Chasing Biodiversity Off the Scientific and Conservation Tracks

Donald S. Maier*

The idea of conserving biodiversity has become central to the very meaning of biological conservation—in the public imagination and for conservation organizations worldwide. Identification of conservation with *biodiversity* conservation owes extensively to the idea that warrant for biodiversity's conservation is anchored in the empirical thesis that biodiversity causally determines ecosystem functioning, and thereby, somehow, important ecosystem services. This idea has fueled an enormous research program dedicated to producing the requisite causal evidence. This essay first reviews the data that are supposed to constitute direct evidence for biodiversity's causal influence. It proceeds by answering a heretofore unasked, yet foundational, question for causal hypotheses: Do these data meet basic requirements for credibility as causal evidence? By virtue of mistakenly reading causal significance into (i) massive numbers of causally *ir*relevant data points, (ii) an equation that simply equates a stipulated definition to an algebraically equivalent expression, and (iii) correlations produced by arbitrary computations over previously collected data sets, these data fall well short of meeting these requirements. Mistakes also suffuse the notion that biodiversity's supposed causal influence gives reasons to conserve it. These mistakes are exposed when the conservation argument that was supposed to proceed from the premise of biodiversity's causal influence is clearly spelled out. Once made explicit, each step is seen to rely on a questionable assumption, invalid logic, or both. The essay concludes with implications for conservation, biodiversity research, and scientific inquiry more generally.

1 Introduction

In the 1990s, two ideas—one normative, the other empirical—captured the public imagination. The normative idea is that, in light of evidently increasing rates of extinction, "saving biodiversity" should be conservation's premier goal. The empirical idea is a causal thesis—namely, that

*Independent Scholar, 8035 Yule Tree Ln., Sebastopol, CA 95472 USA, dmaier@dslextreme.com
https://orcid.org/0000-0002-1025-347X

biodiversity is *causally* responsible for the functional attributes of ecosystems. Ecologists yoked these two ideas together. Their thought (which §2.2 sets out in their own words) was that, by "demonstrating" biodiversity's causal role in determining ecosystem functions and by tying these functions to services that benefit humankind, they could provide a scientifically unassailable basis for a mandate to conserve biodiversity.

That, in a nutshell, is what launched a program of ecological research dedicated to the mission of making biodiversity's causal salience irrefragable. This program became the ecological subdiscipline of Biodiversity Ecosystem Function or 'BEF' research. From a handful of modest studies in the 1990s, it ballooned into what is now one of the biggest, most lavishly funded, and institutionally entrenched research programs of the late twentieth and early twenty-first centuries. Its size and impact on ecological science can be read off its enormous publication footprint. BEF research now consumes the scientific energy of many hundreds of scientists who, at a dizzying pace, add to many thousands of existing publications, which all revolve around the supposition that biodiversity is a significant or even dominant causal force in ecosystem functioning.[1]

Enormous, too, is the commitment of resources to conservation organizations that, to a significant degree, operate under the influence of the idea of preserving biodiversity in order to preserve ecosystem services. The Nature Conservancy (2021 revenues of $1.8 billion), the World Wildlife Fund (2021 revenues of $451 million), and Conservation International (2021 revenues of $218 million) collectively spend the lion's share of the world's conservation dollars on programs at least partly underwritten by this idea.[2] The mission of preserving biodiversity *specifically* in order to preserve ecosystem services is also conspicuously sponsored by the United Nations: The United Nations Environmental Programme (UNEP) (2021 revenues of $1.02 billion) provides secretariat services to the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES) (2018–2021 revenues, including from thirty-one of its 140 member nations, and in-kind contributions of $22 million). That latter organization's very name all but announces its dedication to promoting ecosystem services *by means of* conserving biodiversity. By its own accounting (https://ipbes.net/about), the IPBES alone taps the energy of thousands of scientists across the globe who continually labor to expand an already monumental library of documents (numbering nearly 2,000 listed at https://ipbes.net/library as of this writing), including 247 "Assessments on Biodiversity and Ecosystem Services" (https://catalog.ipbes.net/assessments/), which promote, elaborate, and declaim policy based on this mission.

---

[1]A rough indication of how pervasively scientists tie biodiversity to ecosystem functions may be read off the 1.9 million hits for an uncurated search for "Biodiversity Ecosystem Functions" in Google Scholar. Research at the Cedar Creek Reserve Field station *alone* has generated multiple hundreds of papers on this topic largely funded by many millions of dollars of NSF funds (https://www.cedarcreek.umn.edu/research/publications) through its Long Term Ecological Research Network (LTER), which supposes (in its 2019 *Decadal Review Self Study* https://lternet.edu/documents/decadal-review-report-2022/) that "The relationship between biodiversity and ecosystem function ranks among the most important research issues in modern ecology."

[2]The Nature Conservancy and the World Wildlife Fund are "Core Partners" of The Natural Capital Project (https://naturalcapitalproject.stanford.edu/who-we-are/partners), which is dedicated to the idea that "The world's ecosystems can be seen as capital assets; if well-managed, their lands, waters, and biodiversity yield a flow of vital life-support services" (https://naturalcapitalproject.stanford.edu/who-we-are/natural-capital-project). Conservation International also subscribes to the so-called Natural Capital Protocol, which seeks to promote nature, and most particularly, its biodiversity, in terms of "natural assets and the ecosystem services they provide" (https://capitalscoalition.org/capitals-approach/natural-capital-protocol/). Maier (2018) offers more documentation of the nexus involving BEF science, the Natural Capital Project, The Nature Conservancy, and other large conservation organizations.

In short, overwhelming and overwhelmingly entrenched acceptance of the empirical thesis that biodiversity is a significant causal determinant of ecosystem functions aligns with similarly overwhelming and overwhelmingly entrenched acceptance of the normative conservation thesis that, via some supposed connection of ecosystem functions to ecosystem services, this causal relationship mandates biodiversity's conservation. Yet no systematic review of either thesis has been undertaken *by anyone who is not already directly or indirectly committed to them*.[3] This essay fills this enormously consequential lacuna.

Hundreds of thousands of pages have been written in support of the causal thesis and thousands more have rationalized biodiversity's conservation by reference to it.[4] Even back in 2011, Cardinale et al. observed, "there has been an exponential increase in the number of diversity experiments, and the number of publications has more than tripled since 2006 when earlier databases were put together" (2011, 573). The pace of publication has only increased since then. Nevertheless, concise examination of the causal thesis is possible because essentially all experimental data held to be evidence for it derives from experiments that (a) attempt to test the causal influence of species richness by manipulating levels of that variable while observing consequent levels of some ecosystem function, and (b) share interpretive and analytic assumptions (perhaps the only ones possible) to parse these data as evidence of species richness' causal influence on the ecosystem function.[5] We can therefore set aside other details to ask the heretofore-unasked question: Do the data produced by experiments characterized by these essential methodological and analytic elements constitute credible causal evidence for biodiversity's causal influence?

Section 2 lays a conceptual and methodological foundation for answering this question by pinpointing minimal requirements for causal evidence, which are dictated by the distinguishing characteristics of causal relationships. Section 3 then asks whether data from experiments that manipulate species richness meet these evidential requirements. It finds that these data fall short of credibility as causal evidence because experiments count (and most likely cannot avoid counting) every data point at every level of richness as evidence of the impact of richness at that level, yet some enormous number have no causal relevance at all. Section 4 then examines the particular, standardly utilized analysis of experimental data, which "additively partitions" species richness' supposed causal influence by resolving it into summed terms. It shows that the presumed causal significance of this sum-of-terms representation is wholly undermined by the fact that it merely restates a stipulative definition in algebraically equivalent terms.

Section 5 follows BEF science's tack from species richness to functional diversity as the biodiversity factor that is supposed to impact ecosystem functions. It again asks whether data that are held to be evidence for this twin thesis meet basic requirements for causal significance. In this case, a negative answer follows from its particular use of HARKing (constructing Hypotheses After the Results Are Known) in which (a) data for some ecosystem function are culled from previously conducted *species richness* experiments and (b) correlation of functional diversity

---

[3]The italicized phrase excludes numerous self-evaluative writings by the very scientists who are responsible for the science, conservation organizations and conservationists who promote their conservation mission on its basis, and science writers who uncritically echo back what these seemingly reliable sources tell them.

[4]It must be emphasized that this essay is all and only about the enormous body of science that purports to demonstrate biodiversity's causal impact on ecosystems and all and only about the ubiquitous conservation rationale yoked to that thesis. It is *not* about and it should not be thought to be about other work on biodiversity that does not attempt to tie or portray it as causally responsible for ecosystem functions.

[5]Meta-analyses and surveys of this body of experimental work have proliferated, but just uncritically *presume* that the experimental data from the meta-analyzed and surveyed work is all causally credible. They merely re-analyze these data utilizing the same interpretive framework as the original experiments. This uncritical stance is evident in the meta-analyses and surveys cited later in other contexts, including Cardinale et al. (2009), Cardinale et al. (2011), and Tilman et al. (2014). Section 5 discusses how these same data, again without critical review, are again re-analyzed in supposing them to be evidence for the causal impact of functional diversity.

with that ecosystem function is achieved by choosing—from among an infinitude of differing and often contradicting, yet biologically qualified formulations of functional diversity—some formulation that *makes* it correlate.

Section 6 then answers another heretofore-unasked question: Setting aside whether or not biodiversity is, in fact, a cause of ecosystem functions, what argument, exactly, runs from that premise to the proposition that biodiversity is a cause of *ecosystem services*, and on to the conclusion that biodiversity ought to be conserved? By fully spelling out this previously unarticulated argument, each step can be seen to rely on a questionable assumption, invalid logic, or both.

Section 7 draws some conclusions about what these findings—about a conservation commitment and a scientific commitment to support it—entail for ecological science, for conservation, and for the conduct of science more generally.

2 Causal relationships and causal evidence in biology and BEF science

2.1 What distinguishes causal relationships

Causal relationships are the gold standard of scientific and even everyday explanation, arguably because they tell us how some outcome can be brought about when some (other) things—its causes—are changed. Causal relationships are in this sense *difference-making* (Woodward 2003). Woodward's account of causality as, saliently, difference-making, is but one of many. However, I believe it to be more accommodating of an hypothesized causal relationship between biodiversity and ecosystem functions than others: If that relationship is thought to be causal, then I believe that Woodward's account gives it pretty much its best shot to qualify.[6]

The difference-making quality that is central to Woodward's account of causal relationships is absent from other sorts of relationships, such as predictive, proxy, classificatory, descriptive, mereological, and logical ones. In the well-worn example: The fact that movement of the pressure indicator on a barometer is reliably *change-relating* with respect to weather conditions makes this movement predictive of oncoming storms. But one cannot cause a storm to approach by moving that indicator. In the biological domain, the fact that opossums are good *proxies* for overall mammalian species richness in the Amazon (Sebastião and Grelle 2009) does not entail that the number of other mammal species resident in the Amazon can be changed by introducing or removing Amazonian Didelphimorphs.

The salient point for BEF research is that, if biodiversity is *not* a cause of ecosystem functioning, but rather, merely (say) a predictor or a proxy for it, then one should *not* expect, for example, that reducing it will reduce levels of ecosystem functions.

2.2 BEF research as research about biodiversity's causal influence

One might still ask: Do BEF scientists actually suppose that their research concerns biodiversity as a causal factor in determining ecosystem functions?[7] At least three factors make it clear that

---

[6]This essay cannot present a comprehensive survey of accounts of causality to properly support this claim. I must therefore leave it as a challenge to the skeptical reader to find some account of causality that would be more accommodating than Woodward's of the thesis that levels of biodiversity causally determine levels of ecosystem functions. I myself began my work on this topic by taking up this challenge, only to find that most accounts of causality would summarily reject this thesis as causal, while no other account would offer any advantage over Woodward's.

[7]Addressing the question of whether BEF research is concerned to establish biodiversity as a cause of ecosystem functioning is made essential by my experience of responses to challenges to the credibility of its evidence as causal evidence, which simply deny that BEF science's central thesis is a causal one. According to these responses, it is merely a predictive thesis or a "general" one, whatever that might mean.

they do. The first is that it would otherwise be incoherent for them to routinely decry loss of biodiversity in the very first sentence of papers and press for its conservation on the grounds that its loss entails loss of ecosystem services. Representative is a paper by three of BEF science's most central figures, Tilman, Isbell, and Cowles (2012). Their very first sentence makes it clear that their principal concern is with the loss of biodiversity and their very last makes clear their concern to conserve (or restore) it for the ecosystem services they suppose it to provide:

> Although the impacts of the loss of biodiversity on ecosystem functioning are well established, the importance of the loss of biodiversity relative to other human-caused drivers of environmental change remains uncertain. (Tilman et al. 2012, 10394)

> … contemporary biodiversity declines are among the dominant drivers of changes in ecosystem functioning, and that restoration of biodiversity in managed and seminatural ecosystems may be an efficient way to restore desired ecosystem services. (Tilman et al. 2012, 10397)

BEF scientists' concern to find evidence for biodiversity as a causal factor is, secondly, also evident from the BEF research enterprise's heavy investment—in money, time, scientific personnel, and careers—in experimentation that seeks to manipulate species richness as an independent variable and observe consequent changes on ecosystem functions. Cardinale et al. (2009, 854):

> summarize the results of 164 experiments (reported in 84 publications) that have manipulated the richness of primary producers, herbivores, detritivores, or predators in a variety of terrestrial and aquatic ecosystems …

Cardinale et al. (2011, 574) then claimed (without documentation) to have found 410 more experiments that attempted to manipulate species richness:

> The updated data set included 574 independent manipulations of species richness published in 192 peer-reviewed papers reporting 1417 diversity effect sizes.

These experiments formed the basis of their opening sentence:

> Over the past several decades, a rapidly expanding field of research known as biodiversity and ecosystem functioning has begun to quantify how the world's biological diversity can, as an independent variable, control ecological processes that are both essential for, and fundamental to, the functioning of ecosystems. (Cardinale 2011, 572)

Manipulating an ecological variable is a time- and labor-intensive task; ecologists undertake it only when they believe it holds promise for producing evidence that the manipulated variable causes changes in some other variable.

Third, the active, agential language that BEF scientists routinely employ to characterize biodiversity's role in ecosystems can only be sensibly interpreted as causal. Tilman et al. (2012, 10394, 10397) claim that, without qualification, biodiversity is *a dominant driver* of productivity.[8] Tilman et al. (2014, 471, 172) claim that, without qualification, biodiversity is *a major determinant* [my italics here and below] of productivity. Cardinale et al. (2007, 18123, 18125, 18126, 18127, 18128), Isbell et al. (2013, 11911, 11915), and Tilman et al. (2014, 471–475,

---

[8]Here and below in this and the next paragraph, the phrases are the scientists' but the italics are mine.

481, 484–488) all claim, again with no qualification, that biodiversity *impacts* biomass production. Isbell et al. (2018, 763), like many other BEF scientists, build biodiversity's supposed agency into their paper's title: "Quantifying the *effects of biodiversity on* ecosystem functioning." Also in their title, Cardinale et al. (2013, 1697) assert, "Biodiversity simultaneously *enhances* the production and stability of community biomass, but the effects are independent." And Cardinale et al. (2012, 60) state, "In the sections that follow, we summarize how biological variation per se *acts as an independent variable to affect* the functions and services of ecosystems."

These statements are but drops in an ocean of similar ones that continually and unambiguously declare biodiversity's causal salience. Indeed, some BEF researchers more straightforwardly employ the word "cause." For example, Roscher et al. (2005, 419) declare that they are interested in "clear causal relationships between species richness and ecosystem productivity …." And Selmants et al. (2012, 723) suppose that "Over the past 20 years, many manipulative experimental studies have demonstrated *causal linkages* between species richness and a wide variety of ecosystem processes including productivity, resource use and resistance to invasion."

2.3 Mechanisms and non-mechanistic causal relationships

BEF scientists understandably express concern about identifying *mechanisms* through which biodiversity (they suppose) exerts its causal influence. And they commonly characterize additive partitioning, the basis of BEF causal analysis discussed in Section 4, as representing mechanisms for biodiversity's causal influence. Multiple confusions regarding mechanisms in the BEF literature therefore make it important to understand the sense of "mechanism" in which the quest to identify them is legitimate, and to avoid conflating that quest with one for a *mechanistic* model. Additionally, it is important to understand that the hypothesis that biodiversity is a causal factor in ecosystems should not be summarily dismissed merely on the grounds that there is no settled theory of the mechanisms that might underlie its operation.

Insofar as mechanisms are understood to be intermediary or mediating causes, the concern of BEF scientists to find them for biodiversity's supposed causal action aligns with broad scientific concern to find finer-grained causal explanations. For example, a drug for which there already is compelling evidence for its efficacy in causing recovery is often the subject of pharmacological research that seeks mechanisms—the drug's biochemical and physical-chemical propensities for interacting with particular target molecules at particular sites—through which it produces its pharmacological effects. At the same time, the fact that the mechanism of action for many efficacious drugs is *un*known does not (or should not) call into question the credibility of causal evidence for their efficacy. In a similar way, the evidence for biodiversity's causal action on ecosystem functions can stand or fall without reference to any mechanism through which it may be thought to operate.

On the other hand, if there did turn out to be credible evidence that biodiversity is a cause of ecosystem functioning, no one should expect the mechanisms of this causal relationship to fit a narrowly *mechanistic* model. This model is characterized by operation of a stable set of stable things arranged in some stable spatio-temporal organization, which have stable, modular causal roles (Dupré 2013). An archetypical example is the rotation of a bicycle's pedals, which causes it to move forward via the fixed, pre-determined intermeshing of chainring, chain, rear cog, and rear wheel.

Narrowly mechanistic models do have widespread application in biology. One example is Davidson et al.'s (2002) causal explanation of how undifferentiated embryonic cells develop into specialized tissue by reference to the fixed spatial (and temporal) organization of a fixed inventory of cis-regulatory elements and transcribing genes. However, many apparently causal

explanations in biology are not narrowly mechanistic. Among them are an entire genre of network-based explanations in which topological properties are supposed to play a causal role in determining other properties of the network (Huneman 2010). These sorts of explanations are prominent in explaining the behavior of gene regulatory networks, biochemical networks that feature bioprotein-protein interactions and metabolic pathways, and, in the realm of ecology, trophic networks (Dunne 2006; Montoya et al. 2006). For example, it has been proposed that a trophic network's property of being scale-free may help to causally explain why species that are very highly connected to others in the network are relatively unaffected by random exodus of other species. It has also been proposed that a trophic network's property of being small-world causes a network to be more vulnerable to destabilization. If biodiversity were, in fact, causally determining of ecosystem functions, it would simply be one among these other causes that are not mechanistically elaborated.

Of course, topological properties of networks have internal structure that is wholly absent from an ecosystem's property of being more or less biodiverse. But this difference does not obviously disqualify biodiversity as a candidate causal factor in (non-mechanistically) determining other ecosystem properties.

2.4 Requirements for causal evidence

Evaluating the credibility of evidence for biodiversity's supposed causal influence on ecosystem functions requires first asking a foundational question: What sorts of data constitute credible evidence for its operation? Very crudely speaking, evidence may derive from two different sorts of observational procedures. Evidence might be drawn from observations of unmanipulated systems, commonly called *observational data*. Alternatively, evidence might be produced experimentally, by manipulating the candidate causal variable, setting it to various values, and observing any consequent changes in the value of other variables that characterize the system. These are commonly called *experimental data*.

Observational data may be compiled into either *comparative datasets* or *change-relating datasets*. Certain pervasive problems tend to limit the merits of both these kinds of datasets as causal evidence for any hypothesized causal relationship, including biodiversity's to ecosystem functioning.[9] A comparative dataset records, and consequently allows comparison of, observed levels of species richness and concomitant levels of productivity *across* some number of distinct ecosystems. With regard to such a dataset, species richness is "varied," not in an active sense relating to its manipulation, but rather only in the figurative sense that different ecosystems, with characteristics that differ in any number of respects, also differ in the number of resident species. This severely limits what can be inferred from them about the causal influence of species richness on productivity, which is a matter of the difference in productivity that would be made in an ecosystem by increasing or decreasing *that* ecosystem's number of species.

Even an imaginary world unlike ours, where the productivity of ecosystems always and without exception varied directly with the number of species, would leave open the question whether that correlation evidenced species richness acting as cause on productivity, productivity acting as cause on species richness, or some other factor acting as cause on both. The deficiencies in this evidence would be similar to deficiencies in the evidence for the causal influence of shadows consisting in observations that longer shadows invariably correlate with the greater height of adjacent flagpoles.

---

[9]Limitations of observational data derive both from human factors (Gueyffier and Cuchurat 2019) and theoretical obstacles in which the Causal Markov Condition plays a central role. Additionally, statistical techniques that attempt to surmount some of these obstacles in biomedical and epidemiological research (Khuller and Jena 2021) may not even be available to BEF research. A full discussion of this rich topic is beyond this essay's scope.

In our actual world, jointly observed levels of species richness and productivity display nothing approximating an exceptionless pattern of covariance. For example, not uncommonly, ecosystems with low levels of species richness are highly productive. That can be true in aquatic ecosystems with respect to phytoplankton and algal species richness, salt marshes, freshwater marshes, and forests of bamboo, redwood, Douglas fir, and eucalyptus, as well as tropical and temperate zone riparian forests (Huston and McBride 2002). More generally, co-variation of species richness with productivity appears to defy any simple correlational pattern. It can even vary with trophic level within a single ecosystem. However, even a well-established, universally observed pattern of co-variation would, on its own, constitute fragile evidence for how species richness *causally* connects with ecosystem functions.[10]

Scientists therefore look to observational data sets that are change-relating. Unlike comparative datasets, these track the co-present levels of variables over time—or alternatively, over some limited space—*within* each of some number of monitored ecosystems. Because they relate changes over time, their causal relevance is not undermined by the existence of speciose ecosystems with low levels of productivity or of species-depauperate ecosystems with high levels of the productivity. Yet they still lack the capacity to robustly evidence a *difference-making* relationship in which a change in productivity is *consequent* on a change imposed on species richness.

Multiple factors underlie this evidential limitation. One is that correlations between species richness and productivity may be due to productivity's determination of species richness. Even a cyclic relationship would undermine the claim that observational data compiled into change-relating datasets would be evidentially equivalent to experimental manipulation of species richness. More obvious problems arise from the confounding effects of uncontrolled and possibly highly correlated factors—among them: availability of nutrients and water; temperature mean, range, and fluctuation; frequency and nature of disturbances; and the presence of both competitive and facilitative organisms.

Little is gained by presuming that changes in some ecosystem variable *over time* can be validly represented by its value gradient *over space* within the ecosystem. This does not solve problems that arise from the possibility of cyclic relationships. Nor does it generally solve the problem of confounding candidate causes of productivity because those factors commonly covary with spatial productivity gradients.

All told, while observational data can provide valuable confirming evidence for causal relationships, their limitations as the sole basis for inferring causality generally place the burden of causal evidence on experiments that manipulate the candidate causal variable.[11] That fact is acknowledged explicitly by some BEF scientists (Roscher 2005), and implicitly by the many hundreds of scientists who have undertaken multiple hundreds of experiments that attempt to manipulate species richness in order to produce evidence for its causal determination of productivity and other ecosystem functions (§2.2). However, the credibility of evidence that these experiments afford depends on their meeting some basic requirements that apply to all causal relationships—whether mechanistic or not, and including the one hypothesized for species richness' determination of productivity.

When species richness, represented below by the variable name 'R', is the candidate causal variable, and R is posited to be a contributing cause of productivity, represented as 'P', then

---

[10]The ecological questions of what patterns of co-variation predominate—for example, whether they're unimodal or bimodal—and what latitudinal gradients imply are extensively debated in ecology. However, as the main text suggests, compelling answers to these questions would have but marginal force as evidence that biodiversity causally determines ecosystem functions.

[11]This discussion of observational evidence cannot dive into the big and nuanced topic of how, exactly, the limitations of observational evidence arise and what, exactly, they are. Its role here is only to highlight the preeminence of experimental evidence for causal theses.

the difference-making quality of causes suggests[12] that manipulation of R can provide credible evidence that R causally influences P when an experiment:

(1) employs some means of *intervening* on R or (as Section 3 explains) of *simulating* interventions on R that

(2) *change* the value of R by some particular amount Δr, while

(3) *fixing* the values of other variables in the system—variables other than R that might also causally influence P yet are *not* on the path via which R is supposed to change P—such that[13]

(4) the collection ΔR of particular changes Δr imposed on R are *stably associated* with concomitant or consequent changes ΔP in P.

3   Experimental evidence for species richness as cause

Conditions (1)—(4) describe what should be expected of a species richness-manipulating experiment to produce credible evidence that species richness causally affects productivity. I shall focus on productivity in assessing how well these conditions are met both because productivity has been the most prominent effect variable in BEF experimentation and because the basic requirements for causal credibility appear to be indifferent to which ecosystem function—whether productivity, zoonotic disease suppression, or stability—species richness is supposed to influence.

I shall also focus on what is known as the *random draw design* for manipulating species richness as a candidate causal variable. There are again two reasons: This design is the benchmark and reference for producing essentially all the data that the BEF research community holds up as direct evidence for BEF science's central claim—that species *richness* causally determines an ecosystem's productivity independent of *which* species happen to reside in it.[14] As well, if you

---

[12]This formulation hews closely to Woodward's (2003) classic one, while specializing it for species richness as cause and productivity as effect.

[13]Condition (3) looms large in subsequent discussion at the point marked by Note 20.

[14]Cardinale et al. (2011, 576) confirm the preponderant use of random draw in BEF research:

> the design of most biodiversity experiments was altered so that researchers held the total initial seeding density or biomass of producers constant across several levels of richness, and then grew (1) all possible combinations of species at each level of richness or (2) species combinations selected at random from all possibilities. Thus, the most common hypothesis tested by these experiments was that, when averaged across all species and species combinations, the efficiency of resource use and producer biomass increases as a function of the initial number of species seeded or grown in an experimental unit.

Option (1) is rarely taken because, as Naeem et al. (1996) explain, the combinatoric explosion of cases for each of all levels of species richness with the size of the species pool makes it impractical: That said, the presence of all combinations in an experiment would not, in any essential way, affect the credibility of the evidence.

The random draw design, with the essential elements described in the main text, is nearly ubiquitous in BEF experiments, including those that Cardinale et al. (2011) list in their 574-experiment survey. Naeem et al. (1996) is an early exemplar. Another is the work by Reich et al. (year not documented) on experiment E141 in the vast complex of Cedar Creek experiments at the University of Minnesota. It is noteworthy that this work is not published in a journal and therefore unreviewed and largely hidden from public view. However, the website (https://www.cedarcreek.umn.edu/research/experiments/e141) makes clear its use of random draw: "The BioCON experiment (E141) directly manipulates plant species numbers (1, 4, 9, or 16 perennial grassland species randomly chosen from a pool of 16 species, planted as seed in 1997) ...." Another exemplar, also part of Cardinale et

wish to manipulate species richness in a way that isolates that variable from covarying variables that represent the presence of particular species or particular collections of species, then randomizing the selection of species is not merely the obvious design choice, it may be the only available one.[15]

As condition (1) suggests, BEF experiments utilizing random draw do not actually intervene on R. Instead, they *simulate* interventions on R. They accomplish this by comparing the productivity observed in R-differing treatments, the specific composition of which are determined by random draw (without replacement) of species from an antecedently determined pool of, typically, 8 to 32 species—a number represented below by the variable name 'N'. With respect to *any* two treatments that differ in species richness R, the standard analysis interprets one treatment as the state of an ecosystem's species richness before intervention and the other as the state of *that* ecosystem's species richness afterwards. Analyses of BEF experiments treat these *R-differing* pairs of treatments as simulating interventions that are *R-changing*.

Simulating interventions on a candidate causal variable in this way is a perfectly legitimate means of probing hypothesized causal connections. Drug trials are a convenient reference. These experiments, too, often must contend with factors that may confound the action of the relevant Boolean causal variable D (0 = drug not taken; 1 = drug taken). Drug trials typically resort to simulating D-changing interventions because no single patient is both given the drug and also not given it.

However, in a BEF species-richness random-draw "trial," a crucial question—nowhere considered in the BEF scientific literature—arises: In simulating R-changing interventions, what variable, exactly, is being manipulated? According to a *number-changing interpretation*, a pair of treatments qualifies as simulating an R-changing intervention strictly on the basis of a difference in the number of species assigned to them—that is, *without qualification or restriction* with regard to which particular species with which particular characteristics enter into either the before-intervention count *or* the after-intervention count. Although another interpretation is discussed later in this Section, this one may initially appear to be the only one possible. And it is the one and only one that BEF researchers, without applying that label, employ when they interpret differences in productivity between *any* two R-differing treatments as evidence that their difference in species richness is causally responsible—*without qualification or restriction* with respect to which particular species are in these treatments.

---

al.'s 574-experiment survey, is Spehn et al. (2005) who resuscitated data from the European Union's BIODEPTH experiments six years after BIODEPTH ended. Spehn et al. (2005, 41) state, "Species mixtures were assembled by random draws from the local pool of typical co-occurring species …." They assayed biomass production (and other ecosystem functions) relative to species richness at eight, widely dispersed European sites. Of additional interest is that they were among the first to recycle data from *species richness* experiments to make claims regarding *functional diversity*, which they "measured" by counting the number of functional groups from among a total of three that they retrospectively identified for the species in the eight BIODEPTH species pools (see Section 5 of this essay). In experiments where the draw of species is constrained by functional group, such as the Jena Experiment reported on by Weisser et al. (2017), species richness is no longer an independent variable. Although their manipulation of species richness was constrained, Weisser et al. neglected to mention this significant condition in claiming that species richness is causally influential.

[15]The random draw protocol emerged after widespread recognition of fatal flaws in the design of Shahid Naeem's early "Ecotron" experiments (Naeem et al. 1994; Naeem et al. 1995). Those experiments assigned species from a pool of size N to mini-ecosystems in such a way that: (i) for any richness level $r_1 > r_0$, the species represented in r_1 included all species in r_0 and (ii) this hierarchy of species was defined so that the most productive species were represented only in treatments with the highest levels of species richness close to N. This design all but ensured that species richness was strongly correlated with productivity whether or not these variables are causally related. The only other alternative to the random draw design to have emerged is the little discussed random *partitions* design (Bell et al. 2005; Bell et al. 2009). That design has flaws closely related to those of the random draw design but space precludes its discussion in this essay.

The number-changing interpretive stance has a striking implication. It entails that, in statistically analyzing the productivity data to assess the causal impact of a change in species richness, all treatments that have the same number of species are grouped together, regardless of which species they contain. Yet this assessment cannot avoid attributing causal significance to differences in productivity between R-differing pairs of treatments that cannot plausibly be viewed as simulating how a change in the species richness of some particular ecosystem affects the productivity of *that* ecosystem.

Consider a study utilizing the species pool, represented in set notation as $\{A, B, C, D\}$, where each letter stands for a particular species in the pool. Suppose that random draw produces the treatments $\{A\}$ and $\{B, C, D\}$. The number-changing interpretive stance views this pair as causally relevant by virtue of simulating an R-changing intervention on the first, single-species treatment $\{A\}$ by way of increasing its richness—by two—to arrive at the second, three-species entourage $\{B, C, D\}$. If the productivity of this latter, three-species treatment is observed to be greater than that of the first, one-species treatment, this is supposed to be evidence that increasing a particular ecosystem's species richness (by two) *causes* that ecosystem's productivity to increase.

In this example, the change in R consists in total replacement of species resident in the original ecosystem with an entirely different entourage. It is difficult to make sense of differences in these two ecosystems' productivity as evidence that intervening to increase R in an ecosystem causes P to increase *in that ecosystem*. A more reasonable interpretation is that one ecosystem was obliterated, and the more speciose one that replaced it differed in productivity.

While any observed productivity difference in this case clearly has no credibility as evidence that changing the richness in some ecosystem causes *that* ecosystem's productivity to change, one might initially dismiss it as a bit of data pollution so inconsequential that BEF analyses of can safely ignore it. But that is not so. A tedious but straightforward combinatoric analysis shows that for a species pool of size N there are[16]

$$\sum_{n=1}^{<N/2} \left[\binom{N}{n} * \sum_{k=n+1}^{N-n} \binom{N-n}{k} \right] \tag{1}$$

entirely disjoint treatment pairs—that is, treatment pairs having no species in common. This number is 16 for $N = 4$; 2,472 for $N = 8$; 18,859,512 for $N = 16$; and 846,948,553,422,744 for $N = 32$, respectively. 847 trillion data points cannot legitimately be dismissed as "safely ignored noise." Rather, these data points constitute a substantial portion[17] of every dataset in

---

[16]Eq. 1 utilizes the commonly used notation $\binom{N}{k}$ to denote the number of possible combinations of N things taken k at a time. The first factor in the outer summation (denoted by sigma, 'Σ') represents one possible collection of species that consist of fewer than half the species from a pool of N. For each such collection, the second factor is the sum of all possible collections of species from the same pool that do not include any species in the first collection.

[17]The discussion in the main text explains the problem of data pollution by reference data points from treatments that are entirely disjoint. However, data pollution is far more extensive than that. For example, it cannot be credibly claimed that one ecosystem with six species (from a pool of sixteen) and a second ecosystem with twelve species that includes just one of those six (but not any of the five others) simulate one and the same ecosystem that simply has undergone a species-doubling intervention. More generally, if two treatments must, let us conservatively say, share more than 25% of their species to simulate the same ecosystem before and after a species number-changing intervention, then for $N = 4$, all the productivity data points from treatment pairs that share one species are data-polluting alongside those that are entirely disjoint; for $N = 8$, this data pollution further extends to all the data points from treatment pairs that share two species; for $N = 16$, it extends yet further to treatment pairs with three or four species in common; and for $N = 32$, all pairs with eight or fewer species in common also fail this basic test of evidential relevance. An important consequence is that, although the proportion of entirely disjoint

hundreds of species richness manipulation experiments—data points that are treated as though they constitute evidence that changing an ecosystem's species richness is causally responsible for changing that ecosystem's productivity, even though they actually have no such causal significance.

A dataset so suffused with data points that lack evidential credibility is itself not credible. This places BEF science on the horns of a dilemma: On one side, it appears that there is no escaping the conclusion that the number-changing interpretation cannot make reasonable causal sense of data from all the many experiments that simulate manipulations of species richness. On the other side, there appears to be no alternative interpretation that can make sense of the sweeping causal claim that BEF scientists have striven to support—that species richness *quite generally and without reference to the particular species or combination of species involved*, is a significant cause of ecosystem functioning. Yet BEF science's mission is to establish *this* claim as the "scientific" basis for the claim that diminishment of *species richness* will make ecosystem functioning disintegrate and therefore, somehow, diminish ecosystem services (§2.2).

This pervasive, because unavoidable, contamination of datasets with bad data points undercuts essentially all experimental evidence for species richness' causal influence. Yet a second, independent problem is no less subversive of the evidence. This second problem consists of irresolvable entanglement of causal factors. It arises from the coupling of two complicating factors: (i) It is not possible to intervene on species richness R without also intervening on many other, independently operating causal factors that are highly correlated with R via their *logical/conceptual* rather than empirical coupling with it. And crucially, as we shall see: (ii) It is not possible to disentangle the contribution of these confounding causal variables by intervening on each of them individually without concomitantly intervening on R.

Empirically correlated causal factors not uncommonly arise in experimental setups. Drug trials illustrate why they need not pose insuperable obstacles to effectual experimental test of a factor's causal influence: It may not be possible to intervene (or simulate interventions) on D (drug taken or not) without also intervening (or simulating interventions) on C (capsule taken or not). However, a drug trial may still distinguish the causal contribution of these two factors by intervening on C *without* also intervening on D. That's what enables a drug trial's placebo group to perform its cause-disentangling role, thereby surmounting complication (ii) of the preceding paragraph.

Similar recourse is *not* available to BEF experiments. *They* are subject to complication (i) because intervening on R also *necessarily* (logically and conceptually) involves intervening on multiple causally entangled variables. These variables include some number of the N causal variables S_i (the ith species S_i present or absent) representing *individual* species from a species pool of size N that are added or removed as part of intervening on R. Many more variables arise from some number of the $\sum_{k=2}^{N} \binom{N}{k}$ *combinations* of species that are added or removed as a consequence of intervening on R. For example, the R-changing intervention that adds $N-1$ species to (let us say) one initially present species concomitantly adds, minimally, the $N-1$ causal factors S_i. It also adds many more independent causal factors, including all $\sum_{k=2}^{N} \binom{N}{k}$ *combinations* of two or more species in that pool.[18] The number of added combinations is: $(2^N - 1) - N$, or 11 for $N = 4$; 247 for $N = 8$; 65,519 for $N = 16$; and 4,294,967,263 for

---

treatments among all treatment pairs decreases as N increases, an increase in bad data from not-quite-disjoint pairs compensates for this.

[18] $\sum_{k=2}^{N} \binom{N}{k}$ sums the number of combinations of two species (the starting value for k) taken from among N, the number of combinations of three species from among N, and so on up to the number of combinations of N species (the final value for k) from among N.

$N = 32$, respectively.[19] Each one of these causal factors plausibly operates independently of richness: Each species has a characteristic productivity that is independent of its combination with other species. Each combination of two or more species also makes some characteristic causal contribution. For example, a grass that employs the form of photosynthesis known as "C4 photosynthesis" (rather than C3 or CAM photosynthesis) in combination with a legume is characteristically more productive than the C4 grass alone.[20]

The particular conundrum posed by complication (ii) might suggest that it could be remedied by viewing pairs of treatments that are identical in R but differing in composition as simulating a richness-*preserving* intervention. Nothing comparable is possible for drug trials. BEF experimenters certainly have not sought to do this. And it seems clear that a richness-preserving intervention could not be interpreted as an intervention on *species richness*. But more importantly, data points from *these* simulated interventions would still be irretrievably compromised as causal evidence by the first, dataset pollution problem: Two ecosystems with the same number of wholly different species cannot coherently be regarded as some one particular ecosystem that was intervened upon. Nor would these R-preserving interventions even remedy the second, irresolvable entanglement flaw: Their inclusion could not avert the introduction of a tangle of causal contributions from all the individual S_i as well as all the combinations of S_i that occur only in either the pre-intervention ecosystem or the species-equinumerous, post-intervention ecosystem.

Absent any other suggestion, a remedy for both the dataset contamination problem and the irresolvable entanglement problem might be sought by simply abandoning the number-changing interpretation. This, no doubt, is the nuclear option for BEF science, for it entails also abandoning the sweeping empirical claims, cited in §2.2, for the causal impact of the sheer *number* of species in an ecosystem on its functioning without reference to which species they are. And abandoning *that* thesis, in turn, requires abandoning the similarly sweeping plea (discussed in Section 6) to conserve biodiversity lest declines in the number of species cause declines in ecosystem services.

That said, and although not employed in actual BEF experimental practice, analyses, or statements of findings, one might nonetheless look to a *collection-modifying interpretation* of R-changing interventions in which the only changes to species richness that play an evidential role are ones in which species enter or exit an existing collection. Indeed, this picture might more closely align with actual concerns about biodiversity. However, under this interpretation, R no longer signifies what everyone means by "species richness"—the number of species in an ecosystem. Rather, R now means something relational, tied to some initially defined collection of *particular* species that persist through interventions that add others. Yet imposing the specific restrictions of a collection-modifying interpretation of R-changing interventions *still* falls short of erasing the causal entanglement problem: Even when interpreted in this way, the simulated addition or removal of species from an ecosystem *still* does not permit disentanglement of mul-

---

[19]It is straightforward to show that $\sum_{k=0}^{N} \binom{N}{k} = 2^N$ using the binomial theorem.

[20]The strong and necessary correlation of these many factors to the hypothesized cause R precludes dismissing them as randomly varying background conditions that can be credibly presumed to "average out." The reason resides in violation of Condition (3) of the requirements for causal evidence stated at the end of Section 2. This condition entails that a causal explanation of any observed change in P in terms of R is *not* warranted when the values for other, possibly causally relevant values are set initially, but afterwards permitted to change as a consequence of changes, not in species richness—that is, they are not the instruments through which species richness may be thought to operate on productivity—but rather as a consequence of changes in other causal factors that are *correlated* with species richness. For example, while the presence or absence of a legume may be correlated with levels of species richness, its presence (or absence) is an independent, soil-affecting (and consequently, productivity-affecting) causal factor that is *not* caused by species richness.

tiple, independently operating causal factors that are highly correlated with species richness via logical coupling.

To see why, consider two experimental treatments, the first containing one species ($n = 1$) drawn from a pool of N species. Suppose that this treatment is paired with a second, post-simulated-intervention treatment that adds to it all other $N{-}1$ species in the pool. This pair of treatments satisfies the relational constraint of the collection-modifying interpretive stance in simulating an intervention that is analyzed as a change in species richness $\Delta r = N{-}1$. However, it concomitantly simulates addition of the $N{-}1$ independent causal factors consisting of each of the added species. It also simulates the addition of the other independent causal factors consisting in all $\sum_{k=2}^{N} \binom{N}{k}$ *combinations* of two or more species in that pool. This number (of added combinations) is $(2N{-}1) - N$, or 11 for $N = 4$; 247 for $N = 8$; 65,519 for $N = 16$; and 4,294,967,263 for $N = 32$, respectively. 4.29 billion cannot reasonably be presumed to *not* represent some major causal contribution and importantly, one that is strongly correlated with R, therefore subverting any claim that R is causally responsible. As well, the billions of data points that arise from two treatments where $n = 1$ for the first are joined by billions more that arise from two treatments where $n > 1$ for the first. Finally, the collection-modifying interpretation fares no better than the number-changing interpretation in making available viable means of independently intervening on the independently operating causally confounding variables in order to disentangle their effects from any effects that might be due to species richness.

The details that this section elaborates in order to expose an apparently sweeping failure of BEF experiments to produce credible evidence should not obscure how simple and basic the essence of that failure is: By virtue of the very definition of "species richness," experiments that attempt to manipulate that variable ineluctably produce huge numbers of causally irrelevant data points, which are analytically indistinguishable from all the others. In the history of science, instances of widespread blindness to evidential flaws this conspicuous have most frequently occurred in circumstances such as these, where it was rooted in widespread commitment to affirming one particular result.

4 Additive partitioning

Early critics of BEF research (for example, Aarssen 1997, Huston 1997, Huston et al. 2000, and Huston and McBride 2002) also questioned whether the experimental data truly isolated species richness as a causal factor that operates independently of species composition. In reacting to this challenge, BEF scientists did not see a need to change any essential design element of their experiments. Nor did they see a need to question, modify, or qualify their claim that data from these experiments evidence species richness' causal influence on ecosystem functions. They instead changed only how they *analyzed* these data—following the suggestion of Loreau and Hector (2001) to represent the supposed causal influence of species richness as the simple sum of two (or more) factors, which they refer to as "mechanisms"[21] or "components" of this influence. This representation, in terms of what is called the *additive partitioning equation*, quickly became so integral to every experimental data-based claim for species richness' causal influence and to the defense of these claims that, were this essay not to consider them, its assessment of the evidence for these claims would be quickly (and justifiably) brushed aside.

In the context of the BEF mission to show that species richness causally impacts ecosystem functions, there may be a temptation to view the additive partitioning equation as a mathe-

---

[21]In this context, the term "mechanism" is most sensibly understood in the non-mechanistic sense of Section 2.3: an independently operating underlying or intermediate factor through which species richness is supposed to exert its influence."

matical expression of the idea that changes in levels of ecosystem functions due to additions or removals of particular species or combinations of species are themselves *caused* by changes in species richness. But that would be a conceptual mistake. The stipulated meaning of species richness $= \sum_{i=1}^{N} S_i$ *logically entails* that species richness increases with the addition of some new species. The addition does not *cause* species richness' increase. In much the same way, it is a category mistake to suppose one can cause a circle's area to increase by increasing its radius. That is because (as Archimedes showed) the Euclidean definitions for circles and triangles together logically entail that the area of a circle with radius r is $2\pi r$.[22]

However, the additive partitioning equation's incapacity to contribute causal or even empirical information relevant to the question of species richness' causal influence is made fully apparent by looking at what, exactly, that equation means. In its most basic form, the additive partitioning equation asserts that any observed difference in productivity (or other ecosystem function) associated with a difference in species richness—labeled NE for "the net biodiversity effect"—is a simple sum of two terms—namely, "the selection effect" SE and its conjoined twin, "the complementarity effect" CE:

$$NE = SE + CE \tag{2}$$

Interpreting this equation as representing a causal relationship whereby changes imposed on species richness bring about a net change in productivity NE requires interpreting the "effects" represented by SE and CE as factors that mediate the causal operation of species richness in producing that productivity change. Thus interpreted, SE is supposed to represent the portion of species richness' influence that is due to the particular "selection" of species—whether naturally or through experimental random draw—that count towards an ecosystem's species richness.[23] In contrast, CE has no settled interpretation: It represents whatever portion of species richness' influence that is supposed to be due to any joint action and interaction of organisms that might arise from their specific differences—that is, some open-ended notion of composition.

Because Loreau and Hector's (2001) seminal exposition of the additive partitioning equation frames it as a version of the Price equation (1970), it might be thought that the causal relevance of the former equation can be inferred from the causal relevance of the latter. It is therefore important to dispel that misconception. George Price's original proposal arose from thinking about trait selection in biological organisms and consequent evolution of species. However, Price viewed his equation as a sweepingly general representation of any difference between (a) observed values of some trait in one population of (not necessarily biological) objects and (b) observed values of that trait in a second population of objects, which can be viewed—in some expansive sense not requiring clear-cut causal linkages—as having *descended* from objects in the first, so-called *parent* population. Illustrative is Price's (1995) own example of the population of Modest Mussorgsky's musical *Pictures at an Exhibition*, which (in his interpretation)

---

[22]Whether or not interpretations of the additive partitioning equation by BEF scientists fall prey to this error, they not infrequently do attribute causal significance to logically or conceptually necessary relationships of various sorts. For example, the arrow that James Grace and colleagues (2016, Figure 2) draw from the productivity of a site to the productivity of a plot within that site portrays that mereological relationship as causal. More generally (albeit not universally) accepted is Tilman et al.'s (1998, 282) insistence that statistical artifacts operate as causes in the real world: "Greater stability of more diverse ecosystems is just as real, and just as important, whether it is caused by interspecific competition or statistical averaging." According to them, a statistical average can operate in the real world as a cause of ecosystem functioning every bit as well as some legume's fixing of nitrogen: "the importance of the statistical averaging effect and the negative covariance effect as ecological principles relating stability to diversity comes not from whether they have a biotic or statistical origin but from their very existence."

[23]The selection effect is commonly (although not universally) conceived of as generalizing—by allowing for negative as well as positive effects on productivity—what was earlier called "the sampling effect." That earlier moniker reflects its association with the sampling of the species pool in BEF random draw experiments.

"descended" from a parent population of paintings that the composer viewed at a memorial exhibition of paintings by Viktor Hartmann.

For the purposes of this essay, it suffices to say that the Price equation is a formal representation of a difference in some trait that may be observed in objects in one population with respect to that trait as observed in objects in another population that is somehow *and not necessarily causally* linked to the first. It presents a difference, which, given the observed trait values, is indisputably true—but only by virtue of meanings of the terms that enter into its expression. It is true, that is, in the non-empirical way that the equation $2 + 2 = 4$ is true.

The non-empirical nature of the Price equation is obvious from how it is derived.[24] It starts with an equation that stipulatively *defines* a symbol representing a difference between the respective means of some trait for objects in two linked populations. The stipulated meanings of the terms that appear in this definition ensure that the Price equation can be derived from it via a sequence of algebraically correct transformations, which cannot add empirical content. Because the Price equation only restates the original definition's expression of an equivalence between two different symbolic representations of observed differences in populations of objects, it cannot add causal content.

However, the additive partitioning equation's causal relevance is not thereby doomed because, although Loreau and Hector (2001) tried to model it on the Price equation, they did not succeed in making it a version of the Price equation: The additive partitioning equation violates Price equation semantics in multiple ways, including some that help to illuminate the semantics of the former equation later in this section.[25]

Equation 2 is supposed to be a simplified but equivalent re-presentation of Equation 3:[26]

$$\Delta Y = N \operatorname{Cov}(\Delta RY, M) + N \operatorname{E}(\Delta RY * M) \tag{3}$$

in which NE stands in for Equation 3's ΔY, while the notation $\operatorname{Cov}(\ )$ in Equation 3 makes Equation 2's SE look like a standard statistical covariance and the notation $\operatorname{E}(\ )$ makes CE look like a statistical expectation. The key to understanding what Equation 3 means is that it, in turn, is a provably equivalent re-presentation of Equation 4:

$$\Delta Y = \sum_i^N Y_{Oi} - \sum_i^N Y_{Ei} \tag{4}$$
$$= \sum_i^N RY_{Oi} M_i - \sum_i^N RY_{Ei} M_i$$

Equation 4 stipulatively defines the quantity ΔY to mean: any difference between (i) the sum of actually *observed* yields Y_{Oi} for each species i of plant in multi-species assemblages that draw on a pool of N distinct species, and (ii) the sum of (as we shall see, labeled with crucial ambiguity) *expected* (*not* observed) yields Y_{Ei} for each species i of plant in multi-species assemblages.

---

[24]Space does not permit presentation of the Price equation's derivation here. The best references are Price (1970, 1972, 1995) himself, Frank (1997a, 2012), Fox (2005, 2006 Appendix A), van Veelen (2005), and van Veelen et al. (2012), although idiosyncratic choices for key symbols make it a challenge to line up these various sources.

[25]The literature on the Price equation in all of biology is extremely thin, and few biologists that I've quizzed have even heard of it. Some number of BEF scientists are aware of the Price equation, but mostly only by virtue of their awareness that it is supposed to be a model for additive partitioning. Fox (2005) is one of the few BEF scientists who evidences some awareness of the Price equation's semantics and points out one respect in which additive partitioning deviates from it, although not the deviations that are key to this essay's discussion.

[26]Equation 3 re-presents what Loreau and Hector (2001) present in their Box 1 with some minor notational changes. Its meaning and the meaning of its symbols are explained just below, via Equation 4 from which it is derived.

Y_{Oi} and Y_{Ei}, in turn, are quantities that (as shown on the second line) are derived by weighting species i's yield in monoculture (M_i) by its proportional contribution to yield in multi-species assemblages. RY_{Oi} is the proportion of the total yield that species i is actually *observed* to contribute in those mixed assemblages. On the other hand, RY_{Ei}, like Y_{Ei}, is not an observed quantity. Rather, it is *stipulatively* defined—as the proportional contribution of species i to the yield, computed on the basis of the surely false assumption—the experimental null hypothesis—that no plant's yields are ever affected by the presence of other plants.[27] In Equation 3, ΔRY merely denotes the difference between RY_E and RY_O in abstraction from the particular contribution of each species i.

A few observations regarding Equation 3's derivation from Equation 4 make it apparent that neither is relevant to the question of whether or not species richness causally determines productivity. First, as already observed, Equation 4 merely *stipulates a definition for* a symbol (ΔY). When yields do, in fact, deviate from the null hypothesis—as they very commonly will—the symbol ΔY merely gives us a notational convention for compactly expressing this fact. But its use cannot and does not add empirical information to the already given fact. Illustrative is the reported response of the great Dutch footballer Johan Cruijff to the question of how he won games. Paraphrasing Cruijff: "By scoring one more goal than your opponent" (van Veelen et al. 2012). This "explanation" merely invokes the football convention—whereby outscoring one's opponent in a match constitutes winning. Consequently, Cruijff's response supplies no empirical information. Equation 4 is also merely a descriptive convention, which also cannot add empirical content to the state of affairs that it describes.

Mere reference to a convention is particularly evident in the definition that is stipulated for RY_{Ei}—species i's proportional contribution to yield. As noted above, the relative yields RY_{Ei} are *not* observed quantities and are *not*, the subscript E notwithstanding, even expected in the customary sense of "what one should expect on the basis of empirical evidence." That is because they are *computed* on the basis of the null hypothesis that a plant's yields are never affected by the presence of other plants. And a long-standing mountain of evidence suggests that this hypothesis is routinely violated: Nitrogen-fixing legumes can promote the growth of many neighboring plants, while plants that produce allelochemicals (mostly excluded from experimental species pools) can suppress the growth of many others. Consequently, the sensible way to interpret any difference between yields computed with RY_{Ei} and actually observed yields, RY_{Oi}, is that it reflects the well-established fact, owing nothing to BEF research, of widespread violation of the assumption underlying RY_{Ei}'s computation. It makes little sense to interpret any such difference as reflecting a *change* in some actual property of actual ecosystems, let alone the operation of some *causal* factor.

This basic interpretive error may be obscured by multiple additional errors that Loreau and Hector (2001) make with respect to RY_{Ei} in taking the Price equation as their model. As indicated above, the Price equation simply describes how observed values of a *trait* in the *objects* in a *parent* population relate to observed values of that trait in a *descendent* population linked to it in Price's expansive sense. In a species richness experiment, the *trait* that differs between parent and descendent populations clearly must be the *yield* of the varying assemblages of organisms in experimental treatments. The objects bearing these traits can only be *assemblages*. Therefore,

---

[27] At this point of the argument, the observation that this assumption is false serves only to emphasize the thoroughly conventional character of ΔY's definition. The immediately ensuing argument does *not* draw any conclusions from the mere facts that (i) the null hypothesis is (like many null hypotheses) generally false and that (ii) this hypothesis enters into the additive partitioning question. Rather, the argument revolves around the other facts that (iii) the additive partitioning equation is a necessary truth about fictional data and that (iv) a necessary truth about fictional data cannot say anything about whether actual data from actual observations are causally significant.

in Price equation terms, a BEF experiment's *parent and descendent populations* are populations of experimentally created assemblages.

Given that, we should ask: What would the *parent* population (of assemblages) be? BEF thinking on this is reflected in Fox's (2005) characterization of the additive partitioning equation. According to him, it depicts "monoculture biomass as a trait on which selection can act." If follows that the parent population is the population of assemblages in which monoculture biomass is a trait. That population is simply the population of monoculture assemblages. And from this parent population of monocultures, multi-species assemblages are supposed to descend. In fact, BEF experiments generate just such a descendent population—via random draw of species from a pool. This makes it clear that the "selection," which Fox supposes to explain any overall yield difference ΔY between these populations of assemblages, consists in nothing other than selection by random draw of species from a pool to create multi-species assemblages in a BEF experiment.

The fact that the additive equation treats particular assemblages as members of linked populations raises challenging questions: First, what, exactly, is the *biological* meaning of the "selection" that is supposed to operate on one population to produce the other? And how, exactly, are the monocultures in the parent population supposed to *causally* influence the multi-species assemblages in the descendent population? These two questions have no obvious satisfactory answer. But I will set them aside to focus on questions about the stipulation that RY_{Ei} in Equation 4 means the *frequency* with which various values of the yield trait occur in the *parent population*, which as we have seen, is a population of monoculture assemblages. By assigning RY_{Ei} this meaning, Loreau and Hector (2001) sought to give RY_{Ei} the role that, in the Price equation, is played by the proportion of the *parent* population for which the given trait is *observed* to have a particular value. However, this gives rise to multiple incongruities:

(i) RY_{Ei}, according to the definition that Loreau and Hector's (2001) stipulate for it, is *not* a frequency of occurrence. Loreau and Hector (2001) may obscure this fact about RY_{Ei}'s assigned *meaning* by at one point also assigning it a *value*—equal to the proportion of species i organisms in an assemblage's total. But asserting that some thing assumes some *value* does not alter what it *is*. We cannot change that a Chiriqui harlequin frog is a variegated amphibian by assigning it one of its many possible color values.

(ii) RY_{Ei} does not refer to the *parent* populations—the monocultures—for which it is supposed to be a frequency of occurrence. Rather, it refers to *descendant* populations of multi-species assemblages by way of invoking the null hypothesis.

(iii) RY_{Ei}'s values are not actual frequencies of occurrence of trait values. Actual frequencies, but not RY_{Ei}'s values, are determined by *observing* actual values and how frequently they occur in some actual population of objects.

(iv) RY_{Ei} is not even an *expected* quantity in the sense of "what we can rationally expect" because it expresses the null hypothesis, which is known to be false. Nor is it an expectation in the standard statistical sense because that would require it to be the arithmetic mean of a set of independent *observations* of an independent variable that varies over an actual population of objects.[28]

One remaining mystery concerns the appearance of the covariance and expectation operators in the Equation 3's presentation of additive partitioning. In statistics, these operators have a

---

[28]The observations required to make RY_{Ei} a true statistical expectation would presumably have to be observations of species i's yield in assemblages that include some unspecified number of unspecified other species.

meaning that relates to *observed* values. Yet no such content could possibly come from Equation 4, which merely stipulates what ΔY means, while Equation 3's derivation from it involves only algebraic manipulation. This mystery is dissolved by noticing that, in Equation 3, these statistical operators do *not*, in fact, have the statistical *meanings* that they standardly have in probability theory or statistics:

An expectation in statistics is a frequency-weighted average for all the *observed* values for a variable that characterizes objects in *some one particular population*. The expression $\mathrm{E}(\Delta RY * M)$ in Equation 3 does not satisfy this semantic requirement. ΔRY refers to differences in a relative yield variable between two *different* populations, the second of which—a descendent population of multi-species assemblages—is a fictional construct, not something observable, let alone actually observed.

For its part, a covariance in statistics is a property of, or relationship between, the *observed* values of two variables, each of which varies randomly *over one, single population of objects*. It summarizes the direction and magnitude of one variable's observed deviation from *its* mean in that one population with respect to the direction and magnitude of the second variable's observed deviation from *its* mean in the *same* population. $\mathrm{Cov}(\Delta RY,\ M)$ in Equation 3 does not satisfy *this* semantic requirement, and for reasons similar to why $\mathrm{E}(\Delta RY * M)$ does not satisfy the semantic requirements for an expectation: Its variables span different, parent and descendent populations, while the values for one variable are not even observed but rather computed on the basis of the null hypothesis.[29]

In short, neither the notation $\mathrm{E}(\ )$ nor the notation $\mathrm{Cov}(\ )$ on the right hand side of Equation 3 have the statistical meanings that these notations suggest. Rather, they are semantically misleading syntactic shorthand for expressions that derive from Equation 4 solely via algebraic manipulation.

Summing up: Equation 4, the starting point for the additive partitioning equation's derivation, is a stipulative definition (for ΔY). The additive partitioning equation, Equation 3, is merely an algebraically equivalent re-presentation of Equation 4. The truth of both is independent of contingent facts about the actual world.[30] That means that they are true independent of whether or not species richness causally influences productivity; they contribute no information one way or the other. In fact, these equations cannot even express predictive relationships: Prediction depends on reliable correlations between two factors that are *observed* in the real world. This requirement is not satisfied by the "expected" yields Y_{Ei}, which are computations based on a proposition—the null hypothesis—that is routinely contravened by what is actually observed. This circumstance even strips the additive partitioning equation of *descriptive* value.

A remarkable feature of additive partitioning's incapacity to serve as a means of causal analysis is the sheer number of missteps integral to lending it any appearance of causal credence. It is difficult to imagine how every one of those errors could have been overlooked in the face of scrutiny unrestrained by prior commitment to the causal hypothesis it was supposed to prop up.

5   Functional diversity as cause

BEF researchers universally regard the 'B' in BEF research as standing for biodiversity, not for any one particular representation of it, such as species richness. So starting in the early 2000s, swift embrace of *functional traits* as an organizing concept in ecology led to increasing reinter-

---

[29]While the fact that the null hypothesis is undoubtedly false is not central to this argument, it reinforces the salient point that that hypothesis does not apply to any actual, mixed-species assemblage.

[30]There should be no doubt that it is possible for an equation to be *false* independently of all contingent facts, as is $2 + 2 = 5$.

pretation of biodiversity in terms of *functional diversity*. This trend manifested in BEF research as some efforts to establish the causal influence of biodiversity pivoted from species richness to functional diversity.[31] BEF scientists cited in this section make clear their supposition that functional diversity is a significant cause of ecosystem functions by variously stating that it is "an important *driver* of ecosystem functioning" (Villéger et al. 2008), that it "can *influence* ecosystem functioning" (Clark et al. 2012), that it "*control[s]* biodiversity *effects* on biomass production" (Flynn et al. 2011), that functional diversity can (causally) explain ecosystem functioning when richness does not (Cadotte et al. 2011), that functional diversity "*enhances* ecosystem functions such as primary productivity" (Schittko et al. 2014), and that "functional diversity [is] a *key factor to maintain* important functions and services of ecosystems" (Laureto et al. 2015, 112). Schittko et al.'s supposition that functional diversity's supposed causal determination of ecosystem functions gives warrant for its conservation (via some unexplained connection to ecosystem services) is echoed by others, including Caddotte et al. (2011, 1079) who state, "The goal of conservation and restoration activities is to maintain biological diversity and the ecosystem services that this diversity provides."

It is a comparatively simple task to analyze the credibility of evidence for this causal hypothesis: As BEF researchers such as Lefcheck and Duffy (2015) themselves note, support for this thesis derives almost entirely from post hoc analysis of data produced by prior attempts to manipulate *species richness*. Little effort has been made to produce direct experimental evidence via experimental manipulation of functional diversity as a candidate causal variable.[32]

Crucial to evaluation of any causal thesis is a generally agreed-upon, precise characterization of the candidate causal variable. It is therefore remarkable that no such characterization of functional diversity exists. Instead, there is general agreement only on three characterizing elements, which are so undemanding that they can be embodied in (literally) infinitely many ways that give rise to infinitely many ways to measure functional diversity, which yield assessments that often radically disagree and even conflict.[33] The three agree-upon elements are present in Díaz and Cabido's (2001, 654) commonly cited definition. According to them, *functional diversity* is:

> the value and range of functional traits of the organisms present in a given ecosystem.

The first element is a tacit and questionable empirical assumption—the assumption that every individual organism is a repository of some fixed collection of traits, each with a fixed value that is wholly determined by that organism's species. This assumption is routinely violated, for example because (as is well known) many organisms are phenotypically plastic or vary greatly with the biotic and abiotic conditions in which they live. The second element is another empirical assumption—the assumption that these species-determined traits, with their species-determined values, causally determine functional properties of ecosystems in a way that makes

---

[31]A number of BEF scientists—for example Villéger et al. (2008) and Laureto et al. (2015)—have themselves called attention to the increasing attention paid to functional diversity in BEF research.

[32]Lefcheck and Duffy (2015) emphasize the uniqueness of their own work by pointing out that but two other experiments had attempted to manipulate functional diversity as a candidate causal variable. Because those two experiments involved but two species apiece, the credibility of their causal evidence is close to nil.

A better known experiment, the Jena Experiment (Roscher et al. 2004, Weisser et al. 2017), classified the species in its pool into functional groups, and "measured" functional diversity as the number of functional groups present (one of an infinitude of possible measures). But it employed this classification only to constrain the selection of species. That means that it did not truly manipulate either species richness (which, by definition, takes no account of functional groupings) *or* functional diversity (which, by definition, takes no account of how many species contribute the functional traits). See also Note 14.

[33]An example of how two of the most commonly used measures can differ without bound is provided below.

the organism supplying these traits/values wholly interchangeable with any other or combination of the many other organisms that could also supply them.

The third element is an accompanying definition of *functional trait*, the source of functional diversity, which Díaz and Cabido (2001, 654) define as:

> the characteristics of an organism that are considered relevant to its response to the environment and/or its effects on ecosystem functioning.

This definition makes functional traits out of most traits of most organisms: The preponderance of traits that are of ecological, evolutionary interest, or even general biological interest in *some* way enter into how organisms respond to the particular conditions in which they find themselves. And it is difficult to think of biologically interesting traits that do not collectively affect at least one ecosystem-wide property or ecosystem function.

Permissiveness in assessing an ecosystem's functional diversity is further compounded by the fact that its definition places no constraint whatever on which of dozens of functional traits (Roscher et al. 2004; Cornelissen et al. 2003) of organisms to include and which ones to exclude. This license is reflected in the wide range and disparity of traits that different studies actually use to compute an ecosystem's functional diversity. Yet addition or removal of just one trait can dramatically alter what an ecosystem's functional diversity is supposed to be. Consider a Euclidean distance-based metric (discussed below) with trait T_1 for species $\{A, B, C\}$ with values $\{0.5, 1.0, 2.0\}$, respectively; and trait T_2 with values $\{10, 0, 10\}$. Based on T_1 alone, the Euclidean distance $d(A,\ B) = 0.5$, $d(B, C) = 1.0$, and $d(A,\ B) < d(C,\ D)$. Yet when T_2 is included, $d(A,\ B) = 10.0125$ and $d(B,\ C) = 1.5$: Adding T_2 reverses the inequality.

Permissiveness in assessing an ecosystem's functional diversity is yet further compounded by lack of any constraint which few *species* of any ecosystem's multitude of species to include and which many others to exclude. As with traits, no study selects for inclusion anything but a tiny subset of species that might occur in an ecosystem; and that set varies widely from study to study. BEF scientists routinely employ the phrase "*the* functional diversity of an ecosystem" to refer to functional diversity as a generally operating cause. Yet contrary to what the definite article suggests, every computation of an ecosystem's functional diversity is relative to a choice of tiny minorities of its species and tiny minorities of the functional traits that are supposed to inhere in those organisms. This indeterminacy is made radical by the fact that there appears to be no independent test of the biological soundness of any of the myriad, sanctioned choices: No one can say what an ecosystem's functional diversity actually is apart from its measurement by some measure.

Additionally, species and traits are but two sorts of unconstrained choices among others[34] that contribute to a literal infinitude of choices that engender an *embarras de richesses* of measures, which BEF scientists have proposed or actually utilized (Petchey and Gaston 2006; Mouchet et al. 2010). I shall focus on two of the most discussed and utilized. These fall into the category of *'M'-type measures*, which represent any species-determined value of any real-valued, species-determined functional trait as the coordinate value for one dimension of a continuous, multidimensional, *M*etric space. The location of any species in the space is given by the coordinate values for (typically) multiple traits. The metric defined on the space then supplies the distance between any two species-defined points. That distance is interpreted as a measure of the species' differentness—the quantity of diversity that they embody.

---

[34]Among other choices that enter into functional diversity's measurement are: Use (or not) of species abundances, how (if at all) discrete-valued traits are combined with continuously varying ones, use (or not) of a metric space, use (or not) of dendrograms, and for the latter two, which of an unbounded number of ways to perform the computation over the distances that are supposed to represent functional differentness.

Numerous problems arise with this proposal. One concerns discrete-valued variables because (a) a continuous, real-valued metric space is incapable of representing them and (b) they often have no sensible numeric representation at all. Yet many of the most salient functional traits are discrete-valued: season of first leafing, nitrogen-fixing or not, type of photosynthetic pathway, etc. BEF scientists routinely represent these fixed values with arbitrarily assigned integers—for example, photosynthetic pathways as $\{C3 = 1,\ C4 = 2,\ CAM = 3\}$—and treat the numeric differences as biologically meaningful, real number-valued representations of degrees of differentness.

Another problem arises from the metric. BEF scientists rarely if ever even discuss their *choice* of metric; they simply presume that trait spaces are Euclidean. Yet there appears to be no *biological* basis for presuming that accurate representation of differentness among species compels choice of a Euclidean metric. Why (biologically speaking), must the metric even be isotropic, homogeneous, and flat?[35] Yet non-Euclidean metrics can always be found that, for a given set of coordinates, yield distances and so functional diversity numbers that differ from the Euclidean assessment by any pre-specified amount.

The gravity of these and other problems is compounded by the fact that, even within the domain of 'M'-type measures and even with respect to the *same* set of species and the *same* set of functional traits, conflicting assessments of functional diversity abound. Two of the most popular 'M'-type measures are illustrative.[36] Rao's quadratic entropy Q (Rao 1982; Ricotta 2007) is a schema for computing functional diversity as the sum of pairwise trait-based Euclidean distances between pairs of species located in a Euclidean space. Q weights these distances by the species' abundances even though there appears to be no biological basis for determining whether or not abundance weighting improves the accuracy of functional diversity assessments. Schumacher et al. (2009), Roscher et al. (2012), Clark et al. (2012), Song et al. (2016), Funk et al. (2017), Weiser et al. (2017) all make use of Q.

Functional Richness or $FRich$ (Lefcheck and Duffy 2015; Schittko et al. 2014), on the other hand, like most 'M'-type measures but unlike Q, ignores abundances. However, like Q, $FRich$ starts by computing a set of points in an abstract metric space. Each point "locates" one of the tiny number of species chosen for $FRich$'s computation, utilizing that species' pre-chosen collection of traits, each assigned the value that all organisms of that species are presumed to have. $FRich$ then veers into topological territory, focusing on the *convex hull* that this collection of species-points defines. Intuitively: if you draw lines that connect every pair of points in a space, then their convex hull is the minimal subset of points that forms an "envelope" fully containing all the lines. According to $FRich$, the ecosystem's functional diversity is the volume of this convex hull. Yet nothing in *biology* explains why computing the volume—with a biologically arbitrary metric—of the convex hull of points that represent an arbitrary sample of an ecosystem's species with an arbitrary sample of their traits and species-determined values produces the biologically correct value of that ecosystem's functional diversity.

Both Q and $FRich$ (and dozens more computational schemas that this discussion omits) evidently lack the biological mooring that would make them something more than a highly unconstrained and biologically dubious computational exercise. For example, utilizing $FRich$ to

---

[35]An isotropic metric defines a space that is the same in every direction. A homogeneous metric defines a space that is the same in every region and around every point.

[36]Space precludes discussion of another 'M'-type measure, Functional Diversity (*FD*), which several of the references cited here employ. It is defined as: the total number of segments in all leaf-to-root paths of the hierarchically classifying (dendrogram) structure computed by employing some one clustering algorithm to cluster resident species based on their pairwise distances as points in a metric space with coordinates that are the trait values ascribed to them. While its computation is more complicated that the computation of Q and *FRich*, none of this additional complexity alters any of the substantive points made with reference to Q and *FRich*.

measure functional diversity commits one to the principle that an ecosystem containing species that differ vastly in some one functional trait—for example, plants that reach multiple meters in height versus others that never exceed a centimeter or two—has no functional diversity at all, should no other trait differ: A one-dimensional difference in volumes, no matter how stupendously great, is no difference at all.

Similarly vexing is that $FRich$ and Q provide assessments of functional diversity that, it is easy to see, diverge *systematically and without bound*: Neither adding nor deleting a point located inside the envelope formed by $FRich$'s convex hull affects its assessment of functional diversity. Yet the total of the added or deleted point's pairwise distances to all other points are added to or subtracted from Rao's quadratic entropy Q.

When NASA's Wilkinson Microwave Anisotropy Probe (WMAP) and The European Space Agency's Planck spacecraft returned independent measurements of 13.722 and 13.82 billion years, respectively, for the age of the universe, astronomers regarded their relative closeness as encouraging in light of much greater disagreement among previous measurements. Yet they took the 0.5% discrepancy that still remained to decisively signal the need for further work to "get it right." In contrast, while Q and $FRich$ can differ by orders of magnitude greater than 0.5%, BEF scientists appear to have no sound scientific basis for "getting it right" or even for accounting for these measures' radically diverging assessments of functional diversity. These possibilities are precluded by the facts that Q and $FRich$ both fully qualify as measures of functional diversity and that functional diversity is characterized only as "the" property that is measured by qualifying measures.

The implications of these conundrums are profound. BEF scientists claim that the functional diversity property of an ecosystem is a causal determinant of its ecosystem functioning. Validation of this claim minimally requires knowing the value of this causal variable. Yet there is no independent, biologically grounded basis for adjudicating between the radically differing infinitude of assessments that are on offer.

Faced with this multiplicity of conflicting measures and no sound biological basis for choosing from among them, BEF scientists propose to resolve this radical indeterminacy by "[comparing] their performances" (Clark et al. 2012) and selecting those that are "recommended" (Mouchet et al. 2010, 867) or "most suitable" (Villéger et al. 2008, 2290). They proceed to define a measure's "performance," its "recommendability," or its "suitability" quite straightforwardly in terms of how well it correlates with the productivity numbers found in previously performed attempts to experimentally manipulate *species richness* (as observed at the beginning of this section).

This method for retrospectively "finding" how productivity is "affected" by functional diversity is a paradigmatic form of "HARKing" (formulating the Hypothesis After the Results are Known)—the scientific embodiment of the Texas Sharpshooter. Facing the side of his barn, that character blindfolds himself, pulls out his six-shooter, and then riddles the barn with holes with a random spray of bullets. Removing his blindfold, he achieves the highest possible correlation of shots to target hits by drawing those targets around the holes so as to minimize their distances from bullseyes. Of course, no one who is aware of how this remarkably high correlation was achieved would place credibility in the Texas Sharpshooter's story—that he *caused* this to happen by virtue of his remarkable marksmanship.

Functional diversity researchers also achieve and are quite candid about achieving some high degree of correlation between functional diversity and productivity by retrospectively constructing a "high-performing" measure that "draws" the most advantageously located targets around previously-collected productivity numbers. No one who is aware of how *this* high correlation is achieved should place credibility in the BEF scientific causal story—that it came about by virtue

of functional diversity's remarkable ability to (in the words of the scientists cited above) "influence," "control," "enhance" or be a "key factor to maintain" ecosystem functions and services.[37] It is difficult to explain how BEF scientists gave credence to this story without, once again, referring to their ubiquitous claim that conservation of biodiversity is of paramount importance on account of its supposed causal impact on ecosystems (Section 1).

6   The ecosystem service rationale for conserving biodiversity

Throughout the more than two decades that BEF researchers have claimed to have evidence for the thesis that biodiversity is a major causal determinant of ecosystem functions, these scientists have routinely expressed their conviction that these efforts secure "scientific" warrant for conserving biodiversity. It is therefore remarkable that, to my knowledge, none of these scientists and no conservation organization has articulated a coherent argument in support of this conviction.[38] Instead, the passages quoted in §2.2 and in Section 5 exemplify what is ubiquitous: Right on the heels of warning that declines in biodiversity will cause declines in ecosystem functioning and with no intervening logic, they then claim that this decline in functioning entails a decline in ecosystem services, which conservation of biodiversity can prevent.

   Because much of the world's thinking about biodiversity and its conservation is beholden to this routinely expressed idea, it is important to see if *some* coherent argument for it can be spelled out. What follows is an attempt to do that by trying to think of all the steps that seem indispensable for arguing that, as scientists and conservationists continually suggest, biodiversity's conservation is warranted on the grounds that its diminishment will diminish ecosystem functioning and consequently also ecosystem services. As in earlier sections of this essay, this one references productivity as a representative ecosystem function.

Empirical Premises:

(P1) Higher levels of biodiversity in ecosystems elevate their productivity; lower levels reduce it.

   This is the central causal thesis of BEF science that this essay has discussed.

(P2) P1 applies to the world's unmanaged ecosystems because the causal effects of the appearance and disappearance of species in them is sufficiently similar to their random selection in experiments.

   P2 has drawn substantial prior critical comment. Doubts arise from evidence that sequences of species disappearances and introductions in actual ecosystems are not at all random but rather appear to depend on a variety of factors, including how the presence of any given species affects others (Wardle 2016, Srivastava and Vellend 2005, Solan et al. 2004, Vandermeer et al. 2002, Fridley 2001).

---

[37]Correlations found via retrospective data mining can legitimately suggest hypotheses that are worthy of subsequent scientific investment in experiments that *then* subject them to legitimate test. However, correlations of this sort are not, and should not be mistaken for, the experimental test itself.

[38]Not even among the thousands of documents in the IPBES library (Section 1) can I find a serious attempt to flesh out this idea for an argument into an actual argument.

Normative Premises (normative terms are *italicized*):

(P3) An ecosystem's empirically observable higher levels of productivity mean that it is *higher performing* in the normative sense that it is better at meeting a standard for how it *ought* to perform.

The argument's normative conclusion requires a premise like P3, which defines a general *norm* for how an ecosystem *ought* to perform in terms of how productive it is rather than some empirical claim about (say) higher productivity numbers.

Obviously, some ecosystems are more productive than others. But ecosystems do not resemble automobiles for which standards of performance are based on specifications that indicate the degree to which they fulfill their human-designed purposes. Such specs ground the claim that your Porsche Spyder is *higher performing* than my Volkswagen "Bug" with respect to acceleration. Similar standards also apply to human-designed agricultural systems and may ground claims that their greater productivity makes them higher performing. But there exists no similar, biology-based specification that would ground norms for the performance of natural ecosystems.

(P4) The *higher performance* of more highly productive ecosystems (according to P3) is specifically due to their providing higher levels of *service*.

It is well known and I am not the first to observe that P4 (which the argument requires to hold quite generally) is routinely contravened. Blooms of algae in aquatic ecosystems and proliferation of kudzu in terrestrial ecosystems make them extraordinarily productive. In these and many other cases, high levels of productivity are responsible for—in the non-causal sense of "constitute"—*dis*services.

(P5) An ecosystem's provision of *services* is *good* by virtue of serving the interests and satisfying the desires of persons willing to pay for them.

P5 is the cornerstone principle of economic valuative thinking, which presumes that a measure of a thing's *goodness* is the degree to which persons vouch for it as desirable. P5 may be questioned most centrally because many people desire and vouch for their desire with a willingness to pay for things that are, overall, *bad*. One only need consider the dire consequences of desire for petroleum products and palm oil.

(P6) Ecosystems that provide higher *service* levels are *better*.

The argument requires P6 in order to make the supposed ability of biodiversity to enhance ecosystem services a reason to conserve it. However, P6 follows from P5 only by assuming that the *goodness* of a thing is generally proportional to how much of it there is. But more is generally not *better*. That is amply illustrated by mundane facts such as that it's better to stop eating when you're full.

(P7) An ecosystem's *goodness* preeminently consists in its provision of *services*.

P7 asserts that the econometric, *service*-providing capability of an ecosystem over-shadows all other *goodness*-conferring factors and therefore settles the question of whether or not it is *good* overall. If there are any grounds for thinking that it is true, I am unaware of any. Consideration of its combination with P8 (just below) casts doubt on there being any.

(P8) Conservation *ought to* promote ecosystem *services* by undertaking projects that make ecosystems as high *performing* as they can be in this *service*-providing way.

P8 asserts that an opportunity to increase an ecosystem's *service* levels engenders a presumptive *obligation* to bolster its service-providing capabilities. Like P7, P8 on its own is a questionable normative principle. In combination with P7, P8 encourages disregard for whatever else might be *good* about an ecosystem. In other words, it affords no consideration for species, populations, and organisms that either make no evident contribution to any *service* or that are responsible for *disservices*. It therefore arguably counsels that the promotion of *services* trumps eradication of organisms, if that is what it takes to underwrite strongly desired services. It also welcomes terraforming practices, from the addition, removal, and modification of soil and nutrients to the transformation of major land and water features, if these measures would ratchet up *performance*.

(P9) We *ought to* act so as to promote higher levels of biodiversity (and counter lower levels of biodiversity).

Conclusion:

(C) We *ought to conserve* biodiversity.

That we *ought not* to allow levels of biodiversity to diminish is a straightforward corollary of P9.

There may be other ways to construct the sought-for argument and their details may vary. But all of them must bridge the gap from the causal premise about biodiversity to a normative one about services and from there to service provision as the preeminent object of conservation. Because of that, I believe that the version offered here exposes the fragility of assumptions and logic, which cannot be avoided in arguing that that biodiversity ought to be conserved lest ecosystem services be attenuated. Even aside from P1, which, this essay has argued, lacks credible evidence, multiple other assumptions are questionable and multiple steps appear to be logical missteps. If that is so, then the long-standing *conservation* mission of BEF research, which, for so long has worked so hard to secure P1, was always doomed to fail—even aside from failure of the *empirical* mission to secure that premise.[39]

---

[39]Of course, unsoundness of the ecosystem-services-based argument for conserving biodiversity would not preclude the possibility that some different argument for conserving biodiversity might be sound.

7 Perspective: Implications for biodiversity research and conservation

Summing up: The enormous and enormously influential program of BEF research has been a decades-long mission to find evidence for the thesis that biodiversity causally determines ecosystem functions. That empirical investigative mission was pursued in the service of another, conservation, mission to "save" biodiversity. Both appear to have been undermined by multiple mistaken assumptions that led to multiple missteps. If many of the defects, in retrospect, seem obvious, then the blindness of the scientists to them may well be explained by supposing that their emphatic commitment to these dual missions made them unable to say "no" to the question of whether evidence actually supported their faith in biodiversity as a principal determinant of ecosystem functions.

In short, the BEF research program may be one of those uncommon yet not so rare historical episodes in which deep commitment to an idea from outside science commandeered critical evaluation of the evidence and subsequently led science off the tracks. In this case, the extra-scientific idea that derailed BEF science is the arguably noble one of saving biodiversity. But that idea is morally hollowed out by making it ride, not on whether the many different organisms whose existence is at stake can simply live their lives, but rather on how capably they serve people.

Acknowledgments

Literature cited

Aarssen, L. W. 1997. "High Productivity in Grassland Ecosystems: Effected by Species Diversity or Productive Species?" *Oikos* 80 (1): 183–184. https://doi.org/10.2307/3546531.

Bell, T., A. K. Lilley, A. Hector, B. Schmid, L. King, and J. A. Newman. 2009. "A Linear Model Method for Biodiversity-Ecosystem Functioning Experiments." *The American Naturalist* 174 (6): 836–849. https://doi.org/10.1086/647931.

Bell, T., J. A. Newman, B. W. Silverman, S. L. Turner, and A. K. Lilley. 2005. "The Contribution of Species Richness and Composition to Bacterial Services." *Nature* 436: 1157–1160. https://doi.org/10.1038/nature03891.

Cadotte, M. W., K. Carscadden, and N. Mirotchnick. 2011. "Beyond Species: Functional Diversity and the Maintenance of Ecological Processes and Services." *Journal of Applied Ecology* 48 (5): 1079–1087. https://doi.org/10.1111/j.1365-2664.2011.02048.x.

Cardinale, B. J., J. P. Wright, M. W. Cadotte, I. T. Carroll, A. Hector, D. S. Srivastava, M. Loreau, and J. J. Weis. 2007. "Impacts of Plant Diversity on Biomass Production Iincrease Through Time Because of Species Complementarity." *Proceedings of the National Academy of Sciences USA* 104 (46): 18123–18128. https://doi.org/10.1073/pnas.0709069104.

Cardinale, B. J., D. S. Srivastava, J. E. Duffy, J. P. Wright, A. L. Downing, C. Jouseau, M. W. Cadotte, et al. 2009. "Effects of Biodiversity on the Functioning of Ecosystems: Summary of 164 Experimental Manipulations of Species Richness." *Ecology* 90 (3): 854. https://doi.org/10.1890/08-1584.1.

Cardinale, B. J., K. L. Matulich, D. U. Hooper, J. E. Byrnes, E. Duffy, L. Gamfeldt, P. Balvanera, M. I. O'Connor, and A. Gonzalez. 2011. "The Functional Role of Producer Diversity in Ecosystems." *American Journal of Botany* 98 (3): 572–592. https://doi.org/10.3732/ajb.1000364.

Cardinale, B., K. Gross, K. Fritschie, P. Flombaum, J. M. Fox, C. Rixen, J. van Ruijven, et al. 2013. "Biodiversity Simultaneously Enhances the Production and Stability of Community Biomass, but the Effects are Independent." *Ecology* 94 (8): 1697–1707. https://doi.org/10.1890/12-1334.1.

Cardinale, B. J, J. E. Duffy, A. Gonzalez, D. U. Hooper, C. Perrings, P. Venail, A. Narwani, et al. 2012. "Biodiversity Loss and Its Impact on Humanity." *Nature* 486: 59–67. https://doi.org/doi:10.1038/nature11148.

Clark, C. M., D. F. B. Flynn, B. J. Butterfield, and P. B. Reich. 2012. "Testing the Link between Functional Diversity and Ecosystem Functioning in a Minnesota Grassland Experiment." *PLoS ONE* 7: e52821. https://doi.org/10.1371/journal.pone.0052821.

Cornelissen, J. H. C., S. Lavorel, E. Garnier, S. Díaz, N. Buchmann, D. E. Gurvich, P. B. Reich, et al. 2003. "A Handbook of Protocols for Standardised and Easy Measurement of Plant Functional Traits Worldwide." *Australian Journal of Botany* 51: 335–380. https://doi.org/10.1071/BT02124.

Davidson, E., J. P. Rast, P. Oliveri, A. Ransick, C. Calestani, C-H Yuy, T. Minokawa, G., et al. 2002. "A Genomic Regulatory Network for Development." *Science* 295 (5560): 1669–1678. https://doi.org/10.1126/science.1069883.

Díaz, A., and M. Cabido. 2001. "Vive la Différence: Plant Functional Diversity Matters to Ecosystem Process." *Trends in Ecology & Evolution* 16 (11): 646–655. https://doi.org/10.1016/S0169-5347(01)02283-2.

Dunne, J. A. 2006. "The Network Structure of Food Webs." In *Ecological Networks: Linking Structure to Dynamics in Food Webs*, edited by M. Pascual, and J. A. Dunne, 27–86. Oxford: Oxford University Press.

Dupré, J. 2013. "Living Causes." *Proceedings of the Aristotelian Society Supplementary Volume* 87 (1): 19–37. https://doi.org/10.1111/j.1467-8349.2013.00218.x.

Flynn, D. F. B., N. Mirotchnick, M. Jain, M. I. Palmer, and S. Naeem. 2011. "Functional and Phylogenetic Diversity as Predictors of Biodiversity–Ecosystem-Function Relationships." *Ecology* 92 (8): 1573–1581. https://doi.org/10.1890/10-1245.1.

Fox, J. W. 2005. "Interpreting the 'Selection Effect' of Biodiversity on Ecosystem Function." *Ecology Letters* 8 (8): 846–856. https://doi.org/10.1111/j.1461-0248.2005.00795.x.

Fox, J. W. 2006. "Using the Price Equation to Partition the Effects of Biodiversity Loss on Ecosystem Function." *Ecology* 87 (11): 2687–2696. https://doi.org/10.1890/0012-9658(2006)87%5B2687:UTPETP%5D2.0.CO;2.

Frank, S. A. 1997a. "The Price Equation, Fisher's Fundamental Theorem, Kin Selection, and Causal Analysis." *Evolution* 51 (6): 1712–1729. https://doi.org/10.2307/2410995.

Frank, S. A. 1997b. "The Design of Adaptive Systems: Optimal Parameters for Variation and Selection in Learning and Development." *Journal of Theoretical Biology* 184 (1): 31–39. https://doi.org/10.1006/jtbi.1996.0241.

Frank, S. A. 2012. "Natural Selection. IV. The Price Equation." *Journal of Evolutionary Biology* 25 (6): 1002–1019. https://doi.org/10.1111/j.1420-9101.2012.02498.x.

Fridley, J. D. 2001. "The Influence of Species Diversity on Ecosystem Productivity: How, Where, and Why?" *Oikos* 93 (3): 514–526. https://doi.org/10.1034/j.1600-0706.2001.930318.x.

Funk, J. L., J. E. Larson, G. M. Ames, B. J. Butterfield, J. Cavender-Bares, J. Firn, D. C. Laughlin, A. E. Sutton-Grier, L. Williams, and J. Wright. 2017. "Revisiting the Holy Grail: Using Plant Functional Traits to Understand Ecological Processes." *Biological Reviews* 92 (2): 1156–1173. https://doi.org/10.1111/brv.12275.

Grace, J. B., T. M. Anderson, E. W. Seabloom, E. T. Borer, P. B. Adler, W. S. Harpole, Y. Hautier, et al. 2016. "Integrative Modeling Reveals Mechanisms Linking Productivity and Plant Species Richness." *Nature* 529: 390–393. https://doi.org/10.1038/nature16524.

Gueyffier, F., and M. Cucherat. 2019. "The Limitations of Observation Studies for Decision Making Regarding Drugs Efficacy and Safety." *Therapie* 74: 181–185. https://doi.org/10.1016/j.therap.2018.11.001.

Huneman, P. 2010. "Topological Explanations and Robustness in Biological Sciences." *Synthese* 177: 213–245. https://doi.org/10.1007/s11229-010-9842-z.

Huston, M. A. 1997. "Hidden Treatments in Ecological Experiments: Re-evaluating the Ecosystem Function of Biodiversity." *Oecologia* 110: 449–460. https://doi.org/10.1007/s004420050180.

Huston, M. A., L. W. Aarssen, M. P. Austin, B. S. Cade, J. D. Fridley, E. Garnier, J. P. Grime, et al. 2000. "No Consistent Effect of Plant Diversity on Productivity." *Science* 289 (5483): 1255a. https://doi.org/10.1126/science.289.5483.1255a.

Huston, M. A., and A. C. McBride. 2002. "Evaluating the Relative Strengths of Biotic Versus Abiotic Controls on Ecosystem Processes." In *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives*, edited by M. Loreau, S. Naeem, and P. Inchausti, 47–60. Oxford: Oxford University Press, Oxford.

Isbell, F., P. B. Reich, D. Tilman, S. E. Hobbie, S. Polasky, and S. Binder. 2013. "Nutrient Enrichment, Biodiversity Loss, and Consequent Declines in Ecosystem Productivity." *Proceedings of the National Academy of Sciences USA* 110 (29): 11911–11916. https://doi.org/10.1073/pnas.1310880110.

Isbell. F., J. Cowles, L. E. Dee, M. Loreau, P. B. Reich, A. Gonzalez, A. Hector, and B. Schmid. 2018. "Quantifying Effects of Biodiversity on Ecosystem Functioning Across Times and Places." *Ecology Letters* 21 (6): 763–778. https://doi.org/10.1111/ele.12928.

Khullar, D., and B. J. Anupam. 2021. "'Natural Experiments' in Health Care Research." *JAMA Health Forum* 2 (6): e210290. https://doi.org/10.1001/jamahealthforum.2021.0290.

Laureto, L. M. O., M. V. Cianciaruso, and D. S. M. Samia. 2015. "Functional Diversity: An Overview of Its History and Applicability." *Natureza & Conservação* 13 (2): 112–116. https://doi.org/10.1016/j.ncon.2015.11.001.

Lefcheck, J. S., and J. E. Duffy. 2015. "Multitrophic Functional Diversity Predicts Ecosystem Functioning in Experimental Assemblages of Estuarine Consumers." *Ecology* 96 (11): 2973–2983. https://doi.org/10.1890/14-1977.1.

Loreau, M., Hector, A. 2001. "Partitioning Selection and Complementarity in Biodiversity Experiments." *Nature* 412: 72–76. https://doi.org/10.1038/35083573.

Maier, D. S. 2018. "Should Biodiversity and Nature Have to Earn Their Keep? What It Really Means to Bring Environmental Goods into the Marketplace." *Ambio* 47 (4): 477-492. https://doi.org/10.1007/s13280-017-0996-5.

Montoya, J. M., S. L. Pimm, and R. V. Solé. "Ecological Networks and Their Fragility." 2006. *Nature* 442: 259–264. https://doi.org/doi:10.1038/nature04927.

Mouchet, M. A., F. Guilhaumon, S. Villéger, N. W. H. Mason, J-A. Tomasini, and D. Mouillot. 2008. "Towards a Consensus for Calculating Dendrogram-Based Functional Diversity Indices." *Oikos* 117 (5): 794–800. https://doi.org/10.1111/j.0030-1299.2008.16594.x.

Naeem, S., J. H. Lawton, L. J. Thompson, S. P. Lawler, and R. M. Woodfin. 1995. "Biotic Diversity and Ecosystem Processes: Using the Ecotron to Study a Complex Relationship." *Endeavor* 19 (2): 58–63. https://doi.org/10.1016/0160-9327(95)93585-3.

Naeem, S., L. J. Thompson, S. P. Lawler, J. H. Lawton, and R. M. Woodfin. 1994. "Declining Biodiversity Can Alter the Performance of Ecosystems." *Nature* 368: 734–737. https://doi.org/doi:10.1038/368734a0.

Naeem, S., K. Hakansson, J. H. Lawton, M. J. Crawley, and L. J. Thomson. 1996. "Biodiversity and Plant Productivity in a Model Assemblage of Plant Species." *Oikos* 76 (2): 259–264. https://doi.org/10.2307/3546198.

Petchey, O. L., and K. J. Gaston. 2006. "Functional Diversity: Back to Basics and Looking Forward." *Ecology Letters* 9 (6): 741–758. https://doi.org/10.1111/j.1461-0248.2006.00924.x.

Price, G. R. 1970. "Selection and Covariance." *Nature* 227: 520–521. https://doi.org/doi:10.1038/227520a0.

Price, G. R. 1972. "Extension of Covariance Selection Mathematics." *Annals of Human Genetics* 35 (4): 485–490. https://doi.org/10.1111/j.1469-1809.1957.tb01874.x.

Price, G. R. 1995. "The Nature of Selection." *Journal of Theoretical Biology* 175 (3): 389–396. https://doi.org/10.1006/jtbi.1995.0149.

Rao, C. R. 1982. "Diversity and Similarity Coefficients: A Unified Approach." *Theoretical Population Biology* 21 (1): 24–43. https://doi.org/10.1016/0040-5809(82)90004-1.

Ricotta, C. 2007. "A Semantic Taxonomy for Diversity Measures." *Acta Biotheoretica* 55: 23–33. https://doi.org/10.1007/s10441-007-9008-7.

Roscher, C., J. Schumacher, J. Baade, W. Wilcke, G. Gleixner, W. W. Weisser, B. Schmid, and E-D. Schulze. 2004. "The Role of Biodiversity for Element Cycling and Trophic Interactions: An Experimental Approach in a Grassland Community." *Basic and Applied Ecology* 5 (2):107–121. https://doi.org/10.1078/1439-1791-00216.

Roscher, C., J. Schumacher, M. Gubsch, A. Lipowsky, A. Weigelt, N. Buchmann, B. Schmid, and E-D. Schulze. 2012. "Using Plant Functional Traits to Explain Diversity–Productivity Relationships." *PLoS ONE* 7: e36760. https://doi.org/10.1371/journal.pone.0036760.

Roscher, C., V. M. Temperton, M. Scherer-Lorenzen, M. Schmitz, J. Schumacher, B. Schmid, B., N. Buchmann, W. W. Weisser, and E-D. Schulze. 2005. "Overyielding in Experimental Grassland Communities—Irrespective of Species Pool or Spatial Scale." *Ecology Letters* 8 (4): 419–429. https:/doi.org/10.1111/j.1461-0248.2005.00736.x.

Schittko, C., M. Hawa, and M. Wurst. 2014. "Using a Multi-Trait Approach to Manipulate Plant Functional Diversity in a Biodiversity-Ecosystem Function Experiment." *PLoS ONE* 9: e99065. https://doi.org/10.1371/journal.pone.0099065.

Sebastião, H., and C. E. V. Grelle. 2009. "Taxon Surrogates Among Amazonian Mammals: Can Total Species Richness Be Predicted by Single Orders?" *Ecological Indicators* 9 (1): 160–166. https://doi.org/10.1016/j.ecolind.2008.03.002.

Selmants, P. C., E. S. Zavaleta, J. R. Pasari, and D. L. Hernandez. 2012. "Realistic Plant Species Losses Reduce Invasion Resistance in a California Serpentine Grassland." *Journal of Ecology* 100 (3): 723–731. https://doi.org/10.1111/j.1365-2745.2011.01949.x.

Solan, M., B. J. Cardinale, A. L. Downing, K. A. M. Engelhardt, J. L. Ruesink, and D. S. Srivastava. 2004. "Extinction and Ecosystem Function in the Marine Benthos." *Science* 396 (5699): 1177–1180. https://doi.org/10.1126/science.1103960.

Song, Y., P. Wang, G. Li, and D. Zhou. "Relationships Between Functional Diversity and Ecosystem Functioning: A Review." *Acta Ecologica Sinica* 34 (2): 85–91. https://doi.org/10.1016/j.chnaes.2014.01.001.

Spehn, E. M., A. Hector, J. Joshi, M. Scherer-Lorenzen, B. Schmid, E. Bazeley-White, C. Beierkuhnlein, et al. 2005. "Ecosystem Effects of Biodiversity Manipulations in European Grasslands." *Ecological Monographs* 75 (1): 37–63. https://doi.org/10.1890/03-4101.

Srivastava, D. S., and M. Vellend. 2005. "Biodiversity-Ecosystem Function Research: Is It Relevant to Conservation?" *Annual Review of Ecology, Evolution, and Systematics* 36: 267–94. https://doi.org/10.1146/annurev.ecolsys.36.102003.152636.

Tilman, D. 1998. "Diversity-Stability Relationships: Statistical Inevitability or Ecological Consequence?" *American Naturalist* 151 (3): 277–282. https://doi.org/10.1086/286118.

Tilman, D., P. B. Reich, and F. Isbell. 2012. "Biodiversity Impacts Ecosystem Productivity as Much as Resources, Disturbance, or Herbivory." *Proceedings of the National Academy of Sciences USA* 109 (26): 10394–10397. https://doi.org/10.1073/pnas.1208240109.

Tilman, D., F. Isbell, and J. M. Cowles. 2014. "Biodiversity and Ecosystem Functioning." *Annual Review of Ecology, Evolution, and Systematics* 45: 471–93. https://doi.org/10.1146/annurev-ecolsys-120213-091917.

van Veelen, M. 2005. "On the Use of the Price Equation." *Journal of Theoretical Biology* 237 (4): 412–426. https://doi.org/10.1016/j.jtbi.2005.04.026.

van Veelen, M., J. García, M. W. Sabelis, and M. Egas, M. 2012. "Group Selection and Inclusive Fitness Are Not Equivalent; The Price Equation vs. Models and Statistics." *Journal of Theoretical Biology* 299: 64–80. https://doi.org/10.1016/j.jtbi.2011.07.025.

Vandermeer, J., D. Lawrence, A. Symstad, and S. E. Hobbie. 2002. "Effect of Biodiversity on Ecosystem Functioning in Managed Ecosystems." In *Biodiversity and Ecosystem Functioning: Synthesis and Perspectives*, edited by M. Loreau, S. Naeem, and P. Inchausti, 221–233. Oxford: Oxford University Press.

Villéger, S., N. W. H. Mason, and D. Mouillot. 2008. "New Multidimensional Functional Diversity Indices for a Multifaceted Framework in Functional Ecology." *Ecology*, 89 (8): 2290–2301. https://doi.org/10.1890/07-1206.1.

Wardle, D. A. 2016. "Do Experiments Exploring Plant Diversity-Ecosystem Functioning Relationships Inform How Biodiversity Impacts Natural Ecosystems?" *Journal of Vegetation Science* 27: 646–653. https://doi.org/10.1111/jvs.12399.

Weiser, W. W., C. Roscher, S. T. Meyer, A. Ebeling, G. Luoa, E. Allane, H. Beßler, et al. 2017. "Biodiversity Effects on Ecosystem Functioning in a 15-Year Grassland Experiment: Patterns, Mechanisms, and Open Questions." *Basic and Applied Ecology* 23: 1–73. https://doi.org/10.1016/j.baae.2017.06.002.

Woodward, J. 2003. *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press, New York. https://doi.org/10.1093/0195155270.001.0001.