



**University of  
Zurich**<sup>UZH</sup>

**Zurich Open Repository and  
Archive**

University of Zurich  
University Library  
Strickhofstrasse 39  
CH-8057 Zurich  
[www.zora.uzh.ch](http://www.zora.uzh.ch)

---

Year: 2022

---

**A graphical causal model for resolving species identity effects and  
biodiversity–ecosystem function correlations: comment**

Grace, James B ; Loreau, Michel ; Schmid, Bernhard

DOI: <https://doi.org/10.1002/ecy.3378>

Posted at the Zurich Open Repository and Archive, University of Zurich

ZORA URL: <https://doi.org/10.5167/uzh-211429>

Journal Article

Published Version

Originally published at:

Grace, James B; Loreau, Michel; Schmid, Bernhard (2022). A graphical causal model for resolving species identity effects and biodiversity–ecosystem function correlations: comment. *Ecology*, 103(2):e03378.

DOI: <https://doi.org/10.1002/ecy.3378>

# Comments

*Ecology*, 0(0), 2021, e03378

© 2021 *Ecological Society of America*. This article has been contributed to by US Government employees and their work is in the public domain in the USA.

## A graphical causal model for resolving species identity effects and biodiversity–ecosystem function correlations: comment

JAMES B. GRACE <sup>1,4</sup> MICHEL LOREAU,<sup>2</sup> AND BERNHARD SCHMID<sup>3</sup>

*Citation:* Grace, J. B., M. Loreau, and B. Schmid. 2021. A graphical causal model for resolving species identity effects and biodiversity–ecosystem function correlations: comment. *Ecology* 00(00):e03378. 10.1002/ecy.3378

In a recent paper, Schoolmaster, Zirbel, and Cronin (SZC) (2020) claim “Formal causal analysis show[s] that biodiversity–ecosystem function (BEF) correlations are non-causal associations.” If this conclusion is accepted as true, it suggests a reconsideration of much of our current understanding of how biodiversity relates to the functioning of ecosystems. On the surface, it is easy to spot clear signs of something problematic with SZC’s presentation. They claim, for example, that (1) species richness is incapable of having a causal effect on ecosystem functioning on theoretical grounds, and (2) that trait diversity cannot be causally influenced by species diversity. These remarkable claims are counter to existing thought and evidence. We point to logical errors that lead them to a misapply causal analysis and produce erroneous conclusions.

### SZC’S ARGUMENT

We start with an overview of the four key elements of their logical arguments that lead them to reject existing theory (the Standard Model; e.g., Loreau 2000, Loreau et al. 2001) and to propose an alternative model (SZC’s

Model) for understanding BEF. We follow this with our critique of SZC’s logical arguments and of their use of graphical causal modeling. We offer alternative causal diagrams that omit inappropriate components and use these to provide support for the Standard Model. Fig. 1 provides a storyboard for our examination of SZC.

### *Element 1: The Standard Model*

SZC summarize the current biodiversity–ecosystem function (BEF) literature via a causal diagram for the “Standard Model,”  $B \rightarrow Q \rightarrow F$ , where  $B$  is biodiversity,  $Q$  refers to functional trait diversity, and  $F$  represents ecosystem functions (Fig. 1A). In their words, “if environmental variation [ $E$ ] is accounted for, then variation in species diversity causes variation in functional traits and, subsequently, ecosystem function.”

### *Element 2: SZC’s first criticism of the Standard Model*

SZC state that to evaluate the causal claims in Fig. 1A, “it is necessary to recognize that . . . species diversity and functional trait diversity are calculated. . . .” They go on to present a diagram (Fig. 1B) that is meant to reflect the process they might use to calculate species biodiversity ( $B$ ) and trait diversity ( $Q$ ) from the vectors of species abundances ( $C$ , indicated by boxes  $s_1 \dots s_n$ ) and species traits ( $T$ , indicated by boxes  $t_1 \dots t_n$ ). They claim that by representing the computation process in a diagram, it exposes a logical error in the Standard Model.

### *Element 3: SZC’s second criticism of the Standard Model*

SZC go on to say the hierarchical representation of the Standard Model (Fig. 1B) is itself fatally flawed. They state that the diagram, “makes the claim that in order to calculate trait diversity, some measure of species diversity can be first calculated and then subsequently used to calculate trait diversity, or alternatively, that measures of trait diversity can be rewritten in terms of  $B$ . This claim is simply not true for any definition of trait diversity currently used.” Here they are claiming that trait diversity ( $Q$ ) cannot be computed from species diversity ( $B$ ) and, therefore, species diversity cannot possibly have a causal effect on trait diversity and ecosystem function.

### *Element 4: SZC propose an alternative to the Standard Model, which we refer to as SZC’s Model*

Fig. 1C represents SZC’s diagram for an alternative to the Standard Model. They argue that, “In Fig. 1B, species diversity is a cause of ecosystem function via causal chain, whereas in Fig. 1C, species diversity is correlated with ecosystem function due to their joint

Manuscript received 23 September 2020; revised 17 February 2021; accepted 16 March 2021. Corresponding Editor: Brian D. Inouye.

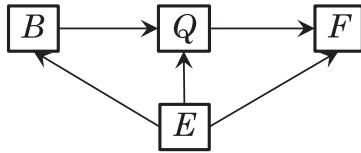
<sup>1</sup>U.S. Geological Survey, Wetland and Aquatic Research Center, 700 Cajundome Boulevard, Lafayette, Louisiana 70506 USA

<sup>2</sup>Theoretical and Experimental Ecology Station, CNRS, 2 route du CNRS, 09200 Moulis, France

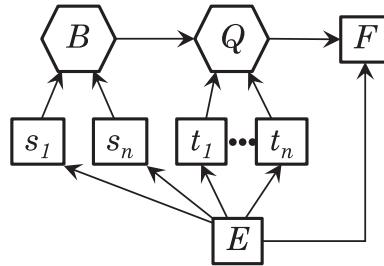
<sup>3</sup>Department of Geography, Remote Sensing Laboratories, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

<sup>4</sup>Corresponding Author. E-mail: gracej@usgs.gov

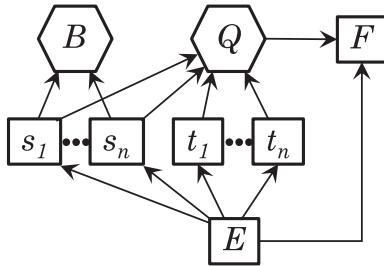
(A) SZC's single-level causal diagram for the Standard Model



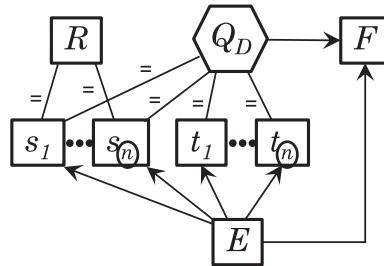
(B) SZC's hierarchical diagram for the Standard Model



(C) SZC's proposed hierarchical diagram

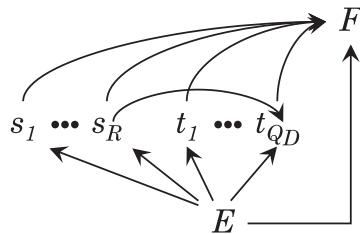


(D) SZC's diagram modified and annotated for discussion



$R$ , richness, = the number of nonzero elements in the vector of species abundances in a sample.

(E) Hierarchical causal diagram for Standard Model



$R$  is number of species in a sample.  
 $Q_D$  is number of traits in a sample.

(F) Single-level causal diagram for BEF experiment, Standard Model

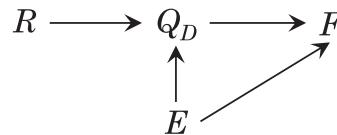


FIG. 1. (A) SZC's representation of a single-level causal diagram for the Standard Model ( $B \rightarrow Q \rightarrow F$ ).  $B$  = a set of possible expressions of species diversity {species richness, Simpson's Index, Shannon's Index},  $F$  = any of a set of possible ecosystem functions,  $Q$  = composite index of a functional trait vector that predicts  $F$  (e.g.,  $F = f(s, t)$ ). (B) SZC's representation of a hierarchical diagram for the Standard Model.  $E$  = environmental gradient,  $n$  = number of species encountered in a study,  $s_i$  = abundance of species  $i$ ,  $t_i$  = traits of species  $i$ . (C) SZC's Model, which removes the directed arrow from  $B$  to  $Q$  in Fig. 1B. (D) Version of SZC's Model that we have modified and annotated by replacing  $B$  with  $R$ , species richness, and  $Q$  with  $Q_D$ , the number of distinct functional traits. We annotate the graph using the = sign to draw attention to arrows that actually represent noncausal, reversible computations as if they were irreversible causal effects. We further remove the heads of those arrows to correct the misrepresentation. We also circle the subscript  $ns$  SZC use to index the total numbers of species and traits across all samples to draw attention to the fact that those represent study-wide information rather than sample plot-level information. (E) Our representation of a valid hierarchical causal diagram for the Standard Model. Here we (a) remove computational elements, including both the computational arrows and computed quantities, (b) index species abundances ( $s_i$ ) to show that the number of species in a sample =  $R$ ; (c) index traits ( $t_i$ ) to show that the number of traits in a sample =  $Q_D$ ; (d) include a causal arrow to show that the number of traits in a sample depends on the number of species; (e) remove variable outlines to conform to the standard format for causal diagrams. (F) Single-level causal diagram for a BEF experiment assuming the Standard Model. We remove the arrow  $E \rightarrow R$ , reflecting the fact that  $R$  is exogenous when randomly assigned. Note hierarchical diagrams (B–E) do not reflect species interactions, the main cause of biodiversity effects in experiments.

dependence on the abundances of the species in the community.” Based on this, they conclude that Fig. 1C is a valid diagram for causal analysis and Fig. 1A, B are not valid causal diagrams.

#### A BRIEF INTRODUCTION TO CAUSAL DIAGRAMS

In order to aid our discussion of causal diagrams, we give a brief statement of their formal properties here. A more complete description of causal diagrams as well as causal calculus (the procedures used to assess causal logic) can be found in Pearl (2009) and Elwert (2013). Causal diagrams are based on directed acyclic graphs (DAGs) and represent hypotheses about the data-generating processes related to a problem of interest. Causal diagrams comprise a set of nodes representing variables, and directed arrows connecting those nodes representing hypothesized causal processes (e.g.,  $X \rightarrow Y$ ). The “directed” requirement for causal diagrams means that all one-headed arrows represent unidirectional, irreversible cause–effect relationships. Thus, we can test the logic represented by any given arrow in Fig. 1 to see if it meets the irreversibility requirement by asking the question, “If we were to manipulate the variable at the tip of the arrow (e.g.,  $Y$ ), would there be a potential for a response by the variable at the base of the arrow ( $X$ )?” For true causal arrows, the answer is “no” for the most fundamental reason that you cannot change conditions in the past (parent or ancestor variables) by making changes in the present. SZC include in their diagrams arrows that represent reversible relationships (noncausal calculations), but argue (in their Appendix S1) that arrows in causal diagrams do not need to represent irreversible relationships. We critique their diagram and their defense of it in detail in our Appendix S1. To be clear, we wish to emphasize that causal diagrams, when used properly, permit researchers to state their assumptions about causal structures explicitly and transparently. This can be of great benefit by allowing other researchers to examine those causal assumptions and subsequent interpretations critically, as is done in this comment.

#### OUR SYNOPTIC REACTIONS TO THE FOUR ELEMENTS OF SZC’S ARGUMENT

##### *Reaction 1: The Standard Model*

We do not disagree with SZC’s single-level representation (Fig. 1A) of a causal diagram for the Standard Model.

##### *Reaction 2: SZC’s first criticism of the Standard Model*

SZC argue that a hierarchical representation (Fig. 1B) is required to evaluate the causal claims made by the Standard Model properly. They claim this is so because it

makes explicit that diversity is computed from the non-zero elements in the vector of species abundances. We do not agree with their argument. Rather, we see this as their first fundamental error, because it introduces noncausal relationships into what is supposed to be a causal diagram. We address this point in detail in Appendix S1, but essentially their decision to include species vectors and community-level properties in the diagram in the way they do, and to treat species diversity as if it is not an inherent property of the community, invalidates their diagrams as suitable for causal calculus. The reasons for this become clear when we discuss Fig. 1E, our alternative representation of a valid causal diagram.

##### *Reaction 3: SZC’s second criticism of the Standard Model*

SZC argue that the hierarchical representation of the Standard Model (Fig. 1B) is invalid because it implies, “that measures of trait diversity ( $Q$ ) can be rewritten in terms of  $B$  (species diversity), which is not true.” We disagree with their statement for reasons given in the section entitled “Critique 1” and directly dispute it in Appendix S2.

##### *Reaction 4: SZC’s Model*

Fig. 1C represents SZC’s attempt to address their second criticism of the Standard Model. In this diagram, they represent the relationship among key quantities as  $B \leftarrow \{C, T\} \rightarrow Q \rightarrow F$ , where  $C$  and  $T$  are the vectors of species abundances  $\{s_{1..n}\}$  and traits  $\{t_{1..n}\}$  across the entire study. This representation becomes the basis for the claims made in their paper, particularly for the claim that species diversity ( $B$ ) is not a cause of either trait diversity ( $Q$ ) or ecosystem function ( $F$ ). It is also the basis for their claims that BEF experiments, “do not actually manipulate biodiversity.” These claims are direct consequences of their diagram, because there is no direct ( $B \rightarrow Q$ ) or indirect ( $B \rightarrow Q \rightarrow F$ ) path from  $B$  to  $F$ . SZC spend a significant portion of their paper illustrating the use of causal calculus to express the empirical expectations that arise from their assumptions. This logic is carried over into their presentation of an empirical example as well as their simulation studies. We argue that their diagram (1) is not a valid causal diagram, (2) is not appropriate for their application of causal calculus, (3) does not justify their theoretical and empirical claims against the Standard Model, and (4) does not support their claim that BEF experiments are incapable of showing that diversity has a causal effect on function.

#### CRITIQUE 1: SZC’S MODEL IS NOT A VALID CAUSAL DIAGRAM AND IS NOT SUITABLE FOR THE ANALYSES THEY PERFORM

We use Fig. 1D to facilitate discussion of the problems we see with SZC’s Model. In that figure, we

annotate and relabel SZC's diagram to simplify discussion, draw attention to SZC's misuse of directed arrows for noncausal computations, and highlight other problems with their diagram. First, we replace the label  $B$ , which SZC use to refer to a suite of possible measures of species diversity, with  $R$ , which refers to species richness, the most elemental measure from their set and the one widely used in BEF experiments. We also replace  $Q$ , which SZC define as "functional trait composites" with  $Q_D$ , which refers to trait diversity, specifically the number of unique traits in a sample. Second, we use the "=" operator to indicate noncausal computations that are misrepresented by SZC as causal operations through their use of the " $\rightarrow$ " symbol and remove the tips of those arrows to correct the notation (see more on this in Appendix S1).

There are multiple problems with SZC's diagram. The most serious is the inclusion of noncausal hierarchical relations as if these are cause-effect relationships. Hierarchical relationships are embedded in their model by including both community-level summary properties ( $B$  and  $Q$ ) and underlying species-level information ( $C$  and  $T$ ). As they have drawn it in Fig. 1C, the elements of the community ( $C$  and  $T$ ) are represented as causes of the properties of the community ( $B$  and  $Q$ ). More specifically, SZC point unidirectional arrows, which represent causal effects by convention, from the community elements to the community properties (Fig 1C). However, community elements are not causes of community properties; rather, the two levels of information are simultaneous descriptions at different levels of detail. To illustrate this critical point further, imagine we have a warehouse filled with items and hold in our hands a ledger that accurately and instantaneously reflects the total number of items and the number of distinct types of items in that warehouse. If there are 10 types of items in the warehouse, the information about how many types of items there are exists simultaneously in the warehouse and on the ledger. The warehouse does not cause there to be 10 types of items, rather, the ledger simply reflects how many types there are, which is a quantitative property of the assemblage of items in the warehouse. We show specifically how their diagram corresponds to this analogy in Appendix S2: Fig. S1. SZC actually recognize there is a problem with their representation when they say, "Because biodiversity is a calculated variable, setting its value independently of the species abundances is impossible." What we would say instead is, "Because biodiversity is represented as a noncausal calculation in the diagram, Fig. 1C is not a valid causal diagram." Despite their recognition that the relationship between  $C$  and  $B$  in their model (shown as  $C \rightarrow B$ ) fails to represent a true causal effect, they go on to perform causal analysis using conditional independence tests as if all the arrows were true causal arrows.

The key subset of relationships in SZC's diagram (Fig. 1C) we wish to draw attention to is  $B \leftarrow S \rightarrow Q \rightarrow F$ .

This sequence presents  $S$  as a common cause for both  $B$  and  $F$  (through  $Q$ ). If this were a true and complete representation, we would describe the associations between  $B$  and  $Q$  and between  $B$  and  $F$  as inherently spurious because of a common cause. As SZC put it (referring here to our figure numbers), "In Fig. [1B (Standard Model)], species diversity is a cause of ecosystem function via causal chain, whereas in Fig. [1C (SZC Model)], species diversity is correlated with ecosystem function due to their joint dependence on the abundances of the species in the community." This interpretation lies at the heart of SZC's entire paper and leads to their claim that the relationship between  $B$  and  $F$  is noncausal. However, their logic hinges on treating the directional arrows from  $S$  to  $B$  as irreversible causal effects in their conditional independence tests. Their conclusions that (1) species richness has no causal effect on ecosystem functioning and (2) trait diversity is not causally influenced by species diversity, emerge from this logic.

Aside from the misuse of arrows, there are other features of their diagram that obscure symptoms of the flaws in their approach. A critical point SZC fail to emphasize is that the number of nonzero elements in the vector  $C$  that are found in a sample is, in fact, species richness ( $R$ ). SZC have chosen to represent the vectors for all species found across all samples in their diagram (i.e.,  $n$  in their diagram is gamma diversity). This choice obscures the fact that  $R$  is an implicit property of samples/plots. When that is recognized, we can see that richness is actually represented twice in their diagram, once as a distinct variable  $R$  at the community level and again as the number of nonzero abundances within each sample plot at the species level (this is shown explicitly in Appendix S1: Fig. S2). It is this duplication of information that leads SZC to claim that  $R$  is unnecessary for computations of causal effects once we have used the information contained in  $C$ .

In Fig. 1E, we remove the inappropriate noncausal calculations from SZC's diagram. We further modify the diagram to conform to the standard format by removing box outlines for the variables, a practice designed to avoid confusing causal diagrams with structural models (see Grace and Irvine 2020). We make other changes as well. We index the vector of species abundances so as to show that the number of species in a sample =  $R$  and we index the vector of traits so as to show that the number of traits =  $Q_D$ . As a result of this reconfiguration, we avoid representing richness twice in the diagram, which is a key hidden flaw in their diagram. Another important change is to include a directed arrow from  $R$  to  $Q_D$  to show that the number of traits in a sample depends on the number of species in a sample. The result of these various changes is to make clear that even in a hierarchical diagram, it is possible to provide a valid representation of the hypothesis  $R \rightarrow Q_D \rightarrow F$  (the Standard Model). Appendix S2 provides simulation results that support

the representation in Fig. 1E and show how an effect of diversity emerges through the filling of niche space as one sums the contributions of species with distinct traits. It is possible and even appropriate to develop a corresponding single-level diagram (Fig. 1F), and in fact the entire literature on causal diagrams is based on single-level representations. In summary, our causal diagrams provide support for the Standard Model. Further, our critique of SZC's diagram leads us to reject their diagram because of its double inclusion of species diversity, both as an attribute of the community sample and as if it were an additional piece of information.

**CRITIQUE 2: SZC DO NOT SHOW THAT BIODIVERSITY HAS NO EFFECT ON ECOSYSTEM FUNCTION. RATHER, THEY BUILD THE ASSUMPTION OF NO EFFECT INTO THEIR GRAPH**

SZC subject their diagram to causal analysis using formal procedures. The results from those analyses are completely dependent on their diagram (Fig. 1C), as SZC freely admit. In their diagram they provide no pathway that points directly or indirectly from  $B$  to  $F$ . Therefore, their diagram encodes the *assumption* that biodiversity is not a cause of community structure. The tests of conditional independence presented by SZC follow from that assumption, but do not test it. SZC indicate that they recognize that their conclusions follow directly from their assumptions when they say, "Using the DAG in Fig. [1C], we will derive the statistical model required to measure a biodiversity effect . . . Fig. [1C] makes clear that the value of this effect is zero [as] there are no causal arrows emitted from the biodiversity node." Thus, their analyses provide no test of their model. This same criticism applies to their simulation study and analysis of field data. For this reason, we do not consider further those elements of SZC's presentation.

**CRITIQUE 3: BIODIVERSITY EXPERIMENTS ILLUSTRATE CONDITIONAL CAUSATION, NOT NONCAUSAL ASSOCIATIONS—STRUCTURAL INVARIANCE IS NOT A REQUIREMENT**

SZC state, "Because the effect of the manipulation on ecosystem function will depend on trait distribution of the set of species present and not the number of species recognized under current taxonomy (i.e., it is not structurally invariant; Pearl 2009), attempting to estimate a 'biodiversity effect' on particular ecosystem functions will never find consistent results and will not be improved by multisite studies, multiscale studies or meta-analysis..." It is not true that structural invariance is a requirement for the estimation of an average causal effect (ACE) of  $R$  on  $F$ . An invariant causal mechanism is one that expresses itself to the same degree for all members of a population and under all conditions. When the ACE for a subpopulation depends on covariates of any sort (which is typical in biological systems),

it is a case of conditional causation (VanderWeele 2015:57). In this particular example, there are expected to be variations in  $F$  among replicate mixtures for any given  $R$ . This in no way invalidates the estimation of a valid ACE for the biodiversity effect. More to the critical point, conditional causation is still causation and would not justify declaring  $R \rightarrow F$  a spurious relationship. For practical reasons, as long as we lack important measures for  $Q$  in the sequence  $R \rightarrow Q \rightarrow F$ , we will have to include paths for  $R \rightarrow F$  for causal modeling and prediction (Appendix S2).

**CRITIQUE 4: SZC'S "ACTUAL-CAUSE" ARGUMENTS ARE FALLACIOUS**

Although SZC's arguments arise from a literal interpretation of their diagram, the authors rely heavily on verbal arguments in their paper. Most notably, they rely on so-called "actual-cause" arguments. First, they say that species richness is not an actual cause of community function; rather, it is the constituent species that are the actual cause. Second, they say BEF experiments do not actually involve manipulations of diversity, but instead manipulations of the constituent species. More specifically, they state, "... it is impossible to manipulate biodiversity directly in an experiment. So-called biodiversity manipulation experiments are *actually* manipulations of community structure." [italics added for emphasis]

Actual-cause arguments are common in human argumentation. They represent a family of fallacious arguments known as "half-truths." In this case, SZC imply that if species composition contributes to variations in  $F$ , then somehow species richness is prohibited from contributing to variations in  $F$ . Expressed more explicitly, they imply "The number of species in a community does not affect function, it is the species themselves." The reason this statement is a half-truth is because the second half of the statement is true, but the first half of the sentence is false. The reason half-truth fallacies are effective in argumentation is that the human mind tends to focus on the part of a sentence that follows a comma and dismiss the part that comes before.

A related type of half-truth found in common discussions of causes is when variables exist in a causal chain and someone argues one of the variables in that chain is the "actual" cause. In truth, all variables in a directed causal chain are causes, it is just that some are more distal to the terminal end of the chain and others are more proximal. So, if manipulating richness produces a change in the number of traits, which subsequently leads to changes in function, then richness is one of the causes in the chain of events. We argue that a more appropriate conclusion to draw regarding causal influences on ecosystem function is, "The evidence suggests that species diversity influences community function, *and* that the characteristics of the individual species in a

community contribute to variations in function.” This represents the form of a “whole-truth” argument.

#### SUMMARY

Schoolmaster et al. (2020) claim that species diversity cannot have a causal influence on ecosystem functions in natural systems or in experiments because the relationship between the two is noncausal (specifically, spurious). They reach this conclusion despite demonstrations that ecosystem functions often respond to manipulations of species richness in experimental studies, as well as other evidence. SZC further claim that they are able to prove their claim using the principles and tools of causal analysis. We draw attention to critical errors in their diagram and analyses that invalidate their conclusions. We provide background on this methodology for the reader and offer alternative diagrams with valid structures, which support the conventional understanding that species diversity can influence ecosystem function indirectly through influences on trait diversity.

#### ACKNOWLEDGMENTS

We thank Brian Inouye, R. Taylor, G. R. Guntenspergen, two anonymous reviewers, and SZC for helpful reviews. JG was supported by the USGS Climate R&D and Ecosystems

Programs, ML by the TULIP Laboratory of Excellence (ANR-10-LABX-41), and BS by the University Research Priority Program for Global Change and Biodiversity of the University of Zurich. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

#### LITERATURE CITED

- Elwert, F. 2013. Graphical causal models. Pages 113–132 in G. Firebaugh, C. Warner, and M. Massoglia, editors. *Handbook of causal analysis for social research*. First edition. Springer, Dordrecht, The Netherlands.
- Grace, J. B., and K. M. Irvine. 2020. Scientist’s guide to the development of explanatory statistical models using causal analysis principles. *Ecology* 101:e02962.
- Loreau, M. 2000. Biodiversity and ecosystem functioning: recent theoretical advances. *Oikos* 91:3–17.
- Loreau, M., et al. 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* 294:804–808.
- Pearl, J. 2009. *Causality*. Second edition. Cambridge University Press, Cambridge, UK.
- Schoolmaster, D. R., Jr., C. R. Zirbel, and J. P. Cronin. 2020. A graphical causal model for resolving species identity effects and biodiversity–ecosystem function correlations. *Ecology* 101:e03070.
- VanderWeele, T. 2015. *Explanation in causal inference: methods for mediation and interaction*. Oxford University Press, Oxford, UK.

#### SUPPORTING INFORMATION

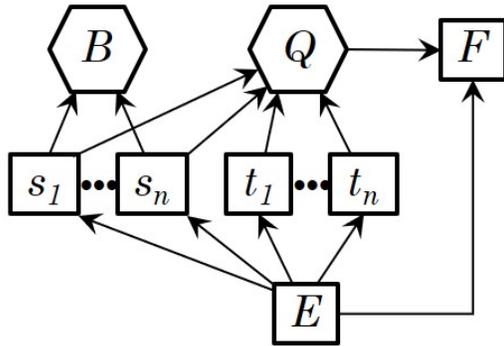
Additional supporting information may be found in the online version of this article at <http://onlinelibrary.wiley.com/doi/10.1002/ecy.3378/supinfo>

**Supporting Information.** Grace, J.B., M. Loreau, and B. Schmid. 2021. A graphical causal model for resolving species identity effects and biodiversity–ecosystem function correlations: comment. *Ecology*.

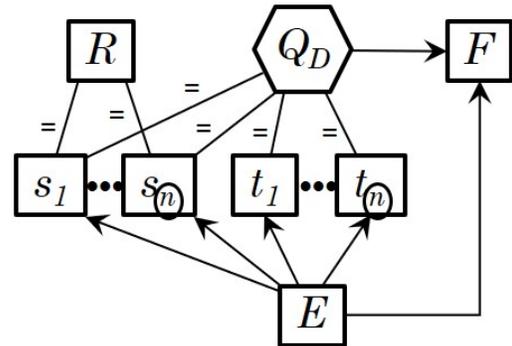
**Appendix S1. Critique of SZC’s claim that their diagrams are proper ‘causal’ diagrams.**

As noted in the main text, SZC present a diagram representing the hypothesis  $B \leftarrow C \rightarrow Q \rightarrow F$  as an alternative to the Standard Model  $B \rightarrow Q \rightarrow F$ , where  $B$  refers to biodiversity,  $C$  the vector of species abundances  $\{s_{1..n}\}$  for all species encountered in a study,  $Q$  is functional trait diversity, and  $F$  is ecosystem function. SZC refer to their diagram as a ‘causal diagram’, which has formal requirements if it is to be used for causal analysis, which they do. We dispute that SZC’s diagram (Fig. S1 panel C) is a proper causal diagram because it includes non-causal arrows, as indicated by = signs in panel D of Figure S1. It is also improper for other reasons we address later in the appendix.

(C) SZC's Proposed Hierarchical Diagram



(D) SZC's Diagram Modified and Annotated for Discussion



$R$ , richness, = the number of non-zero elements in the vector of species abundances in a sample.

Figure S1. Reproduction of panels C and D from Figure 1 in main text. Note that the vector of species abundances  $\{s_{1..n}\}$  for all species encountered in a study is referred to in the text as  $C$  and the vector of traits  $\{t_{1..n}\}$  for those species is referred to as  $T$ .

*Non-Causal Computations do not belong in Causal Diagrams*

We need to look no further than SZC’s own presentation to provide us with a basis for recognizing that the arrows they point from  $C$  and  $T$  (the vector of traits  $\{t_{1..n}\}$ ) to  $B$  and  $Q$  fail to meet the requirements for causal arrows. SZC state in the section Graphical Causal Modeling of their paper, “Arrows represent the presence of a direct causal effect of the node at the tail of the arrow on the node at the head of the arrow and thus suggests a temporal ordering.” Further, they state in their paper, “In causal diagrams, as defined here, there is an implicit flow of time because causes must precede effects.”

The passage of time has the consequence of giving causal arrows the property of irreversibility, since events in the present cannot change events in the past. Formally, this is known as asymmetry. Grace and Irvine (2020) provide the following description:

“There are several definitional features of causal diagrams. First, directed arrows imply cause–effect relationships. This means that if we physically change the values of the variable at the base of an arrow, while holding constant all other predictors, it can alter the values of the variable at the tip of the arrow. There are two corollaries that go with this definition of arrows. One is that there is understood to be a finite passage of time between the cause and the response. Thus, the arrows represent movement of information from some time in the past to present. Another corollary is that we can force changes to the quantity at the tip of the arrow without influencing the quantity at the base of the arrow (also referred to as the assumption of asymmetry [irreversibility]). This means that changing present conditions should not change past conditions. Thus, arrows represent implied experimental manipulations and routinely we use thought experiments to analyze the logical relations in causal diagrams. When we can be more rigorous, we use physical experiments to test our assumptions.”

Despite SZC acknowledging that causal arrows imply the passage of time, they include arrows they admit are non-causal. Specifically, SZC recognize that the arrows directed from **C** and **T** to **B** do not have the usual meaning associated with causal arrows when they say, “Because biodiversity is a calculated variable, setting its value independently of the species abundances [**C**] is impossible.” Despite that, they go on to say, “community composition drives ... biodiversity”, implying **C** has a causal effect on **B** through the use of the word “drives”. More importantly, their application of causal analysis treats the arrows from **C** and **T** to **B** as causal. The consequences of that treatment are manifested in their statement, “In [Fig. S1 panel C], species diversity is correlated with ecosystem function due to their joint dependence on the abundances of the species in the community”. This representation of **C** as a common cause confounder for **B** and **Q** unambiguously indicates that they are treating their non-causal arrows as if they are irreversible causal arrows. They do this despite admitting the relationships are non-causal.

Despite the contradictions and inconsistencies created by including non-causal arrows in causal diagrams, SZC argue that including non-causal arrows in causal diagrams is permissible in their Appendix S1. SZC justify the inclusion of non-causal calculations through quotes of Pearl (2016) such as the following: “A variable  $X$  is the direct cause of the variable  $Y$  if  $X$  appears in the function  $f_Y$  that assigns the values of  $Y$ .” We point out that this select text only describes how effects are calculated and does not present the broader context and requirements for these to be calculations of causal effects, which depends on valid causal diagrams (diagrams made up of causal arrows). SZC further state that  $X$  is recognized to be a *parent* of  $Y$  in their description of causal diagrams without acknowledging that the term *parent* denotes ancestry, which in turn, implies the passage of time. In fact, SZC admit that making changes to **B** will instantaneously result in changes in **C**, an obvious violation of the requirement for an implied time step. Despite all these inconsistencies and the improper use of arrows to denote reversible, non-causal calculations, SZC base their paper on causal analysis techniques that assume all arrows in their diagrams represent irreversible causal relationships. This error is fatal to the scientific conclusions they draw.

*SZC's Approach to including Hierarchical Information Hides Sample Characteristics and Leads to a Duplication of Information*

SZC's preoccupation with embedding computational machinery into their diagram leads them to another problematic choice, which is the inclusion of study-wide data vectors in a model of sample plot outcomes. We draw attention to this issue in Figure S1 panel D by placing circles around the subscript  $n$ , which represents the total number of species in the entire study (across all samples). This means that  $B$ ,  $Q$ ,  $F$ , and  $E$  in their diagrams represent sample-level information, while  $C$  and  $T$  vectors represent study-wide information. We would expect a normal causal diagram to include nodes that summarize comparable quantities; SZC's Model does not. This is, to our experience, unprecedented and another problematic element in their diagram.

An important consequence of SZC's representation of species-level information using study-wide vectors is that it hides from view critically important sample-plot information. To clarify the discussion of this issue, in Fig. S1 panel D we replace  $B$ , which represents a heterogeneous collection of diversity measures with  $R$ , the most elemental measure from that collection, species richness (the number of species in a sample). We also replace  $Q$ , which again represents a heterogeneous collection of functional trait composites, including trait diversity, with  $Q_D$ , which is the number of distinct traits in a sample. Now, what would happen if we additionally replaced SZC's representation of study-wide vectors with sample plot vectors? We illustrate the consequences through another modification of their model in Figure S2.

Modification of SZC's Diagram replacing  $n$  with  $R$  and  $Q_D$ .

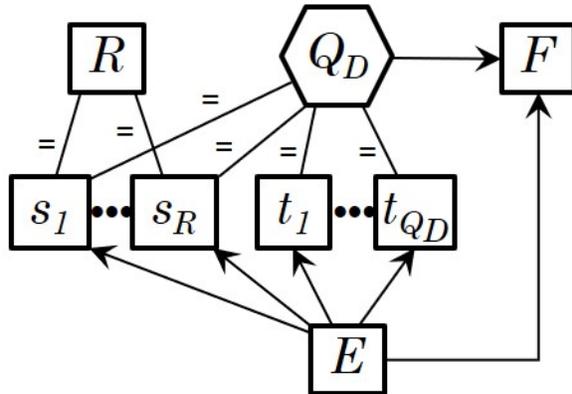


Figure S2. Modification of SZC's diagram replacing study-wide vectors with plot-level vectors.

What Figure S2 shows is that the number of species in a sample ( $R$ ) is an explicit property of the sample-level vector of species abundances - the length of the vector. Also, the number of unique traits in a sample ( $Q_D$ ) is the length of the trait vector for a sample. Here we are deviating from what SZC have done in a number of ways, but most importantly, we are making transparent the fact that  $R$  and  $Q_D$  are explicit sample-level properties. In addition, we avoid the complexity represented in SZC's diagram, as well as the mis-match of what nodes represent.

Based on our further modification of their diagram (Fig. S2), it now becomes apparent that species richness and trait diversity are represented twice, once as the number of species or traits in a sample and again as if it were some separate computation of the same quantity. We believe this clarifies the erroneous claims SZC make about 'species diversity', namely that (a) 'species

richness' is only spuriously correlated with ecosystem function, (b) that trait diversity cannot be causally influenced by 'species richness', and (c) it is impossible to manipulate 'species richness' in BEF experiments. It would seem they are using the term 'species richness' to refer to a redundant copy of the variable whose effects on function have already been taken into account. Stated in another way, SZC appear to be saying that once one accounts for the role of the number of species in a plot, the extra copy of 'species richness' provides no additional critical information for computing trait diversity. This duplication of information is another fatal flaw in their diagram.

### **In Summary**

The diagrams that SZC use to represent their model do not qualify as proper causal diagrams and are not appropriate for the inferences they make. This conclusion is not only supported by our own descriptions of causal diagrams and causal analysis (Grace and Irvine 2020), but also by reference to the requirements for causal diagrams as described by SZC in their introductory section, *Graphical Causal Modeling*.

### **Literature Cited**

- Grace, J. B. and K. M. Irvine. 2020. Scientist's guide to the development of explanatory statistical models using causal analysis principles. *Ecology* 101(4): e02962
- Pearl, J., Glymour, M., and Jewell, N.P. (2016) *Causal inference in statistics*. John Wiley & Sons Ltd, West Sussex, UK.

**Supporting Information.** Grace, J.B., M. Loreau, and B. Schmid. 2021. A graphical causal model for resolving species identity effects and biodiversity–ecosystem function correlations: comment. *Ecology*.

## **Appendix S2. Evaluation of Causal Diagram for Standard Model**

### *Introduction*

The Standard Model describes the hypothesis  $B \rightarrow Q \rightarrow F$ , where  $B$  is biodiversity,  $Q$  refers to functional trait diversity, and  $F$  represents ecosystem functions. In our discussion, we first make things more concrete by replacing  $B$ , which SZC use to refer to a suite of possible measures of species diversity, with  $R$ , which refers to one of the forms of  $B$  they mention – species richness, the most elemental measure. We also replace  $Q$ , which SZC define as “functional trait composites” with  $Q_D$ , which refers to one example of  $Q$  – trait diversity, specifically the number of distinct traits in a species mixture. Thus, we represent the Standard Model in terms of  $R \rightarrow Q_D \rightarrow F$ . In their words, SZC state, “if environmental variation is accounted for, then variation in species diversity causes variation in functional traits and, subsequently, ecosystem function.”

SZC criticize the Standard Model on theoretical grounds, stating that the Standard Model “makes the claim that in order to calculate trait diversity, some measure of species diversity can be first calculated and then subsequently used to calculate trait diversity, or alternatively, that measures of trait diversity can be rewritten in terms of  $B$ . This claim is simply not true for any definition of trait diversity currently used.” We directly confront their claim in this Appendix.

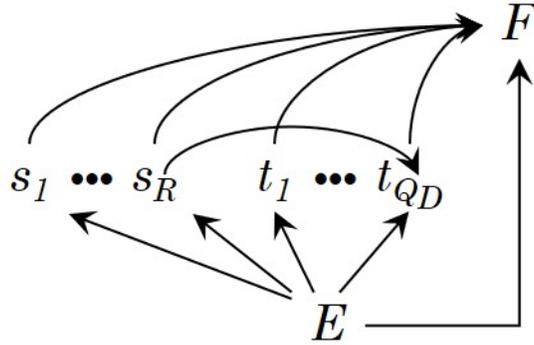
As a general argument, we state that trait diversity is fundamentally a function of species diversity. In a sample made up of a single species, the number of trait types will be those associated with the individual species. As the number of species increases in a sample, the number of trait types increases. For the specific case of  $R \rightarrow Q_D$ , it is easy to show that the number of traits will depend on the number of species. SZC do not seem to dispute the idea that  $Q_D \rightarrow F$ , and so we direct less of our discussion to that point.

### *A Simple Simulation Demonstrating $R \rightarrow Q_D \rightarrow F$*

In Figure 1E of the main text we propose a hierarchical causal diagram for the Standard Model (reproduced here as Figure S1). We demonstrate how this diagram is consistent with equations that provide a simple demonstration of how species richness ( $R$ ) can lead to changes in trait diversity ( $Q_D$ ), and in turn lead to changes in ecosystem function ( $F$ ). In our diagram (Fig. S1) it can be seen that species richness, defined as the number of species in a sample, represents the dimensionality of the vector of species abundances  $\{s_{i..R}\}$  in individual samples. Trait diversity, defined as the number of unique traits found in a sample, is represented by the dimensionality of the vector of traits  $\{t_{i..QD}\}$ .

Figure S1. Reproduction of Fig. 1E from main text. Causal diagram for BEF showing the relationships between the abundances of species in a sample  $\{s_{1..R}\}$ , traits found in that sample  $\{t_{1..Q_D}\}$ , and ecosystem function  $F$ .

**(Fig. 1E)** Hierarchical Causal Diagram for the Standard Model



$R$  is number of species in a sample.  
 $Q_D$  is number of traits in a sample.

In Box S1, we describe equations consistent with Figure S1. In this demonstration, we treat the environment ( $E$ ) as constant for all samples. Our simulation is meant to represent a biodiversity experiment where numerous levels of richness are created (see Schmid et al. 2002 for a review of the designs used in BEF experiments). For simplicity, each of the  $N$  species in the species pool can be classified into one of  $Q_T$  types, where each type represents a “trait type” and  $Q_T \leq N$ . Thus,  $Q_T$  refers to the total number of trait types in the species pool. As the number of species in a sample increases, we expect a natural increase in the number of traits found in a sample, i.e.,  $E(Q_D) = f(R)$ . For our demonstration, traits are quantified in terms of Resource Conversion Coefficients ( $RCCs$ ), which express how much each trait is capable of contributing to total biomass production in a sample (our measure of  $F$  for this example). Niche-complementarity is simulated by summing unique trait contributions over the trait vector, keeping species densities constant as  $R$  increases and allowing production to saturate as total contributions reach some maximum.

Box S1. Code illustrating expectations from causal diagram in Figure S1 (Fig. 1E in main text).

```
##### Simulating BEF Experiment
library(reshape)

### Step1: Create pool of species.
#traits quantified in units of Resource Conversion Coefficients
## For a pool of 20 species where there are 10 unique trait types
W <- c(1,1,2,2,3,3,4,4,5,5,6,6,7,7,8,8,9,9,10,10)

### Step 2: Sample different numbers of species from the species pool
set.seed(132)
n.spp = length(W) #species pool size
N=c(1:n.spp) #index variable for number of species
n.sims <- 10 #number of simulated replicates
ran.var = 1 #error variance for computation of function values
max.val = 60 #saturation value due to biotope-space saturation(1)
max.var = 5 #error variance for saturation value
Rich=c(1:n.spp) #vector of species richness values to be established
Qd=array(NA, c(n.sims,n.spp)) #n.sims by n.spp matrix of NAs for Qd matrix
F.dat=array(NA, c(n.sims,n.spp)) #n.sims by n.spp matrix of NAs for E(F)
F.dat.error=array(NA, c(n.sims,n.spp)) #n.sims by n.spp matrix of NAs for the errors of F

### Step 3: Simulate Data With Niche Complementarity
for (i in 1:n.spp){ #set different levels of plot richness

  for (s in 1:n.sims) { #create reps for each level of richness
    plot.sample <- sample(W, size=Rich[i]) #randomly sample species from species pool
    Qd[s,i] <- length(unique(plot.sample)) #determine number of traits in a sample
    F.dat[s,i] <- sum(unique(plot.sample)) #sum species contributions to plot
    F.dat.error[,i] <- rnorm(n.sims, 0, ran.var) #generate random errors for function ests
    F.dat[,i] <- F.dat[,i] + F.dat.error[,i] #add random error to function values
    F.dat[,i] <- ifelse(F.dat[,i] <max.val, F.dat[,i], max.val+rnorm(n.sims, 0, max.var))
  }

  N1.dat=melt(Qd)
  labs1 <- c("Rep", "R","Qdiv")
  colnames(N1.dat) <- labs1;
  N2.dat=melt(F.dat)
  labs2 <- c("Rep", "R", "RCC.sum")
  colnames(N2.dat) <- labs2;
  Nall.dat=cbind(N1.dat,N2.dat)
  Nall.dat2 <- Nall.dat[,-c(4,5)]
  head(Nall.dat2, 12L)

### Simple Inspections of the Data with Loess Fit
with(Nall.dat2, scatter.smooth(RCC.sum ~ Qdiv))
with(Nall.dat2, scatter.smooth(Qdiv ~ R))
with(Nall.dat2, scatter.smooth(RCC.sum ~ R))

### Model for Qdiv as function of R
formula1<-formula(Qdiv~a*R/(b+R))
a_start=10
b_start=8
m<-nls(formula1,start=list(a=a_start,b=b_start), data=Nall.dat2)
summary(m)

#Goodness of fit
cor(Nall.dat2$Qdiv,predict(m))

#plot
plot(Nall.dat2$R,Nall.dat2$Qdiv, xlab="Richness/Sample", ylab="Trait Diversity/Sample")
lines(Nall.dat2$R,predict(m),lty=2,col="red",lwd=3)

```

1– For a definition of “biotope space” see Dimitrakopoulos & Schmid (2004).

For the simple illustration we provide, it is possible to express the number of traits found in a sample ( $Q_D$ ) as a function of the number of species in that sample ( $R$ ). In our particular demonstration, there will be one trait type found in a sample comprised of one species. Further, if there are some species that are of the same trait types as others, the total number of traits in a

sample will be less than or equal to the number of species. This establishes that  $Q_D = f(R)$ . Procedurally, in our simulation random samples of  $R$  species are pulled from the species pool ( $W$ ) and the number of unique types of species are recorded ( $Q_D$ ).

Various mathematical expressions for  $Q_D = f(R)$  can be constructed. We can understand the process mechanistically by imagining we are drawing marbles of different colors from a jar, where the jar is our species pool for an experiment. We focus on drawing different numbers of marbles corresponding to our species richness treatments and this is replicated, allowing for chance to influence the exact number of colors found in a sample of marbles. Large samples will tend to contain all the colors in the jar, while samples of size one will contain only one color. If we were to know the proportional representation for each color in the jar, we could write a complete expression for the function. The relationship between  $Q$  and  $R$  can be approximated by a saturation function, such as:

$$Q_D = a * R / (b + R) \quad \text{eqn S1}$$

where  $Q_D = Q_{div}$  in the R code.

Figure S2 (first panel) shows how the number of traits  $Q_D$  increases with increasing  $R$  in our simulation, saturating as  $Q_D$  approaches the maximum number in the species pool  $Q_T$ . Also shown are how the sample sum of resource conversion coefficients  $RCC_{sum}$  increases with  $Q_D$  (second panel) as well as the net causal effect of  $R$  on  $RCC_{sum}$  (third panel). There is substantial experimental support for this empirical prediction. Fitting the data to a nonlinear model of the form in equation S1, leads to the result in Figure S3 ( $R^2=0.95$ ).

Figure S2. Descriptive loess plots.

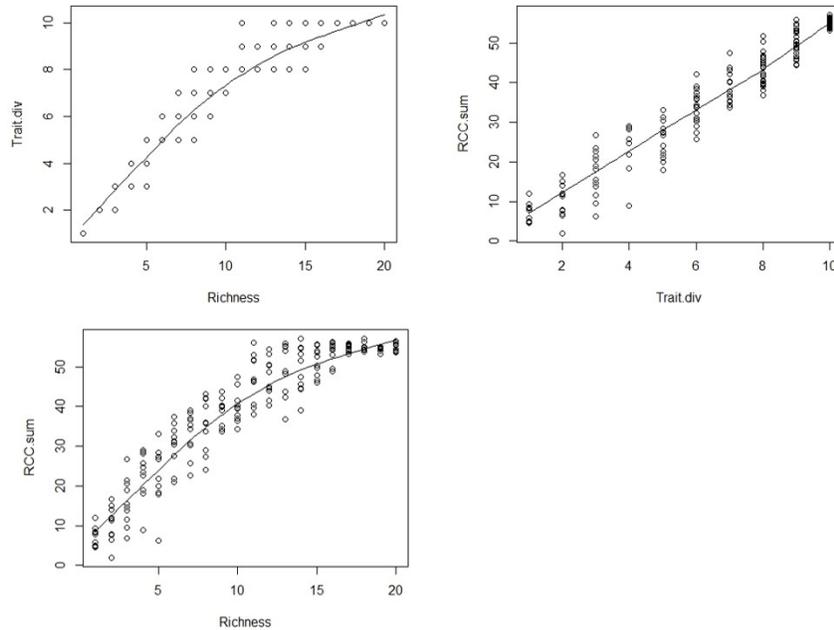
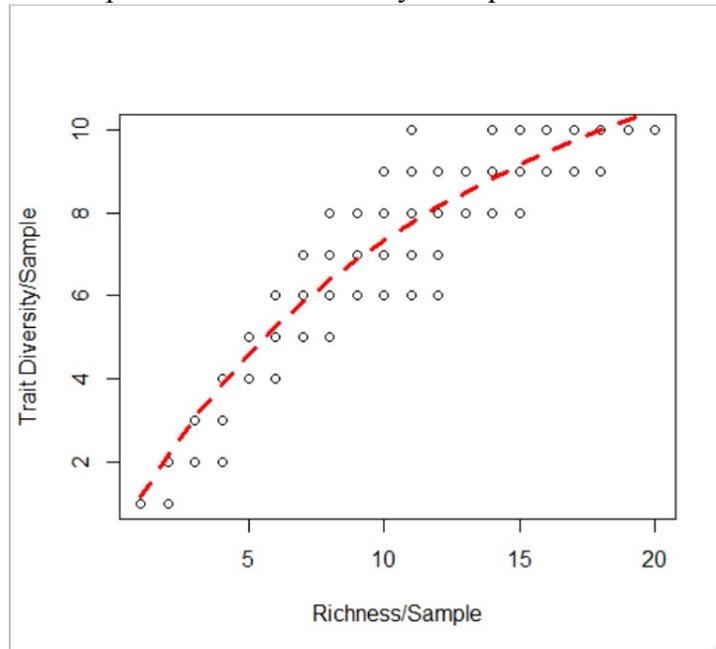


Figure S3. Fitted relationship between trait diversity and species richness.



*Relationship of this Simulation Demonstration to the Published Theoretical Literature Describing the Standard Model*

The above illustration simply shows that the number of traits in a sample does indeed depend on the number of species, contradicting SZC's second criticism of the Standard Model. There exists a significant body of published literature that provides a more complete exploration of hypothesized mechanisms for how variations in species diversity may propagate into variations in trait diversity and then into ecosystem function.

Classic theoretical models that predict the effects of biodiversity on ecosystem functioning consider this propagation of effects explicitly. For instance, Tilman et al. (1997) proposed a simple static model in which each species is able to cover only a small part of a 2-dimensional niche space. This model predicts that the portion of niche space covered by the community – which can be thought of as a measure of the community's trait diversity – increases asymptotically with species richness, very much like in Fig. S3. This increased trait diversity is then assumed to lead to a corresponding increase in some ecosystem function.

Loreau (1998) presented a mechanistic dynamical model of a nutrient-limited plant community in which plant individuals compete for soil nutrients. This model predicts a range of different relationships between (intra- or interspecific) diversity and ecosystem functioning depending on the degree of trait complementarity between plant individuals. This model made the propagation of effects from species richness to trait diversity to ecosystem function explicit, and was the basis for the development of the Standard Model (Loreau 2000; Loreau et al. 2001).

Subsequent models that further explored and predicted BEF relationships in spatially and/or temporally varying environments (e.g., Yachi & Loreau 1999; Loreau et al. 2003) or with different extinction scenarios (e.g., Gross & Cardinale 2005) implicitly or explicitly considered trait diversity as a variable that mediates the effects of biodiversity on the functioning and stability of ecosystems.

### *Comment on Complexities in Real World Settings*

BEF relationships are more complex in nature than represented in either SZC's presentation, our causal diagram (Fig. S1), or the simple simulation presented in this Appendix. In our presentation in the main text, we focus on describing the problems we find with SZC's presentation and claims. We contrast their model with the Standard Model, but we do not attempt to provide a comprehensive treatment for BEF relationships in experiments or natural systems, as this is a far more complex issue. First of all, numerous factors influence species richness aside from the environment, most notably competitive interactions and feedbacks from production to richness (Grace et al. 2016). Secondly, a full discussion of how traits mediate the effects of species identity on ecosystem function would need to consider a number of factors, which we briefly summarize here.

With regard to how real-world complexities relate to causal diagrams, it must be recognized that causal diagrams place no implied constraints on the functional form of the causal effects. Thus, Figure S1 simply implies that  $F = f(R, Q_D, \{s_{1..R}\}, \{t_{1..QD}\})$ , but the ways that causal elements combine is left undescribed and all possibilities (including interactions and other nonlinearities) are permitted. That said, it must be recognized that causal diagrams are very limited devices for expressing many of the interesting complexities that fascinate students of BEF. A result of this is a tendency for discussions based on causal diagrams to oversimplify discussions of mechanisms. In this section we seek to correct this limitation by making a more in-depth statement of what is known about processes consistent with the Standard Model. We further recognize that some of these issues are not yet settled.

### *Species Interactions can Affect Species Abundances and Species Traits*

First, the diagrams in Fig. 1 of the main text may seem to imply that species abundances are not influenced by species richness. This is an assumption implemented in rarely used additive designs applied in biodiversity experiments (Balvanera et al. 2006). This assumption can be interpreted as an extreme case of niche complementarity (Loreau et al. 2001), where individuals compete for resources only within but not between species. More frequently, biodiversity experiments have used substitutive designs, where species abundances are corrected for species richness in order to keep total community density constant. In this case, that is if species densities remain constant throughout the experiment, positive biodiversity effects must emerge via changes in individuals' traits as richness increases, allowing them to reach larger average individual sizes when interacting with individuals of other rather than the same species (this can affect many species or only one or few, corresponding to complementarity or selection effects (Loreau and Hector 2001)).

It is also true that many BEF experiments (*e.g.*, grassland BEF experiments) and in natural ecological systems, species abundance is allowed to vary through time, such that the initial substitutive design is not preserved and changes in EF result from additional changes in abundance. Furthermore, variations in trait diversity not only result from changes in species abundances but are also due to correlated changes in genotype distributions within species and plastic responses of plant genotypes to their biotic environment (see below). This raises the issue of whether we define traits as fundamental or realized traits – realized traits do change but fundamental traits (expressing the fundamental potential of a species or an individual) are expected to remain constant, *i.e.*, do not allow for evolutionary or plastic changes during the course of a study.

In real world settings, both increased community density (as simulated in additive designs) and increased average individual size (as observed in substitutive designs) may contribute to ecosystem function, as for example in subtropical forest (Barrufol et al. 2013). These effects of species interactions on species abundances  $s_i$  and on traits  $t_i$  are the essence of biodiversity effects on ecosystem functioning. Note that statistical analysis of additive designs in the absence of species interactions on average would find linearly increasing ecosystem function with increasing species richness, because monoculture performances of species could simply be added up. None of this complexity is suggested, much less sorted out, in the causal diagrams presented.

### *Trait Diversity is Difficult to Define and Measure in Practice*

Although the Standard Model is well justified on theoretical grounds (Loreau 1998, 2000), it has been proven extremely challenging so far to obtain a priori measures of trait diversity that predict ecosystem functioning (Huang et al. 2018, Barry et al. 2020, van der Plas et al. 2020). In a theoretical model using plants, Loreau (1998) showed that if all species in a local species pool compete for a single resource that is evenly distributed in the soil (biotope space), species with different rooting architectures or depths may better exploit the total resource pool, allowing rooting-depth diversity to explain diversity effects (Dimitrakopoulos and Schmid 2004). However, in more realistic situations, a plant community would need to have a trait diversity that allows it to cover all axes of the biotope space to extract as much as possible of the total available resource pool and convert it into some ecosystem function. At the same time, it should have a trait diversity that allows it to minimize losses to enemies and maximize gains from mutualists. Even without enemies and mutualists, the perfect plant community should contain species that allow it to place its roots and leaves in such a way as to fill total biotope space across seasons and years and to take up all the different types of resources, *e.g.*,  $N_2$ , nitrate, ammonia or amino acids for nitrogen. Such a trait diversity must necessarily be multidimensional (Villéger et al. 2008) and reflect different aspects that allow the community to fill all the axes of the biotope space.

Furthermore, trait diversity should also include within-species diversity. Obviously, single species can already place their roots and leaves at a variety of depths or heights. Thus, particular species may completely cover one biotope axis whereas several species may be needed to cover it in the absence of these particular species. If species differ in which biotope axis they cover more or less completely, then in one community species richness may be important because it increases rooting-depth diversity (because the component species do not have large intraspecific rooting-depth diversity) and in another community species richness may be important because it increases foliage-height diversity. In this case, species richness is a good predictor of ecosystem function because it predicts rooting-depth diversity in one and foliage-height diversity in another community. However, combining these two mechanisms by which species richness, depending on the particular species composition, causes ecosystem function may be very difficult.

Finally, trait means and variances are not constant across environments, especially not across different richness levels. Trait values measured for a species in one community can differ strongly from trait values of the same species measured in another community, with corresponding consequences for ecosystem functioning (Roscher et al. 2018). Trait means and variances can evolutionarily change in response to diversity, leading to high within-species trait diversity under selection in monoculture (van Moorsel et al. 2018) and to high between-species trait variation under selection in mixture (Zuppinge-Dingley et al. 2014). This emphasizes the need to include species interactions in realistic interpretations of the Standard Model. Thus,

while the Standard Model holds, because of species interactions, diagrams that fail to explicitly include them can only be used to discuss some general issues, as we do in the main text.

To illustrate this further, in some real-world studies the only way in which functional trait diversity has been able to accurately explain and predict ecosystem functions is by using a posteriori clustering approaches that cluster species in ad-hoc functional groups based on the realized functional properties of species in their environment (Jaillard et al. 2018). This is a sobering message, which suggests that in practice ecosystem functions cannot easily be predicted from a priori knowledge of species traits, independently from their biotic and abiotic context.

### *Literature Cited*

- Balvanera, P., A. B. Pfisterer, N. Buchmann, J. S. He, T. Nakashizuka, D. Raffaelli and B. Schmid. 2006. Quantifying the evidence for biodiversity effects on ecosystem functioning and services. *Ecology Letters* 9:1146-1156.
- Barry, K. E., J. Van Ruijven, L. Mommer, Y. Bai, C. Beierkuhnlein, N. Buchmann, H. De Kroon, A. Ebeling, N. Eisenhauer, C. Guimarães-Steinicke, and A. Hildebrandt. 2020. Limited evidence for spatial resource partitioning across temperate grassland biodiversity experiments. *Ecology* 101(1), p.e02905.
- Barrufol, M., B. Schmid, H. Bruelheide, X. Chi, A. Hector, K. Ma, S. Michalski, A. Tang, and P. A. Niklaus. 2013. Biodiversity promotes tree growth during succession in subtropical forest. *PLoS one*, 8(11), p.e81246.
- Dimitrakopoulos, P. G. and B. Schmid. 2004. Biodiversity effects increase linearly with biotope space. *Ecology Letters* 7:574-583.
- Grace, J. B., T. M. Anderson, E. W. Seabloom, E. T. Borer, P. B. Adler, W. S. Harpole, Y. Hautier et al. 2016. Integrative modelling reveals mechanisms linking productivity and plant species richness. *Nature* 529, no. 7586: 390-393.
- Gross, K. and B. J. Cardinale. 2005. The functional consequences of random vs. ordered species extinctions. *Ecology Letters* 8:409–418.
- Huang, Y., Chen, Y., Castro-Izaguirre, N., Baruffol, M., Brezzi, M., Lang, A., Li, Y., Härdtle, W., von Oheimb, G., Yang, X., Pei, K., Both, S., Liu, X., Yang, B., Eichenberg, D., Assmann, T., Bauhus, J., Behrens, T., Buscot, F., Chen, X.-Y., Chesters, D., Ding, B.-Y., Durka, W., Erfmeier, A., Fang, J., Fischer, M., Guo, D., Guo, L.-D., Gutknecht, J.L.M., He, J.-S., He, C.-L., Hector, A., Hönig, L., Hu, R.-Y., Klein, A.-M., Kühn, P., Liang, Y., Michalski, S., Scherer-Lorenzen, M., Schmidt, K., Scholten, T., Schuldt, A., Shi, X., Tan, M.-Z., Tang, Z., Wang, Z., Welk, E., Wirth, C., Wubet, T., Xiang, W., Yan, J., Yu, M., Yu, X.-D., Zhang, J., Zhang, N., Zhang, S., Zhou, H.-Z., Zhu, C.-D., Zhu, L., Bruelheide, H., Ma, K., Niklaus, P.A., Schmid, B. 2018. Impacts of species richness on productivity in a large-scale subtropical forest experiment. *Science* 362:80-83.
- Jaillard, B., C. Richon, P. Deleporte, M. Loreau, and C. Violle. 2018. An a posteriori species clustering for quantifying the effects of species interactions on ecosystem functioning. *Methods in Ecology and Evolution* 9:704-715.
- Loreau, M., 1998. Biodiversity and ecosystem functioning: a mechanistic model. *Proceedings of the National Academy of Sciences* 95:5632-5636.
- Loreau, M., 2000. Biodiversity and ecosystem functioning: recent theoretical advances. *Oikos*, 91:3-17.
- Loreau, M. and A. Hector. 2001. Partitioning selection and complementarity in biodiversity experiments. *Nature* 412:72-76.

- Loreau, M., S. Naeem, P. Inchausti, J. Bengtsson, J. P. Grime, A. Hector, D. U. Hooper, M. A. Huston, D. Raffaelli, B. Schmid, D. Tilman, and D. A. Wardle. 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* 294:804-808.
- Loreau, M., N. Mouquet, and A. Gonzalez. 2003. Biodiversity as spatial insurance in heterogeneous landscapes. *Proceedings of the National Academy of Sciences of the United States of America* 100:12765–12770.
- Roscher, C., J. Schumacher, M. Gubsch, A. Lipowsky, A. Weigelt, N. Buchmann, E. D. Schulze and B. Schmid. 2018. Interspecific trait differences rather than intraspecific trait variation increase the extent and filling of community trait space with increasing plant diversity in experimental grasslands. *Perspectives in Plant Ecology, Evolution and Systematics*, 33:42-50.
- Schmid, B., A. Hector, M. A. Huston, P. Inchausti, I. Nijs, P. W. Leadley, and D. Tilman. 2002. The design and analysis of biodiversity experiments. Pages 61-75, In M. Loreau, S. Naeem, and P. Inchausti, editors. *Biodiversity and ecosystem functioning: synthesis and perspectives*. Oxford University Press.
- Tilman, D., C. L. Lehman, and K. T. Thomson. 1997. Biodiversity and ecosystem productivity: theoretical considerations. *Proceedings of the National Academy of Sciences of the United States of America* 94:1857–1861.
- van der Plas, F., T. Schröder-Georgi, A. Weigelt, K. Barry, S. Meyer, A. Alzate, R. L. Barnard, N. Buchmann, H. de Kroon, A. Ebeling, and N. Eisenhauer. 2020. Plant traits alone are poor predictors of ecosystem properties and long-term ecosystem functioning. *Nature Ecology & Evolution*, 4:1602–1611.
- van Moorsel, S. J., T. Hahl, C. Wagg, G. B. De Deyn, D. F. Flynn, D. Zuppinger-Dingley and B. Schmid. 2018. Community evolution increases plant productivity at low diversity. *Ecology Letters*, 21:128-137.
- Villéger, S., N. W. Mason and D. Mouillot. 2008. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology*, 89:2290-2301.
- Yachi, S. and M. Loreau. 1999. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. *Proceedings of the National Academy of Sciences of the United States of America* 96: 1463–1468.
- Zuppinger-Dingley, D., B. Schmid, J. S. Petermann, V. Yadav, G. B. De Deyn and D. F. Flynn. 2014. Selection for niche differentiation in plant communities increases biodiversity effects. *Nature*, 515:108-111.