversity undoubtedly have this character. Imperfect detection, uncertain species distributions, and poorly estimated abundances are empirical problems, while ambiguities in terminology can sometimes be removed by stipulation. Treating diversity as an ECC should not obscure these resolvable sources of disagreement.

The stronger claim concerns what remains after such problems have been addressed. Increasing mathematical sophistication, improved sampling methods, and unifying frameworks such as Hill numbers can clarify relationships among measures and resolve genuinely technical disagreements, but they do not determine how rarity and dominance should be weighted, which biodiversity dimensions should be prioritized, which traits should instantiate functional difference, or which spatial, temporal, and biological boundaries should define the object of assessment. In such cases, further evidence can reveal the consequences of alternative choices without uniquely determining which choice should be made. Persistent disagreement under conditions of substantial empirical and theoretical refinement is therefore more consistent with essential contestability than with conceptual immaturity alone.

A third concern is that ECCs are typically understood as *appraisive* concepts—terms that inherently involve evaluative judgments (Gallie 1955, MacIntyre 1973)—and that not all scientific concepts meet this criterion. Biodiversity can certainly be used descriptively: an ecologist may calculate a diversity index without making any explicit claim about what ought to be valued or conserved. But biodiversity also plays an evaluative role in conservation science and environmental policy. Evaluations of biodiversity inform decisions about protection, restoration, and resource allocation. Diversity thus functions both as a description of biological variation and as a basis for judgments about what ought to be preserved. This evaluative role does not convert every empirical use of diversity into a conservation judgment, but it helps explain why disagreements over its constitution and representation can become normatively consequential.

Finally, Gallie's criteria need not be treated as a rigid checklist (see discussion in MacIntyre 1973). Our claim does not depend on demonstrating that biodiversity satisfies every condition in precisely the same way as Gallie's original examples. Rather, the ECC framework is diagnostic. It identifies a recognizable pattern: biodiversity is internally complex and evaluatively consequential; it permits multiple defensible descriptions; competing interpretations are mutually recognized and can stimulate conceptual innovation; and the concept remains open to revision as ecological knowledge and conservation purposes change. Most importantly, disagreement persists not only because biodiversity is difficult to measure, but because different reasonable judgments can be made about what constitutes it and how it should be represented.

Treating biodiversity as essentially contested therefore does not imply that its meaning is unconstrained or that scientific progress is futile. It distinguishes disagreements that additional evidence or technical refinement can resolve from those in which evidence can constrain the alternatives and clarify their consequences without necessarily selecting a uniquely correct one. The practical implication is not that ecologists should abandon efforts to define or measure biodiversity precisely, but that precision alone cannot eliminate the substantive choices embedded in those definitions and measures. Recognizing those choices as part of the concept's contestability provides the basis for making them more explicit, transparent, and open to scrutiny.

Box 4. Illustrative reporting of consequential commitments

Biodiversity assessments involve choices concerning what is measured, how it is represented, and the purposes it serves. These choices need not be justified exhaustively; transparency is especially important when reasonable alternatives could materially affect conclusions.

Imagine that Figure 1 represented a real-world conservation problem. The following statements illustrate how the study might be reported in its Methods and Discussion sections, first without making consequential commitments explicit and then with greater transparency.

A. Reporting without making consequential commitments explicit

Methods: "We compared species diversity between the two communities using Hill numbers at $q = 1$."

Discussion: "Community A had greater diversity than Community B and should therefore be prioritized for conservation."

B. Reporting with consequential commitments made explicit

Methods: "Our objective was to compare the two communities in Figure 1 to recommend which should be conserved to maximize the diversity preserved. We chose to interpret diversity in this context as taxonomic diversity at the species level. However, neither this interpretation nor the stated conservation objective uniquely determines how rare and common species should be weighted. We therefore calculated species richness, Shannon entropy, and the Simpson index, which embody different weightings of species abundances. We then expressed these measures as effective numbers of species using the corresponding Hill-number transformations ($q = 0$, $q = 1$, and $q = 2$, respectively), allowing comparisons on a common, count-like scale. Each assemblage was treated as a distinct local community, with diversity assessed using the species and abundances as recorded."

Discussion: "The comparison revealed that Community A ranked higher at $q = 0$ and $q = 1$, whereas Community B ranked higher at $q = 2$. Thus, the identity of the 'most diverse' community depends on normative choices concerning both what constitutes diversity and how its components should be weighted—choices that neither the conservation objective nor the abundance data alone can resolve. Consequently, the diversity comparisons alone do not establish which community should be prioritized for conservation."

The first example implicitly treats one particular representation of diversity as sufficient to determine a conservation priority. The second identifies the relevant choices, examines reasonable alternatives, and distinguishes what the calculations establish from what remains normatively unresolved. Transparency need not require exhaustive justification of every methodological decision, but should make consequential commitments and their implications explicit.

Appendix S1 provides a comprehensive guide to identifying consequential commitments and reporting their implications. We argue that authors should identify and justify consequential analytical choices in the Methods section, then explain how those choices affect the interpretation of results and the limits of ecological inference in the Discussion. The guide provides a practical framework for doing so.

4 | ENGAGING WITH DIVERSITY AS AN ECC

To understand diversity is to confront contestation; our task is not to eliminate it but to reason well within it

The debate over how to quantify diversity is not new. More than fifty years ago, Stuart Hurlbert (1971) dismissed "species diversity" as a *non-concept*, arguing that it conflated distinct components such as richness and evenness into a single ambiguous measure. The ECC framework suggests a different interpretation. Rather than treating this plurality as evidence that diversity is conceptually defective, it recognizes persistent disagreement as a consequence of the multiple legitimate ways biodiversity can be constituted, represented, and valued.

The practical question is therefore not how to eliminate contestation, but how to work responsibly within it. If diversity is not reducible to a single uncontested quantity, ecologists and conservation scientists need ways to make consequential choices explicit, assess reasonable alternatives, and connect biodiversity measures clearly to the purposes they serve. The following five practices offer concrete ways to do this.

Embrace pluralism.

Recognizing diversity as essentially contested fosters pluralism instead of reductionism. Rather than searching for a single "best" measure, conservation science can benefit from complementary dashboards that track multiple dimensions, each aligned with specific management goals. This pluralism legitimizes different perspectives, such as safeguarding evolutionary history, promoting ecosystem resilience, or ensuring equitable access to nature's benefits, while clarifying that no single metric can capture them all. By acknowledging the coexistence of competing but legitimate interpretations, ecologists can better tailor their analyses to the relevant questions and values.

Treat contestation as a source of innovation.

Ecologists and conservation biologists should view contestation as a source of innovation. Disputes over how to measure or prioritize diversity have historically driven methodological innovation, from the development of Hill numbers to the rise of eDNA-based indices. Following this path, researchers should approach conceptual disagreements not as flaws but as opportunities to advance theory, refine tools, and improve conservation practice. Embracing contestation expands the toolbox of methods rather than narrowing it.

Educate experts and engage the public.

Recognizing diversity as an ECC also calls for changes in how we teach and communicate. For experts, training in biodiversity science should include philosophical and normative literacy and not only how to calculate indices, but why particular metrics reflect specific value commitments (e.g., Frank 2022). For the broader public, including decision- and policy-makers, diversity metrics should be communicated not as purely technical outputs, but as tools shaped by priorities that are open to discussion. Framing diversity as a contested but constructive concept empowers learners and stakeholders alike to engage more critically, fostering informed dialogue about what should count in conservation and why.

Make better policy decisions.

Treating diversity as an ECC also has implications for the science-policy interface. Conservation recommendations often carry moral weight and political consequences. Recognizing the contested nature of diversity ensures that policy decisions are not presented as the inevitable result of a single, value-free calculation but rather as reasoned judgments that balance different dimensions and commitments. This honesty does not undermine conservation; it strengthens its legitimacy by aligning scientific practice with democratic ideals of transparency, deliberation, and accountability.

Make normative commitments transparent.

Ecologists should be transparent about the normative commitments inherent in how they quantify diversity. Transparency requires more than simply naming the metric that was used. Where choices are consequential, researchers should explain why the biological properties represented by a metric are relevant to the question being asked, what the choice commits the analysis to, and the limits it places on interpretation. A study using species richness, for example, should make clear that species are weighted equally and that differences in abundance structure or evenness are consequently outside the representation. Similarly, use of a Hill number should distinguish the weighting determined by q from the decision to represent diversity as an effective number of species.

Box 4 develops these into a practical guide for reporting consequential choices. Together, these practices shift the aim from resolving contestation to making its sources, consequences, and reasonable alternatives explicit.

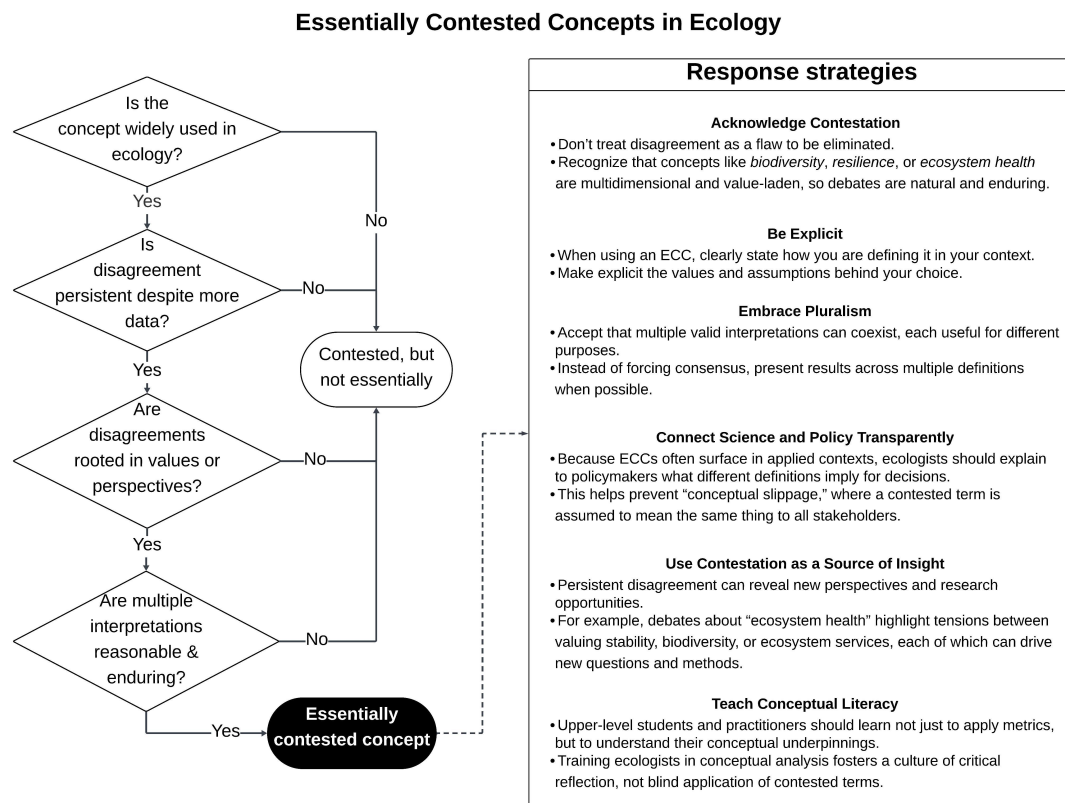


FIGURE 3 Identifying and responding to essentially contested concepts in ecology.

5 | SYNTHESIS AND CONCLUSIONS

Clarity is not achieved by eliminating disagreement, but by understanding what the disagreement is about.

The analysis developed here suggests that disagreements about biodiversity arise at different levels and therefore require different kinds of resolution. Some concern substantive questions about which properties of ecological systems should count as biodiversity, including species richness, abundance structure, functional difference, evolutionary history, or potential community membership. Others concern how those properties should be represented quantitatively, including species weighting, effective numbers, and additive or multiplicative relationships across scales. Still others are genuinely technical, involving sampling, estimation, bias, precision, and uncertainty. These distinctions matter because only the last category can, in principle, be resolved primarily through better data or improved statistical methods.

TABLE 4: Sources of contestation in biodiversity and a heuristic taxonomy of the choices through which particular conceptions of diversity are specified. The table summarizes the principal forms of contestation discussed in this paper and illustrates why additional empirical knowledge alone cannot resolve disagreements concerning biodiversity.

Dimension of contestation	What is contested	Why evidence alone is insufficient, and why it matters
Weighting of elements <i>Richness, Shannon entropy, Simpson diversity</i>	Whether all species contribute equally, common species contribute more, or some intermediate weighting scheme is preferred	Abundance estimates do not determine how abundance should influence diversity. Different weighting schemes can therefore change which sites are prioritized.
Spatial scale α , β , and γ diversity	Whether biodiversity should be assessed within communities, among communities, or across larger regions	Sampling cannot determine which spatial level matters most. Conservation may prioritize locally rich, compositionally distinct, or regionally diverse systems.
Scale representation <i>Additive vs. multiplicative partitioning</i>	Whether diversity across spatial levels is represented as accumulation or proportional differentiation	Evidence does not determine what β diversity should represent. Different frameworks emphasize accumulation, differentiation, representativeness, or complementarity.
Biological dimension <i>Taxonomic, functional, phylogenetic, genetic diversity</i>	Which kinds of biological differences should count as biodiversity	Data can estimate each dimension but cannot determine which should take precedence. Different choices emphasize species, functions, evolutionary history, or adaptive potential.
Unit and boundaries <i>Alternative definitions of ecological communities</i>	Which organisms belong to the relevant ecological unit and where its boundaries should be drawn	Observations do not uniquely determine community boundaries. Different units of analysis can produce different biodiversity estimates and comparisons.
Membership <i>Dark diversity and biodiversity shortfalls</i>	Whether biodiversity includes only observed elements or also unrealized, undetected, or potential components	Sampling can reduce uncertainty but cannot determine whether unrealized possibilities should count. Their inclusion may alter restoration and conservation priorities.
Temporal frame <i>Reference conditions and restoration targets</i>	Whether biodiversity should be evaluated relative to present states, historical conditions, or future potential	Evidence can describe alternative states but cannot determine the appropriate benchmark. Different temporal frames alter assessments of degradation and recovery.
Purpose of assessment <i>Prioritization and biodiversity targets</i>	Whether biodiversity is valued primarily as representation, distinctiveness, functioning, resilience, heritage, or some other conservation objective	Evidence informs the consequences of alternative goals but does not determine which goal should prevail. Different objectives can favor different conceptions and metrics.

The framework of *essentially contested concepts* helps explain why the remaining disagreements persist. If biodiversity were merely vague, poorly measured, or conceptually immature, increasing empirical knowledge and methodological sophistication should progressively eliminate disagreement. Instead, these advances often clarify the alternatives without determining which should be preferred. Competing interpretations can remain reasonable because they embody different judgments about what biological variation matters, how it should be represented, and which comparisons are ecologically or normatively significant. These sources of contestation can also be understood as dimensions along which particular *conceptions* of diversity are specified. If *diversity* is the overarching concept, then any particular conception of it reflects choices about what kinds of biological differences count, how they are weighted, the spatial and temporal scales over which they are evaluated, the boundaries of the system under consideration, and the purposes for which diversity is being assessed. Table 4 synthesizes these dimensions of contestation and shows where empirical uncertainty ends and substantive or representational disagreement begins.

Recognizing this distinction changes the appropriate response to disagreement. Technical disputes should be addressed with better evidence and methods. Substantive and representational disputes require something different: the commitments underlying a biodiversity assessment should be made explicit, linked to the purpose of the analysis, and examined for sensitivity to reasonable alternatives when those alternatives could alter

a description, ranking, or recommendation (Box 4). This is not a retreat from scientific rigor. It is a more precise account of what evidence can establish and where additional judgment is unavoidable. Figure 3 summarizes this diagnostic and practical logic.

The implications extend beyond the choice of diversity metrics. Biodiversity enters conservation planning, environmental assessment, restoration, public policy, and international governance as though it were often a single object of measurement, yet different actors may reasonably understand its relevant dimensions, boundaries, and purposes differently. Recognizing biodiversity as essentially contested therefore places a premium on conceptual transparency: policymakers, scientists, communities, and other affected groups need to know not only what biodiversity measure has been used, but what conception of biodiversity that measure represents and why it is appropriate to the decision at hand.

Treating diversity as an essentially contested concept does not imply that all definitions or measures are equally good, nor that empirical evidence is incapable of constraining choices. Rather, it clarifies the limits of what evidence alone can decide. The central task is therefore not to eliminate plurality, but to distinguish disagreements that can be resolved empirically from those that require explicit normative judgment. Doing so can make biodiversity science more transparent about its commitments, more precise about the meaning of its measurements, and better equipped to connect ecological knowledge to conservation and policy decisions.

AUTHOR CONTRIBUTIONS

Both authors contributed equally to all aspects of this paper.

ACKNOWLEDGMENTS

The authors thank Gary Varner for pointing us toward Gallie's essentially contested concepts. JAN works on the shared traditional territory of the Neutral, Anishnaabe, and Haudenosaunee peoples. These Nations are the traditional stewards, since time immemorial, of the lands and water on which Wilfrid Laurier University is now present. PRPN works on the unceded territory of the Kanien'kehá:ka Nation, who are recognized as the custodians of the lands and waters on which Concordia University is located. These acknowledgements represent our shared responsibility and commitment to learn about and confront colonial legacies and to earnestly engage in the unfinished work of reconciliation. This work was supported by grants from the Canadian Natural Science and Engineering Research Council (NSERC) to JAN (Discovery Grant Program, RGPIN-2022-04581, Grass-*Epichloë* Interactions in Current and Future Climates) and PRPN (Discover Grant Program, RGPIN-2020-07188, Ecology in Spatially Structured Environments); PRPN is also supported by the Canada Research Chairs program (CRC-2023-00139, Spatial Ecology and Biodiversity). We used ChatGPT 5.5 to generate the synthetic example in Figure 1 and to test some of our ideas, and we used Grammarly to improve the writing in some places.

FINANCIAL DISCLOSURE

None reported.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

REFERENCES

- Alroy, J. (2025) Does evenness even exist? *Ecology Letters*, 28(7), e70181. doi:10.1111/ele.70181.
- Anderson, M.J., Ellingsen, K.E. & McArdle, B.H. (2011) Navigating the multiple meanings of β diversity: A roadmap for the practicing ecologist. *Ecology Letters*, 14(1), 19–28. doi:10.1111/j.1461-0248.2010.01552.x.
- Baselga, A. (2010) Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, 19(1), 134–143. doi:10.1111/j.1466-8238.2009.00490.x.
- Cadotte, M.W., Albert, C.H. & Walker, S.C. (2013) The ecology of differences: Assessing community assembly with trait and evolutionary distances. *Ecology Letters*, 16(10), 1234–1244. doi:10.1111/ele.12161.
- Chao, A., Chiu, C.H., Hu, K.H., van Der Plas, F., Cadotte, M.W., Mitesser, O. et al. (2024) Hill–chao numbers allow decomposing gamma multifunctionality into alpha and beta components. *Ecology Letters*, 27(1), e14336. doi:10.1111/ele.14336.
- Chao, A., Chiu, C.H. & Jost, L. (2014) Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through hill numbers. *Annual Review of Ecology, Evolution, and Systematics*, 45(1), 297–324. doi:10.1146/annurev-ecolsys-120213-091540.
- Chichaibelu, B.B., Bekchanov, M., von Braun, J. & Torero, M. (2023) The global cost of reaching a world without hunger: Investment costs and policy action opportunities. In: von Braun, J., Afsana, K., Fresco, L.O., Hag, M. & Hassan, A. (Eds.) *Science and Innovations for Food Systems Transformation* Springer, pp. 625–660.
- Collier, D., Daniel Hidalgo, F. & Olivia Maciuceanu, A. (2006) Essentially contested concepts: Debates and applications. *Journal of Political Ideologies*, 11(3), 211–246. doi:10.1080/13569310600923782.
- Connolly, W.E. (1993) *The Terms of Political Discourse*, 3rd Edition. : Princeton University Press.
- Convention on Biological Diversity (2022) *Kunming-Montreal Global Biodiversity Framework*. : United Nations Environment Program.
- Deutz, A., Heal, G.M., Niu, R., Swanson, E., Townshend, T., Zhu, L. et al. (2020) *Financing nature: Closing the global biodiversity financing gap*. URL <https://www.paulsoninstitute.org/conservation/financing-nature-report/>
- Diniz-Filho, J.A.F., Loyola, R.D., Raia, P., Mooers, A.O. & Bini, L.M. (2013) Darwinian shortfalls in biodiversity conservation. *Trends in Ecology & Evolution*, 28(12), 689–695. doi:10.1016/j.tree.2013.09.003.

- Douglas, H. (2009) *Science, Policy, and the Value-Free Ideal*. Pittsburgh: University of Pittsburgh Press.
- Díaz, S., Lavorel, S., de Bello, F., Quétier, F., Grigulis, K. & Robson, T.M. (2007) Incorporating plant functional diversity effects in ecosystem service assessments. *Proceedings of the National Academy of Sciences*, 104(52), 20684–20689. doi:10.1073/pnas.0704716104.
- Edmonds, D. (2013) *Would You Kill the Fat Man?: The Trolley Problem and What Your Answer Tells Us about Right and Wrong*. : Princeton University Press.
- Faith, D.P. (1992) Conservation evaluation and phylogenetic diversity. *Biological Conservation*, 61(1), 1–10. doi:10.1016/0006-3207(92)91201-3.
- Fisher, R.A., Corbet, A.S. & Williams, C.B. (1943) The relation between the number of species and the number of individuals in a random sample of an animal population. *The Journal of Animal Ecology*, 42–58. doi:10.2307/1411.
- Foot, P. (1967) The problem of abortion and the doctrine of the double effect. *Oxford Review*, 5, 5–15. doi:10.1093/0199252866.003.0002.
- Frank, D.M. (2022) Science and values in the biodiversity-ecosystem function debate. *Biology & Philosophy*, 37(2), 7. doi:10.1007/s10539-022-09835-4.
- Frankham, R. (2005) Genetics and extinction. *Biological Conservation*, 126(2), 131–140. doi:10.1016/j.biocon.2005.05.002.
- Gaggiotti, O.E., Chao, A., Peres-Neto, P., Chiu, C.H., Edwards, C., Fortin, M.J. et al. (2018) Diversity from genes to ecosystems: A unifying framework to study variation across biological metrics and scales. *Evolutionary Applications*, 11(7), 1176–1193. doi:10.1111/eva.12593.
- Gallie, W.B. (1955) Essentially contested concepts. *Proceedings of the Aristotelian Society*, 56, 167–198.
- Garcia, R.K. & Newman, J.A. (2016) Is it possible to care for ecosystems? policy paralysis and ecosystem management. *Ethics, Policy & Environment*, 19(2), 170–182. doi:10.1080/21550085.2016.1204054.
- Gellner, E. (1974) *Legitimation of Belief*. Cambridge: Cambridge University Press.
- Gotelli, N.J. & Colwell, R.K. (2001) Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 4(4), 379–391. doi:10.1046/j.1461-0248.2001.00230.x.
- Gray, J. (1977) On the contestability of social and political concepts. *Political Theory*, 5(3), 331–348. doi:10.1177/009059177700500304.
- Hill, M.O. (1973) Diversity and evenness: A unifying notation and its consequences. *Ecology*, 54(2), 427–432. doi:10.2307/1934352.
- Hortal, J., de Bello, F., Diniz-Filho, J.A.F., Lewinsohn, T.M., Lobo, J.M. & Ladle, R.J. (2015) Seven shortfalls that beset large-scale knowledge of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, 46, 523–549. doi:10.1146/annurev-ecolsys-112414-054400.
- Hurlbert, S.H. (1971) The nonconcept of species diversity: A critique and alternative parameters. *Ecology*, 52(4), 577–586. doi:10.2307/1934145.
- Jost, L. (2007) Partitioning diversity into independent alpha and beta components. *Ecology*, 88(10), 2427–2439. doi:10.1890/06-1736.1.
- Jurasinski, G., Retzer, V. & Beierkuhnlein, C. (2009) Inventory, differentiation, and proportional diversity: A consistent terminology for quantifying species diversity. *Oecologia*, 159(1), 15–26. doi:10.1007/s00442-008-1190-z, epub 2008 Oct 25.
- Lande, R. (1996) Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos*, 5–13. doi:10.2307/3545743.
- Legendre, P. & De Cáceres, M. (2013) Beta diversity as the variance of community data: Dissimilarity coefficients and partitioning. *Ecology Letters*, 16(8), 951–963. doi:10.1111/ele.12141.
- Mace, G.M., Norris, K. & Fitter, A.H. (2012) Biodiversity and ecosystem services: A multilayered relationship. *Trends in Ecology & Evolution*, 27(1), 19–26. doi:10.1016/j.tree.2011.08.006.
- MacIntyre, A. (1973) The essential contestability of some social concepts. *Ethics*, 84(1), 1–9. doi:10.1086/291897.
- Magurran, A.E. (2004) *Measuring Biological Diversity*. Oxford, UK: Blackwell Publishing.
- Newman, J.A. (2025) Communities and ecosystems. *EcoEvoRxiv*, doi:10.32942/X2BW5K.
- Nobles, E.C. (2025) The ethical foundations of biodiversity metrics. *Current Opinion in Environmental Sustainability*, 72, 101503. doi:10.1016/j.cosust.2024.101503.
- Noss, R.F. (1990) Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology*, 4(4), 355–364. doi:10.1111/j.1523-1739.1990.tb00309.x.
- Pavoine, S. & Ricotta, C. (2024) Functional and phylogenetic β diversities and their link with clustering/overdispersion and uniqueness/redundancy. *Ecosphere*, 15(11), e70045. doi:10.1002/ecs2.70045, handling Editor: Michael P. Perring.
- Pollock, L.J., Thuiller, W. & Jetz, W. (2017) Large conservation gains possible for global biodiversity facets. *Nature*, 546(7656), 141–144. doi:10.1038/nature22368.
- Pyron, R.A. & Mooers, A. (2022) The normative postulate problem: Hidden values in ecology, evolution, and conservation. *Biological Conservation*, 270, 109584. doi:10.1016/j.biocon.2022.109584.
- Ricklefs, R.E. (2008) Disintegration of the ecological community: American society of naturalists sewall wright award winner address. *The American Naturalist*, 172(6), 741–750. doi:10.1086/593002.
- Roswell, M., Dushoff, J. & Winfree, R. (2021) A conceptual guide to measuring species diversity. *Oikos*, 130(3), 321–338. doi:10.1111/oik.07202.
- Sarkar, S. (2005) *Biodiversity and Environmental Philosophy: An Introduction*. Cambridge, UK: Cambridge University Press.
- Scheiner, S.M. (2003) Six types of species-area curves. *Global Ecology and Biogeography*, 12(6), 441–447. doi:10.1046/j.1466-822X.2003.00061.x.
- Scheiner, S.M., Cox, S.B., Willig, M.R., Mittelbach, G.G., Osenberg, C. & Kaspari, M. (2000) Species richness, species-area curves and simpson's paradox. *Evolutionary Ecology Research*, 2, 791–802.
- Scheiner, S.M., Kosman, E., Presley, S.J. & Willig, M.R. (2025) The units of biodiversity. *Ecological Monographs*, 95(2), e70019. doi:10.1002/ecm.70019.
- Shannon, C.E. (1948) A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379–423. doi:10.1002/j.1538-7305.1948.tb01338.x.
- Simpson, E.H. (1949) Measurement of diversity. *Nature*, 163(4148), 688–688. doi:10.1038/163688a0.
- Tuomisto, H. (2010) A diversity of beta diversities: Straightening up a concept gone awry. part 1. defining beta diversity as a function of alpha and gamma diversity. *Ecography*, 33(1), 2–22. doi:10.1111/j.1600-0587.2009.05880.x.
- Tuomisto, H. (2010) A diversity of beta diversities: Straightening up a concept gone awry. part 2. quantifying beta diversity and related phenomena. *Ecography*, 33(1), 23–45. doi:10.1111/j.1600-0587.2009.06148.x.

- Turton-Hughes, S., Holmes, G. & Hassall, C. (2024) The diversity of ignorance and the ignorance of diversity: origins and implications of “shadow diversity” for conservation biology and extinction. *Cambridge Prisms: Extinction*, 2, e18. doi:10.1017/ext.2024.21.
- United Nations Conference on Environment and Development (1992) *Convention on biological diversity*. Adopted 5 June 1992. URL <https://www.cbd.int/convention/text/>
- Veech, J.A. & Crist, T.O. (2010) Diversity partitioning without statistical independence of alpha and beta. *Ecology*, 91(7), 1964–1969. doi:10.1890/08-1727.1, corresponding Editor: A. M. Ellison.
- Waldron, J. (1994) Vagueness in law and language: Some philosophical issues. *California Law Review*, 82(3), 509–540.
- Whittaker, R.H. (1965) Dominance and diversity in land plant communities: Numerical relations of species express the importance of competition in community function and evolution. *Science*, 147(3655), 250–260. doi:10.1126/science.147.3655.250.
- Whittaker, R.H. (1972) Evolution and measurement of species diversity. *Taxon*, 21(2-3), 213–251. doi:10.2307/1218190.
- Whittaker, R.J., Willis, K.J. & Field, R. (2001) Scale and species richness: Towards a general, hierarchical theory of species richness. *Journal of Biogeography*, 28(4), 453–470. doi:10.1046/j.1365-2699.2001.00563.x.
- World Health Organization (2025) *Biodiversity*. [Online; accessed 2025-07-28]. URL <https://www.who.int/news-room/fact-sheets/detail/biodiversity>

How to cite this article: Newman J.A., and Peres-Neto R.. From metrics to meaning: diversity as an essentially contested concept. *Ecological Monographs* in press.

APPENDIX

S1 A GUIDE TO REPORTING CONSEQUENTIAL COMMITMENTS IN DIVERSITY QUANTIFICATION

Biodiversity assessment involves a sequence of conceptually distinct choices: which forms of biological variation are relevant, which entities constitute the units of diversity, how those entities are weighted, how diversity is represented quantitatively, and what spatial and temporal boundaries define the object of assessment. These choices are not all normative in the same sense or to the same degree. Some concern substantive judgments about what biological differences should matter; others concern how those differences should be represented quantitatively. Their importance lies in whether reasonable alternatives could materially change interpretation, rankings, trends, or decisions.

This appendix provides a practical framework for identifying and reporting such consequential commitments. For each decision point, we identify common alternatives, explain the principal commitment involved, and provide illustrative reporting language. In general, the Methods section should identify and justify consequential analytical choices, while the Discussion should address their implications for interpreting results, including limitations arising from reasonable alternatives. The examples are illustrative rather than prescriptive. The objective is not to require philosophical justification for routine methodological decisions, but to make explicit those choices for which defensible alternatives embody materially different conceptions of diversity.

Form of biological variation

Possible choices. Taxonomic, functional, phylogenetic, genetic, interactional, or other dimensions of biological variation, considered separately or jointly.

What the choice commits the analysis to. The selected dimension specifies the form of biological difference represented by the diversity measure. Taxonomic, functional, phylogenetic, and genetic diversity need not covary and can yield different rankings of the same sites or communities. Selecting one dimension therefore defines the inferential target while leaving other forms of variation outside the focal representation unless they are analysed separately.

Example of transparent reporting. “We assess taxonomic diversity because the management objective concerns species representation rather than maintenance of functional or phylogenetic distinctiveness. Our conclusions therefore apply to variation in species identity and should not be interpreted as equivalent assessments of functional, phylogenetic, or genetic diversity.”

Units that contribute to diversity

Possible choices. Species, populations, genotypes, individuals, lineages, functional entities, interaction types, or other biological units.

What the choice commits the analysis to. The choice specifies the entities among which variation is individuated. Treating species as the relevant units excludes within-species differentiation from the diversity measure, whereas population-, genotype-, lineage-, or function-based approaches recognize finer or alternative biological distinctions. The unit of analysis therefore determines which differences are represented as diversity and which are treated as variation within units.

Example of transparent reporting. "We treat species as the relevant units because the conservation targets are specified at the species level. Population structure and within-species genetic variation therefore do not contribute independently to the diversity measure. A population- or genotype-based assessment could yield different conservation priorities."

Properties emphasized within a chosen form of variation

Possible choices. Richness, rarity, abundance structure, dominance, evenness, functional distinctiveness, evolutionary distinctiveness, compositional turnover, or other properties of the selected biological units.

What the choice commits the analysis to. Even after the relevant biological dimension and units are specified, diversity can be defined in terms of different properties of those units. Within taxonomic diversity, for example, richness treats species number as sufficient, whereas abundance-sensitive measures allow relative abundance, dominance, rarity, or evenness to influence measured diversity. The choice therefore specifies which differences among assemblages are substantively relevant to the comparison.

Example of transparent reporting. "We use species richness because the inferential target is variation in the number of species present rather than variation in abundance structure. Species therefore contribute equally to the measure, and differences in dominance or evenness do not affect comparisons unless they are associated with species gain or loss."

Weighting of entities

Possible choices. Equal weighting, weighting by relative abundance, progressively greater emphasis on common species, explicit rarity weighting, weighting by functional or evolutionary distinctiveness, or other weighting schemes.

What the choice commits the analysis to. Weighting determines the relative influence of different entities on measured diversity. Within the Hill-number framework, increasing the diversity order q progressively increases the influence of common species relative to rare species. Other frameworks may instead weight entities by rarity, functional distinctiveness, evolutionary history, or another specified property. This choice is distinct from estimating abundances, traits, or distances: estimation determines the empirical inputs, whereas weighting determines their influence on the resulting quantity.

Example of transparent reporting. "We use Hill diversity at $q = 2$ because our objective is to characterize the diversity of species contributing most strongly to an ecological process expected to be driven primarily by common species. This choice gives relatively little influence to rare species compared with lower diversity orders. We therefore interpret the resulting estimates as dominance-weighted diversity rather than as measures intended to represent rare species strongly."

Quantitative representation of diversity

Possible choices. Entropy or uncertainty, probabilities, raw indices, effective numbers or numbers equivalents, additive differences, proportional changes, ratios, or other quantitative representations.

What the choice commits the analysis to. The representational scale determines which quantitative relationships are directly interpretable. Shannon entropy expresses uncertainty on an additive scale, whereas its Hill-number transformation preserves the same ordering and abundance weighting while expressing diversity as an effective number of equally abundant species. Choosing between such representations therefore concerns the meaning assigned to differences, ratios, equivalence, and proportional change rather than the underlying ranking alone.

Example of transparent reporting. "We report Hill diversity at $q = 1$ because our comparison concerns proportional differences in effective diversity rather than additive differences in entropy. The transformation preserves the Shannon weighting of species abundances but permits direct interpretation of ratios in effective numbers of species. We therefore interpret ratios of ${}^1D$, not ratios of Shannon entropy."

Partitioning of diversity across scales

Possible choices. Alpha, beta, and gamma diversity; additive or multiplicative partitioning; differentiation or proportional diversity; local contribution or regional differentiation; and related approaches to representing diversity across spatial scales.

What the choice commits the analysis to. Partitioning determines how diversity at local, among-community, and regional scales is related quantitatively. Additive formulations represent β diversity as an absolute contribution beyond average local diversity, whereas multiplicative formulations represent it as proportional differentiation or scaling. This representational choice is analytically distinct from the weighting of rare and common species within the diversity measure.

Example of transparent reporting. "We use multiplicative partitioning because our objective is to quantify proportional differentiation among assemblages. Under this representation, β diversity is interpreted as the effective number of distinct communities required to scale mean local diversity to regional diversity. An additive partition would instead emphasize absolute contribution to regional diversity."

Definition of the ecological community

Possible choices. A co-occurring assemblage, an interacting community, a habitat-defined community, a taxonomically delimited assemblage, a network-defined community, or another delimitation of the ecological system.

What the choice commits the analysis to. Community definition specifies the biological and spatial domain over which diversity is evaluated. Alternative defensible delimitations can include different organisms, interactions, or spatial extents and therefore produce different estimates even when based on the same underlying ecological system. The relevant commitment concerns the boundaries of the object being measured rather than the statistical properties of the estimator.

Example of transparent reporting. "We define communities as vascular plant assemblages occurring within individual plots because the analysis concerns spatial variation in local composition. Community membership is therefore based on co-occurrence at the plot scale rather than on direct ecological interaction. Network-based or habitat-wide definitions would delimit a different ecological object."

Criteria for inclusion in the community

Possible choices. Observed species only, present but undetected species, inferred current membership, the regional species pool, dark diversity, or other potential members.

What the choice commits the analysis to. Inclusion criteria determine whether diversity is restricted to realized community membership or extends to ecologically plausible but currently unrealized membership. Imperfect detection concerns estimation of present diversity and can, in principle, be addressed empirically. Inclusion of locally absent but ecologically plausible species instead changes the scope of the quantity being represented.

Example of transparent reporting. "We define diversity in terms of species currently present, including species inferred to have been missed through imperfect detection. Species belonging to the regional pool but absent from a site are not included. Our estimates therefore describe realized community diversity rather than dark diversity or restoration potential."

Temporal scope and baseline

Possible choices. The present state, a historical baseline, a moving baseline, a future or potential state, a temporal snapshot, or temporally integrated diversity.

What the choice commits the analysis to. Temporal scope determines whether diversity is represented at a particular time, integrated through time, or evaluated relative to a reference condition. Where a baseline is used, its selection affects the interpretation of loss, recovery, degradation, and departure from a desired state. Empirical data can characterize alternative historical or contemporary states, but they do not by themselves determine which state should serve as the relevant benchmark.

Example of transparent reporting. "We use the contemporary state as the reference condition because the analysis compares current management alternatives rather than estimating departure from a historical condition. Our results therefore describe present differences among sites and should not be interpreted as estimates of historical biodiversity change. A historical baseline could yield a different assessment of degradation or recovery."

Integration of multiple forms of diversity

Possible choices. Separate indicators, multidimensional dashboards, composite indices, weighted combinations, optimization across dimensions, or other approaches to integrating multiple forms of biological variation.

What the choice commits the analysis to. Integration determines whether different dimensions of biodiversity remain analytically distinct or are rendered commensurable. Composite measures require explicit or implicit judgments about relative weighting and about whether gains in one dimension can compensate for losses in another. Retaining separate indicators avoids that compensability assumption but leaves trade-offs among dimensions unresolved.

Example of transparent reporting. "We report taxonomic, functional, and phylogenetic diversity separately because they represent non-equivalent forms of biological variation and we do not assume compensability among them. Combining them into a composite index would require explicit weighting of the dimensions and specification of the extent to which gains in one could offset losses in another."

Purpose for which diversity is represented

Possible choices. Description, hypothesis testing, monitoring, conservation prioritization, restoration, ecosystem functioning, resilience, policy targets, or other scientific or decision-making purposes.

What the choice commits the analysis to. Purpose provides the criterion against which a defensible representation can be judged appropriate. It does not eliminate the commitments embedded in a metric, but it can explain why particular biological properties, weighting schemes, scales, or quantitative representations are relevant to the inferential or decision context. A measure appropriate for one objective need not be appropriate for another.

Example of transparent reporting. "The management objective is to minimize local species loss, so species richness is the primary response variable. This choice gives equal influence to each species and does not incorporate abundance, functional role, or evolutionary distinctiveness. A management objective centred on ecosystem functioning or evolutionary representation would require a different diversity measure."

Comparison of reasonable alternative representations

Possible choices. A single prespecified representation, multiple metrics, diversity profiles, alternative dimensions of diversity, formal sensitivity analysis, or other approaches to evaluating dependence on contestable choices.

What the choice commits the analysis to. Comparison among reasonable alternatives reveals whether a substantive conclusion is robust to the conception of diversity adopted. Such analysis is most informative when alternative metrics, diversity orders, dimensions, or partitioning frameworks alter rankings, trends, or recommendations. The objective is not to report every available measure, but to identify conclusions that depend materially on a contestable representational choice.

Example of transparent reporting. "Community A ranks higher under richness and Shannon diversity, whereas Community B ranks higher under Simpson diversity. We therefore report the alternative rankings rather than treating either ordering as an unconditional property of the communities. The reversal indicates that the conclusion depends on the relative influence assigned to common versus rare species."

Using the guide

The preceding decision points are not intended as an exhaustive checklist. Their relevance depends on the inferential and decision context. Many methodological choices will be conventional or substantively inconsequential; others may alter which communities are judged more diverse, how biodiversity change is interpreted, or which sites receive conservation priority.

Three reporting principles follow. First, state the biological, inferential, or representational rationale for consequential choices rather than allowing the name of a metric or framework to serve as its own justification. Second, restrict interpretation to the dimensions and properties represented by the chosen measure. Third, where reasonable alternatives embody materially different commitments, consider whether they could alter substantive conclusions. Where feasible and informative, assess this sensitivity; otherwise, acknowledge its potential implications for interpretation.

In practice, these commitments need not be reported as a lengthy checklist. A concise Methods statement can often make the principal choices, their rationale, and any relevant sensitivity analysis explicit, while the Discussion can address the implications and limitations of those choices. For example:

"We quantify taxonomic diversity at the species level using Hill numbers with $q = 1$, because our objective is to compare abundance-weighted assemblage diversity while retaining sensitivity across a broad range of species abundances. We report diversity as effective numbers of species to permit proportional comparison among sites and use multiplicative partitioning to quantify differentiation between local and regional scales. Communities are delimited by plot-level co-occurrence, and only species inferred to be currently present are included. We assessed sensitivity to alternative species weighting by repeating comparisons at $q = 0$ and $q = 2$; site rankings were unchanged."

The objective is not methodological pluralism for its own sake, nor exhaustive justification of every analytical decision, but transparency about the consequential commitments on which substantive conclusions depend.