Biodiversity Skepticism and Measurement Practices

Forthcoming in *Biology & Philosophy*

Abstract

This paper challenges "biodiversity skepticism:" an inferential move that acknowledges the proliferation, heterogeneity, and lack of co-variance of biodiversity measurements, and concludes that we should doubt the scientific validity of the biodiversity concept. As a way out of skepticism, philosophers have advocated for eliminating "biodiversity" from scientific inquiry, revising it, or deflating its meaning into a single measurable dimension.

I present a counterargument to the inferential move of the skeptic by revealing how it stands on two unstated premises, namely a reflective view of measurements and the unidirectional dynamics between definitional and measurement practices, and corollary assumptions. These premises and assumptions are misaligned with a richer theoretical understanding of measurement and are sometimes inconsistent with how science operates. A more nuanced view of measurement can better explain measurement proliferation while being consistent with new ways in which the general biodiversity concept could be useful.

To conclude, I urge philosophers of measurement and conceptual engineers to collaborate in tackling the interplay between conceptual change and measurement practices.

Keywords: biodiversity, measurement, theoretical concept, biodiversity practices.

1 Introduction

> "Before measurements can be meaningful they must be directed to the right things and, even in science, finding these things is the major achievement; entitation is more important than quantitation." (Gerard, 1965, p.762)

Philosophical debates increasingly recognize that reliable measurements do not result from mere observations but require complex conceptual machinery (Tal, 2015, Ohnesorge, 2022). The conceptualization of an object of scientific interest—either a phenomenon, property, or entity—is essential for reliable measurements. Often, the way the object of measurement, henceforth "measurand", is conceptualized results in a multiplicity of measurements.

Frequently, the existence of a plurality of partial, even inconsistent, measurements for the same measurand is not a problem. Due to the complex nature of certain measurands, a unique measurement procedure is not expected. Measurements may be *adequate for specific purposes* (*sensu* Parker, 2020; Bokulich & Parker, 2021) and different measurements could be validated in separate domains. Scientists in one domain might find specific measurements more suited to their task, without questioning different measurements of the same measurand performed using different techniques. In other contexts, different measurements might be encouraged in robustness arguments. If independent measurements yield similar trends, then stronger inferences about the object can be made. A plurality of methods might even be required to grasp a measurand from various angles.

However, measurement proliferation and heterogeneity can be understood as a symptom that *something about the object intended to be measured is not grasped yet.* Some might argue that the lack of a unique measurement or a set of covarying measurements might depend on the *non-measurable nature* of the object itself. Intelligence and talent, for example, are phenomena that have been criticized as not really being measurable, especially in the form of quantified measurements or standard tests (see Serpico, 2021). Others argue that a plurality of measurements or the difficulty of settling on the best procedure might depend on the *non-existence of a measurable phenomenon or property in the first place* (Zhao, 2023). We might have a concept for a measurand—think of the ether or the four humors in ancient Greek medicine—yet these concepts lack a referent: failure to measure said phenomena is a clue that they do not exist. Given the weight that science generally attaches to measurements, especially the quantitative information obtained via specific measurement processes, certain objects traditionally investigated in science might not be genuinely susceptible to scientific investigation. Although controversial, this view is widely accepted (Porter, 1996; Fischer & Young, 2007; Muller, 2018).

In sciences dealing with biodiversity, reflections about measurement practices, their validation and interdependence abound. Biodiversity measurement practices include the identification of biodiversity indicators, the formulation of indices, and the development of tools and instruments for data collection and analysis. Biodiversity *indicators* pick aspects of a complex object of measurement qualitatively, facilitating quantification (Duelli & Obrist, 2003, Pereira et al., 2013, Rochette et al., 2019). There are no overall best indicators of biodiversity because they are "expressive of particular sets of concerns" (Williamson & Leonelli, 2022, p.178). Indices are mathematical formulas that map the *variability* of an indicator expressing this variability on a scale (Morgan, 2007), and play the epistemic role of capturing properties of an object using a single number (Alexandrova, 2017b), facilitating comparisons and predictions. Some indices are helpful to understand taxa fluctuations after a disturbance, others are suited to indicate the most fragile ecosystem, and others to map feedback loops between trophic levels (Daly, Baetens, & De Baets, 2018). The most common *index* is Shannon's diversity index (Shannon, 1948). Data collection and analysis, such as recording physical measures of an area's surface to calculate species density, are crucial measurement practices. We have witnessed a spike in data availability and emphasis on biodiversity quantification (Heberling et al., 2021) and biodiversity science depends on the availability of large amounts of heterogeneous biodiversity data (Kelling et al., 2009).

Philosophers have long pointed out how the plurality and heterogeneity of biodiversity measurements coincide with a paucity of theoretical views on how to interpret this proliferation (Maclaurin & Sterelny, 2008a). Oftentimes, this lack of a unifying framework has resulted in suspicions about the referential scope of "biodiversity" and about the usefulness of the theoretical concept itself (e.g. Maier, 2013; Santana, 2014; Meinard, Coq, & Schmid, 2019; Reydon, 2019). This paper addresses the distrust around the utility of the biodiversity concept in relation to the plurality and heterogeneity of measurement practices, which often leads to eliminativist, deflationist, or revisionist positions.

The argument proceeds as follows. I first introduce three common perspectives around biodiversity as *a general concept* (henceforth "biodiversity"), eliminativism, deflationism, and revisionism (Section 2). I argue the three positions share a common inferential move: they challenge the scientific validity of "biodiversity" from the plurality and heterogeneity of its measurements (Section 3). I call this inferential move "biodiversity skepticism" and expose two crucial unstated assumptions and a set of corollary assumptions that must be in place for the inference to work. I then problematize each assumption by arguing that they are at odds with our best understanding of measurement theory and fail to acknowledge how measurement practices often unfold (Section 4). While I do not directly argue for a specific approach to measuring biodiversity or for the utility of the general concept, I demonstrate that the skeptical move is misplaced. In Section 5 I more broadly argue that the interrelation between conceptual development and measurement practices is a topic in need of additional scrutiny. I then encourage conceptual engineers and philosophers studying concept change to give greater importance to measurement practices, particularly when dealing with scientific concepts.

2 Biodiversity Eliminativism, Deflationism, and Revisionism

In its most iconic formulation in the Convention on Biological Diversity (CBD), biodiversity is characterized as:

> "the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part: this includes diversity within species, between species and of ecosystems"(CBD, (1992), article 2).

The CBD does not specify how to quantify the various aspects of biodiversity, it rather characterizes biodiversity in general terms consistent with a plurality of research subfields. Today, biodiversity research includes conservation biology, whose primary business is the assessment and protection of threatened species; ecology, which investigates the resilience and function of ecosystems in relation to various types of diversity, and biological control science (of which crop science is a branch), which tackles the impact of pests

(Duelli & Obrist, 2003). Each field identifies and studies specific aspects of diversity. Conservation biology primarily addresses biodiversity as *species* diversity (a subset of taxonomic diversity). In ecology, where patterns of change or resilience are the main study subject, community composition is the preferred metric (Lovejoy, 1994). In contrast, agronomists tend to focus on genetic diversity, especially genetic characteristics that favor adaptability in novel conditions (Williamson & Leonelli, 2022).

Even if "biodiversity" is standard scientific jargon, an exhaustive scientific definition of biodiversity is subject to an ongoing controversy. Several scientists have agreed that the meaning of biodiversity established in the CBD and then adopted in the scientific community is too broad and vague, to such an extent that they call biodiversity a "buzz word" (Rousseau, 1998), or a word "so all-inclusive that it has little more than "feel- good meaning" (Lautenschlager, 1997, p.683). Others, instead, claim that the orthodox meaning is too limiting and should be expanded. For example, the initiative "Biodiversity Revisited" aims to expand the meaning traditionally associated with the term to reflect our evolving scientific and public understanding of biodiversity (Díaz, 2019). Yet, today we are facing a proliferation of nonoverlapping, sometimes inconsistent meanings that have blossomed out of biodiversity research, and sparked philosophical interest (Sarkar, 2005; Maclaurin & Sterelny, 2008a; Santana, 2014 and 2018; Justus, 2011; Crupi et al., 2018; C. H. Lean, 2017). This conceptual messiness has polarized the conversation around the concept of biodiversity in three standard positions: eliminativism, deflationism, and revisionism.

Concept eliminativism is a well-established way of criticizing certain scientific or pseudo-scientific notions or entire lines of research. Eliminativists often argue that a concept should be eradicated from scientific or everyday conversation due to reference failure. One type of evidence for reference failure is the proliferation of non-overlapping characterizations or measurements for the same term. Concepts that have received resonant eliminativist criticisms are "species" (Ereshefsky, 1998),"pain" (Corns, 2016), "race" (Zack, 2014; Appiah, 2006) "sex" (Watkins & Di Marco, MS). For example, the proliferation of heterogeneous, non-overlapping species concepts constitutes one of the reasons Mark Ereshefsky (1998) used to lean toward an anti-realist, eliminativist position about the species concept. According to one formulation of his species eliminativism, Ereshefsky argues that the plurality of meanings associated with the term "species," stemming from its diverse applications, suggests that

> "there is no unified ontological category called 'species'. It implies that the species category does not exist. This I take to be the strong argument from species pluralism to the non-existence of the species category" (Ereshefsky, 1998, p.113).

According to Ereshefsky, the concept of species should be eliminated and only more specific concepts should be used instead.

Along these lines, Carlos Santana has argued that "the biodiversity concept is a poor fit for the role we want it to play in conservation biology on both empirical and conceptual grounds" (2014, p. 761). He argues that focusing on biodiversity *per se* misses the target of conservation, which is instead the assessment and preservation of "biological values", such as the specific richness of a biodiversity hotspot or the evolutionary history of a taxon, dimensions that are more easily quantified. Following Ereshesfy's strategy, Santana argues that the concept of biodiversity should be eliminated: since biodiversity is a heterogeneous, complex concept, whose various operationalizations diverge, biodiversity is not a real "natural property" or a "natural kind" that can be uniquely and consistently measured. Accordingly, it is not helpful for science to rely on a unique concept of biodiversity and expect it to explain or predict. For Santana, the general concept of biodiversity can be discarded without scientific loss: no research question actually relies on the general concept.[1]

Other philosophers have chosen less extreme critical positions. Concept deflationism is a reductionist view according to which what can be meaningfully said about a concept should be distilled in one and only one of its most significant dimensions. A deflationist is interested in singling out an optimal dimension to signpost a more complex concept, usually motivated by pragmatic considerations, such as frequency of use or historical-social significance. Deflationism is a genuine alternative to eliminativism as it does not suggest getting rid of a general concept; it just calls for a simplification of the intensional landscape a complex concept should occupy.

Deflationists about the concept of biodiversity champion solving the dissatisfaction about biodiversity pluralism by reducing the complexity of the concept to one of its components. Since the term pervades the science-policy

[1] Reydon has likewise argued that "generally, it does not make much sense to talk about the biodiversity of a particular region or the biodiversity of planet Earth, because there are too many distinct aspects to consider and a region can be very diverse in some aspects and much less diverse in others" (2019, p. 169).

debate for decades, it is not likely to be abandoned. However, as complex as it is now, it is likely to create more issues. It is better to then simplify and standardize its usage, even at the cost of sacrificing some of its fecundity. Sahotra Sarkar advocated for *simplifying* the concept of biodiversity by semi-conventionally selecting some of the dimensions associated with the concept (2002, 2005). Restricting the concept of biodiversity to one or more of its proxies, among which species richness appears to be one of the best candidates, would allow a more practical quantification and consistent utilization and more on-target assessment in conservation science.[2] Another deflationist position equates biodiversity with the benefits derived from natural variability, specifically the services provided to humans. This perspective has been debated and partially abandoned by organizations such as IPBES, as noted in Faith (2018). As an instance of this deflationist approach, Mace and colleagues argue that despite the inherent complexities in defining both biodiversity and ecosystem services, there is a significant overlap as biodiversity contributes to the "delivery of ecosystem services" along several dimensions (Mace, Norris, & Fitter, 2012, p.25). This overlap justifies a conceptual pairing of the two concepts. The practical implication of this identification is that conservation efforts, which traditionally focus on protecting ecosystem services, simultaneously contribute to the protection of biodiversity.

Conceptual revisionism is a view according to which ambiguous concepts can be made more precise or ameliorated. The revised concept should better align with new approaches, theories, or empirical evidence, making it more accurate, interoperable with related notions, and more applicable in policy contexts. Two examples of conceptual revisionism in science include the notions of race (Glasgow, Haslanger, Jeffers, & Spencer, 2019) and pregnancy (Meincke, 2022). Both revisionism and deflationism advocate for clearer and more distinct characterizations of concepts, but they differ fundamentally. Revisionism embraces the functional role a concept can have, aspiring to align its usage with promising scientific knowledge or non-epistemic commitment (for example, to fairness), even if this entails dropping the original meaning. In contrast, deflationism focuses on reducing concept complexity

[2]Sarkar named "true surrogates" and "indicator-surrogates" those pivotal aspects of biodiversity that should be understood when using the concept. True surrogates would be more general proxies for biodiversity, according to him, such as taxonomic diversity or ecosystem health, whereas indicator-surrogate would be context-specific, for example, when the status of a certain species is elected as a representative for the health of an entire community.

for simplicity and standardization, independent of best scientific knowledge or any normative intent.

In discussions on biodiversity, revisionists maintain that the concept remains essential to conservation. However, a more robust notion of biodiversity than that suggested by the CBD must accompany conservation efforts. This revised notion should meet some desiderata for good scientific concept, for example, being theoretically fundamental, tractable, representative, and normative (C. H. Lean, 2017). Noteworthy revisionist accounts maintain that biodiversity should encapsulate variety as both a result and a condition for evolution. Accordingly, the goal of conservation is to maintain the mechanisms promoting variety (Fan et al., 2020). Motivated by the need for a general measure of biodiversity for global conservation, two similar revisionist accounts have emerged: one pioneered by Daniel Faith (1992) and the other by Christopher Lean and James MacLaurin (2016). Both accounts suggest revising "biodiversity" to mean "phylogenetic diversity," which measures lineage diversification, indicating how evolution differentiates populations. Phylogenetic diversity serves as an effective general measure of biodiversity, capturing how diversity and variety are created and maintained evolutionarily, supported by sophisticated and comparable mathematical tools. Since revisionists assume biodiversity is instrumentally valuable and should be protected as a safeguard against an uncertain future (Faith, 2021, C. Lean & Maclaurin, 2016), conceptualizing biodiversity as phylogenetic diversity provides a useful framework for global conservation.

Eliminativists, deflationists, and revisionists all critique the general biodiversity concept as formulated in the CBD as too broad and complex to effectively capture a phenomenon of interest, thereby limiting its utility to conservation. The three views have been exhaustively presented, criticized, and defended. Here, I argue that there is a different and more interesting way of looking at this debate. I will pay attention to a striking similarity among these three positions: the centrality of measurement practices and their interplay with the general biodiversity concept. This shared concern underscores a broader issue that the three positions confront: the challenge of measurement proliferation.

3 Biodiversity Skepticism, Measurements, and Theoretical Concepts

One of the major observations in this paper is that what is really at stake in the three positions presented above is a concern around measurements. Collectively, eliminativism, deflationism, and revisionism regarding biodiversity lament that the proliferation of heterogeneous, non-overlapping measurements indicates a fundamental issue with how biodiversity is conceptualized.[3] Measurement plurality and heterogeneity have often been used to cast doubt on the nature of biodiversity as a concept and the phenomenon or entity it purportedly captures, in philosophy as well as in science (see, for example, Hurlbert, 1971; Schmeller et al., 2017; Meinard et al., 2019). A common argument underneath this suspicion is that most measurements are domain-specific and rely on different conceptualizations of biodiversity, distinct theoretical assumptions, and the trends they identify sometimes do not covary.[4] Santana argues that the concept of biodiversity is meaningless due to the multiplicity and inconsistency of measurements, while Sarkar and Lean try to identify the best measurement in response to proliferation. The literature predominantly focuses on how "capricious" biodiversity measurements can be (Sarkar, 2002, p. 133) and the "temptation" to reach a unified measure (Maclaurin & Sterelny, 2008a, p.172)

The difference between the three views lies in the solutions they propose to address this measurement plurality and heterogeneity. An eliminativist argues that faced with difficulties in measuring exactly a general notion of biodiversity, the concept itself is not useful, and conservationists should get rid of it. Rather, they should limit themselves to individually targeting, measuring, and assessing the numerous components of biodiversity, such as

[3]Notice that analogous arguments could be assembled in other disciplines: the concepts of "quality of life" and "well-being" might and sometimes did lead to skepticism. The above argument is thus a token of a broader argumentative strategy that affects social and behavioral sciences.

[4]For example, an ecosystem might show high genetic diversity but low functional diversity, or vice versa. Studies conducted at different scales, such as the Living Planet Index, for example, show a *global* taxonomic diversity decline by around 28% between 1970 and 2008, based on the numbers of individuals monitored. On the other hand, recent studies have shown that *local* taxonomic diversity is increasing. This boost results from a 7.6% increase per decade in the number of species present in sampled plots (Vellend et al., 2013).

species abundance, ecosystem functionality, etc. These dimensions are properly quantifiable with validated measurements and can serve the purposes of science better. Deflationists and revisionists agree that addressing measurement proliferation requires a monodimensional approach to studying biodiversity. Deflationists argue for simplifying measuring "biodiversity" to one of its dimensions or using a specific measurement as a proxy for the entire concept. This proxy should be preferred because it represents the most intuitive, valuable, or manageable aspect of biodiversity, such as species richness (e.g., Sarkar, 2002, Maclaurin & Sterelny, 2008a). Revisionists instead contend that the best measure should be picked using the most updated theories and data within conservation. The development of sophisticated measurement practices provides evidence that the dimension highlighted by these measurements constitutes an improvement of the concept (e.g., Faith, 1992, C. H. Lean, 2017).

I will refer to the shared attitude of distrust toward a general biodiversity concept as "biodiversity skepticism." Biodiversity skepticism is the argumentative stance that begins by acknowledging the complexity of measuring biodiversity and concludes by rejecting "biodiversity" as a useful scientific tool. In short, according to the skeptical argument, "biodiversity" is not a useful concept in science because it is not (consistently and coherently) measurable. As such, "biodiversity" should either be eliminated or modified to fit one of its truly measurable dimensions.

The skeptical viewpoint is certainly attractive, and its advocates have clearly shown a solid understanding of the tradeoffs and conceptual implications of measurement pluralism. They also present compelling arguments for validation criteria within particular domains. However, for biodiversity skepticism to be sound, strong presuppositions about the interplay between theoretical concepts and measurement practices must be assumed, and this is rarely done explicitly (Fig.1). My first goal here is to reveal the conceptual scaffolding underlying the skeptical move from measurement proliferation to eliminativism, deflationism, or revisionism. I display two major assumptions that support skepticism, along with some of their corollaries. In Section 4, I will then demonstrate how these assumptions conflict with our best understanding of the process of measuring and with scientific practices. Thus, I will be able to contend that questioning the value of "biodiversity" due to its measurements is unwarranted.

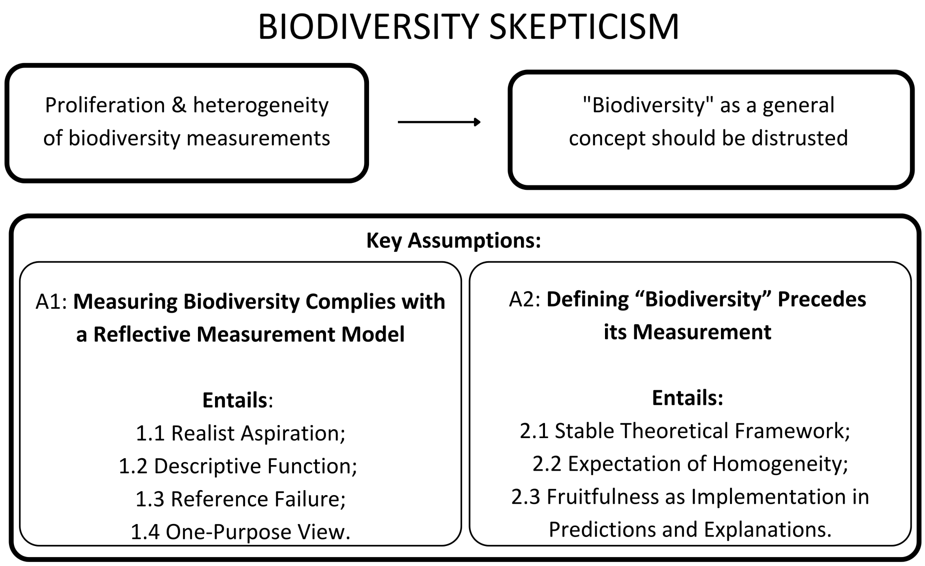

Figure 1: Summary of the skeptics' argumentative move and unstated assumptions.

Assumption 1: **Measuring Biodiversity Complies with a Reflective Measurement Model**

The first key assumption in biodiversity skepticism, I argue, is the implicit support of a specific view of measurement: a "reflective measurement model". A reflective measurement model describes the relationship between measurement results and the objects being measured. This view is quite prevalent in disciplines where measurement practices have led to advanced debates, particularly in psychometrics and econometrics. The underlying idea here is that when attempting to measure an object, the measurements acquired are caused by the object itself (Conway & Kovacs, 2015).[5] Accordingly, changes in the recorded measurements are caused by changes in the measurand (Coltman, Devinney, Midgley, & Venaik, 2008, Borsboom, 2005).

[5]The technical term for a measurand in this context is "latent construct." Reflective measurement models are a subset of models within "latent variable theory", a qualification used in psychometrics to describe structural relationships between measurements (in terms of behaviors, questionnaire responses, MRI readings) and the intended measurand, which does not necessarily have to be causal. (Van Bork, Wijsen, & Rhemtulla, 2017)

Importantly, when multiple indicators are available for the same measurand, a causal model *entails* an expectation of co-variation or some form of tuning among the measurements of the various dimensions, as a reflection of the latent measurand causing the observations (Van Bork et al., 2017). For short, the term "reflective" emphasizes that the outcomes of a successful measurement process should represent the evolving status of a measurand, an inference ensured by assuming a causal connection.

Biodiversity skepticism, I argue, assumes a reflective approach to measurement, when biodiversity (whether an entity or a property) would be the cause of the measurements results. This assumption becomes particularly evident in the problematization of measurement proliferation and heterogeneity itself. If one assumes that biodiversity measurements are caused by a unique underlying phenomenon captured by the general notion of biodiversity, then one expects the various measurements to covary with the trends in the measurand. However, the lack of consistent, covariant measurements for biodiversity suggests that these measurements are not caused by the same measurand. Consequently, the concept of biodiversity itself can be doubted as incapable of reflecting that unique cause of the measurements. In other words, the assumption of a reflective model underpins skepticism by demanding a clear and direct causal link between the concept of biodiversity and consistent measures. Without such a link, skeptics can conclude that the concept cannot properly serve science. I have identified four additional stances entailed by this reflective approach.

Sub-Assumption 1.1: **Realist Aspiration**

First, reflective measurement models posit the existence of the object being measured as the validation condition for measurement (Conway & Kovacs, 2015, Borsboom, Mellenbergh, & Heerden, 2003, Coltman et al., 2008. I will call this attitude a "realist aspiration", which spreads from the rather common practice, in science, of drawing "inferences about the structure of the world on the grounds of measurement results alone" (Zhao, 2023, p. 237). Zhao argued that it is logical to move from recognizing multiple measurements to postulating the existence of an entity as their cause. This is an inference to the best explanation: what better explains the abundance of measurements than the existence of an object causing those measurements?

Biodiversity skepticism arises from failure to meet this realist expectation.

Biodiversity skepticism relies on this linear inference from measurements to the intended objects of the measurement but starts by acknowledging the messiness of measurements. The skeptic grants that most biodiversity measurements are valid in conservation, ecology, and crop science as they can be used in various predictions or explanations. Since biodiversity is captured by a set of valid yet incoherent measurable procedures, it is concluded that there is no unique property or phenomenon being quantified or at least that general biodiversity is intrinsically non-measurable. The proliferation of validated yet non-overlapping measurements demonstrates that there is no existing phenomenon referred to as "biodiversity." Skepticism often relies on an ontological disappointment. Santana supports eliminativism after a failure to satisfy his realist aspiration, exemplified by his argument that "biodiversity" cannot be a natural kind, an explanandum, or a scientific property (Santana, 2018, p.15). On the other hand, Lean advocates for revising the concept of biodiversity to refer to the phylogenetic structure of lineages, which, in his view, quantifies a mind-independent biological property (C. H. Lean, 2017).

Sub-Assumption 1.2: **Descriptive Function of Measurements**

Second, endorsing a reflective model of measurement aligns well with a specific understanding of the primary function ascribed to measurements, namely that of *describing* a phenomenon.[6] It is a common view, in environmental sciences and beyond, that measurements, especially quantitative measurements, operate as a mirror into the world (Turnhout, Neves, & De Lijster, 2014,Pine & Liboiron, 2015, Merry, 2009). Very often, this function grants them privilege in guiding policy (Muller, 2018, Porter, 1996)

Measures of biodiversity should ideally discover and reflect the state of an existing phenomenon or property if the skeptical argument has to hold. Indeed, measurement pluralism becomes a problem insofar as the represented object or objects are not clear. Assuming that the same object cannot exhibit inconsistent values or trends, measurement procedures that represent a measurand inconsistently are failing their representative function with respect to that measurand. One could criticize the individual measurements as valid,

[6]Other functions of measurement consistent with this view are that measurements also function as a guide to policy (normative function) or that they determine how a phenomenon is perceived based on which of its features are better represented (constitutive function)

but this is not the strategy that biodiversity skeptics usually follow. The representative failure is instead taken to reveal the inadequacy of biodiversity measurements to comply with a reflective model that aims to uniquely and consistently model biodiversity. Consequently, something about the alleged object of measurement is misunderstood, such as the fact that other measurands are captured.

Sub-Assumption 1.3: **Reference Failure**

Third, the ontological machinery presented above is also used to signal a failure of reference—when the extension of a term is empty as it does not successfully refer to any actual entity or phenomenon, despite its current usage. In philosophy of science, problems with terms' semantics has led to eliminativist and revisionist projects as it is commonly believed that the retention of a concept should at least be ensured by its ability to consistently refer to something.

In the context of biodiversity, the skeptics assume that because the various measurements of biodiversity do not seem to point to a singular, identifiable phenomenon, the term "biodiversity" itself is referentially empty. Accordingly, the only function that "biodiversity" can have is to almost rhetorically embrace a series of subconcepts that are described and individuated with specific measurements; "biodiversity" becomes a placeholder or umbrella term. Placeholders and umbrella terms are referentially opaque: it is not indeed clear which of the many meanings, the term refers to. Keeping "biodiversity" as an umbrella term can even be misleading as it gives the impression that there is a single object of study that various disciplines share, which the skeptical argument criticizes. While reference failure justifies eliminativism according to Santana (2018), Lean's revisionism could be interpreted as a way of avoiding reference failure. For Lean, an optimal biodiversity dimension represented by a robust measurement would be the ideal referent for the general concept of biodiversity (C. H. Lean, 2017, p.1088).

Sub-Assumption 1.4: **One-Purpose View**

Fourth, the skeptics' acceptance of a plurality of valid measurements is compatible with an adequacy for purpose view of measurement. The idea here is that measurements are to be evaluated based on the degree to which they attain "epistemic or practical aims that interest their users" (Bokulich & Parker, 2021, p. 31). Validity is determined by whether the measurements fulfill the task they were created for. The tasks that biodiversity measurements can perform include classification, quantification, and performativity in explanatory or predictive models. Faith's index, for instance, is valid and perfectly adequate to measure phylogenetic diversity, which is especially relevant in evolutionary biology. Shannon index, by quantifying diversity as a function of species abundance and evenness, is adequate to measure community diversity, which is relevant in ecology.

Since there is no universally agreed purpose for biodiversity conservation (Maclaurin & Sterelny, 2008b), and thus for the general concept of biodiversity (Burch-Brown & Archer, 2017), no surprise that there is no measurement that is perfectly adequate to capture it according to the skeptics. Therefore, while we can trust multiple measurements as adequate for specific purposes, we have reasons to distrust the general concept. To bring some examples, an eliminativist like Santana would argue that the purpose of the biodiversity concept is to identify and maintain biological values. But since biological values are adequately measured by a plurality of non-overlapping techniques, the general concept of biodiversity is not actually being quantified and does not serve any purpose related to the identification of biological values. Deflationists and revisionists, on the other hand, agree that if the general concept of biodiversity is properly defined, it does serve a purpose, typically identified as global conservation. Thus, valid biodiversity measures are those that advance these goals. Although there may be disagreements over the specific purposes of global conservation—whether to maintain specific ecosystem services (Mace et al., 2012), anticipate any unexpected future benefits (Faith, 2021), or preserve variety (C. Lean & Maclaurin, 2016)—deflationists and revisionists concur that the concept of biodiversity should be shaped to serve these conservation purposes effectively.

To review this first assumption, I argued that the skeptical argument commits to an ontological stance, namely that measurement results should be caused by the features of an existing measurand captured by a scientific concept. The linguistic function of a general concept refers to this existing measurand, and a measurement is considered valid if its representation appropriately serves the purpose that prompted the measurement procedure. According to the skeptic, "biodiversity" fails on all these fronts. Thus, the general concept of biodiversity is doomed or requires enhancement, as it is certainly not useful in its present form.

Assumption 2: **Defining Biodiversity Precedes its Measurement**

The second key assumption in biodiversity skepticism, I argue, emerges in the unstated link between the characterization of a scientific or theoretical concept and the statistical and empirical machinery used to measure and quantify it. It is generally agreed that the appropriate characterization of a measurand helps delimit the measurement target significantly and, in addition, helps reduce some types of systematic error thus improving measurement accuracy. Even disciplines more interested in physical magnitude emphasizes how crucial accurate definitions of a *measurand* are. In metrology, the science of measurement themselves, a perfectly conclusive definition for theoretical concepts is not considered to be ultimately attainable—not even for concepts that are basic to scientific practice, like length and time. Some degree of "definitional uncertainty", will always accompany measurement in science.[7] Even if not all details of a measurand can be listed (Tal, 2016), producing appropriate definitions for concepts meant to be quantified must be encouraged, and appropriate definitions, even if they only approximate the concept they are meant to limit, will improve precision and decrease measurement uncertainty. The *directionality* between definitional and measurement practices is at stake in the skeptical argument. I contend that, for the skeptical move to make sense, the skeptic must be implicitly committed to a linear model of how valid measurement practices unfold, namely by being predetermined by a well-defined, fixed concept.

In the context of measuring biodiversity, it is generally agreed that defining "biodiversity" is not an end in itself: a concept's characterization significantly determines how accurate and precise its measurement can be, and influences research and policy. In their landmark volume *Biological Diversity. Frontiers in Measurement and Assessment*, Anne E. Magurran and Brian J. McGill say that:

> "[...] Biological diversity is a multifaceted concept that can be defined and documented in different ways. Being clear about exactly what we mean by biological diversity (or biodiversity)

[7] "Definitional uncertainty" is also known as "intrinsic uncertainty" according to the Guide to the expression of uncertainty in measurement (GUM), entry D.3.4) is the "measurement uncertainty resulting from the finite amount of detail in the definition of a measurand" (Draft of International Vocabulary of Metrology (VIM, 4th edition) p. 38).

> is *the first step towards measuring it*. Even then the user can be confronted by a myriad measures, some of which will do a better job than others." (Magurran & McGill, 2011, p.1, emphasis added)

Delimiting and characterizing the scope of your object of interest is a fundamental step to measuring it. This increases confidence that the right object is being targeted. Generally, scientists agree that we must be able to quantify and measure biodiversity to treat it scientifically. Vague or polysemous concepts associated with measurement targets are more problematic to quantify and measure.[8] Characterizing biodiversity properly, therefore, determines measurement success.

The skeptical argument relies on a top-down interaction between definition of a scientific concept and measurement procedures. For the skeptic, scientists and policymakers should settle on the characterization of a concept before they can develop measurement tools, this is why conceptual clarity around "biodiversity" is critical. Definitions take precedence over practice, following a logico-chronological path in the quantification of a study object.

According to the skeptical line of argument, messiness in measurements is a symptom that the top-down approach is not carried out properly. The concept of biodiversity should clearly and consistently guide measurement practices, but the lack of coherence among measurements suggests that the underlying concept of biodiversity is not sufficiently precise and unified to anchor these varied measurement tools. As we have already seen, biodiversity skepticism is not about the difficulty of measuring the various dimensions of biodiversity, but about the poor guidance the concept itself provides to science.

Sub-Assumption 2.1: **Concepts Provide Stable Theoretical Frameworks**

One of the corollaries to this view that definitions precede measurements, and good definitions guide consistent and reliable measurements is that a theoretical concept should provide a stable and uniform framework for mea-

[8]Consider concepts such as *absolute time* (Bridgman, 1927), *well-being* (Alexandrova, 2017a; Hausman, 2015), *learning*, *evolvability* (Milot, Béchet, & Maris, 2020; Mace & Purvis, 2008) and even *climate* (Frigg, Thompson, & Werndl, 2015). Their measurability has long been challenged by their vagueness or polysemy.

surement practices, ensuring consistency and functioning as a standard for correct measurement extension into new domains. A well-defined concept provides the necessary structure for developing precise and reliable measurement tools. In other words, a concept *epistemically anchors* a system of practices and serves as a benchmark for the correct use of the concept. For example, in biology, the concept of species provides a minimal stable framework that underpins various measurement practices and ensures consistency across domains. Inversely, when a scientific concept is abandoned, it indicates that a series of metaphysical, epistemic, and methodological stances are no longer acceptable. For instance, the concept of the ether in physics and the four humors framework in medicine were discarded, leading to the rejection of theoretical systems that relied on it.

This perspective logically aligns with an underlying hypothesis about the nature of scientific research, emphasizing that proper measurement is inherently tied to well-defined concepts. Without such a framework, measurements can become inconsistent. This is the case with "biodiversity", according to the skeptic. No unique theoretical framework or methods is encapsulated in the biodiversity concept. Various frameworks are accommodated by the concept as it is used in different fields— conservation, ecology and crop science. Failing to provide a stable theoretical framework, the concept does not effectively guide coherent measurements, and its utility is called into question.

Sub-Assumption 2.2: **Expectation of Homogeneity**

As we saw earlier, causality embedded in reflective measurement models entails the expectation of a unique consistent approach to measurement. There is an epistemic counterpart to this expectation entailed by assumption 2. Conceptual priority over practice strengthens the expectation that measurements associated with a single concept should be homogeneous or, if multiple, they should covary.

For example, in physics, temperature is consistently measured using different scales like Celsius, Fahrenheit, and Kelvin. Despite the different units, these measurements covary and maintain consistency, thereby reinforcing the scientific validity of the concept of temperature. Concepts in psychology, economics, anthropology, are much harder to measure and different scales or tests sometimes display opposing values (Alexandrova, 2017b, Merry, 2009).

That is one of the reasons why the debate around social science concepts is much more inflamed and hard to finalize. What is certain is that measurement homogeneity, when unreached, indicates that something is wrong with a concept in the first place. Similarly, in biodiversity, consistency or covariance failure among biodiversity measurements can be explained by a flaw in the concept of biodiversity itself.

Sub-Assumption 2.3: **Fruitfulness of a Concept**

A crucial epistemic consideration derived from 3 concerns a concept's scientific fruitfulness, which grants its retention. Biodiversity skeptics adhere to the view that for "biodiversity" to be truly useful, it would be involved in two of the most important scientific activities that rely on robust measurements: predictions and explanations. Forecasting trends and discovering causes are essential for (global or local) conservation, ecology, and crop science. Again, this is not an unusual view, as many scientific concepts have been assessed using prediction and explanation as main criteria. For instance, this has been the case with the concept of "cancer". As a general notion, "cancer" is controversial due to its various operationalizations and measurements (especially qualitative classifications), but it is still considered overall useful as it helps anticipate health decline (Plutynski, 2018).

The skeptic implies that the general biodiversity concept cannot play a similar role—only quantifying its various dimensions can. Even worse, retaining "biodiversity" as a scientific notion could be counterproductive insofar as modeling one aspect of diversity might mislead us into thinking that other aspects follow similar trends. Since several aspects of biodiversity are not covariant, this is a risky inference. Therefore, the biodiversity concept, being as contentious as the idea of cancer, is not similarly validated by the overall helpful role it plays in predictions and explanations, thus rendering it unusable.

Let me recap this second assumption. I argued that biodiversity skepticism is underpinned by the epistemic desideratum that definitions should precede and guide measurement practices. This assumption entails that concepts should provide a stable theoretical framework for the development of measurement practices, and adds up to the expectation that measurements should be homogeneous. Finally, a well-defined, fruitful concept can also be assessed from its contribution to prediction and explanation. By failing to meet these desiderata, the skeptics argue that "biodiversity" cannot guide science, and should be distrusted.

In conclusion, the skeptical stance on biodiversity is based on two key assumptions and several corollary stances that explain why measurement proliferation should make us doubt the scientific validity of a concept. Eliminativists, deflationists, and revisionist accounts, by doubting the utility of the biodiversity concept upon measurement proliferation, implicitly rely on some or all of these assumptions.[9] In the next section, I offer a way out of this inferential move by challenging the two assumptions. If the general notion of biodiversity has to be problematized, it cannot be because of the proliferation and heterogeneity of its measurements.

4 A way out of skepticism

The section above described the conceptual scaffolding that needs to be in place for the skeptical move to be reasonable and sound. Skeptics argue that diversity and heterogeneity in biodiversity measurements signal a lack of an underlying object that causes the measurements (Assumption1) and the concept's impairment to guide science (Assumption2). This section challenges the conceptual scaffolding by arguing that most of the intuitions supporting biodiversity skepticism are inconsistent with a more nuanced understanding of measurements and do not capture the interplay between measurement practices and concept formation. The argumentation here is inspired by measurement discourse in the physical and social sciences, which could be extended to the case of biodiversity. I here criticize the interplay between measurements and concepts in the skeptical argument, ultimately suggesting the need to rethink eliminativism, deflationism, or revisionism as solutions to measurement proliferation.

Against Assumption 1: **Reflective or Formative Measurement Model?**

Assumption 1 is about biodiversity being measured following a reflective model, where a measurand is said to cause or being described by the recorded

[9]I believe that many of the eliminativists, deflationists, and revisionists whose work I discussed above might openly oppose the two premises and related assumptions if explicitly presented with them in this paper.

variables. In this view, the observed variables are expected to covary and reflect different correlated facets or dimensions of the same construct. However, this is not the only way measurement theorists have conceptualized the relationship between a study object and its measurements. Psychometric and econometric studies, also faced with a plurality of often heterogeneous and inconsistent measurements, often rely on a *formative* measurement model (Borsboom, 2005). It is a fairly standard view in social science that by adopting a formative measurement model, one commits to the view according to which a measurand *results* from multiple observed variables, instead of being their cause. The essential difference between the two views is the direction of "causality": while reflective approaches see the measurand causing the measurement results (or observations), the formative ones see the measurement results as constituting the object of measurement (Hardin, Chang, Fuller, & Torkzadeh, 2011, Coltman et al., 2008). The measurand is a combination of all these measures.

Within a formative approach, the various measurements are conceptualized as independent from one another, and they are not fixed, pre-established or selected based on the coherence they have with one another. Validation of these measurement is not based on how close they capture a pre-existent measurand, but on other contextual considerations, which can be empirical adequacy in a context. The fact that certain aspects are being used as measures is thus assessed contextually, when the specific circumstance requires certain practices to be favored, and there is no restriction to changing or adding more variables. Consequently, proliferation and heterogeneity, under a formative model, could signal a more subtle and detailed understanding of the resulting measurand. Since the construct is a conglomerate of the various measurements, more measurements could actually make the concept more detailed. Proliferation and heterogeneity simply emerge from the specificities of a problem space, the epistemic and non-epistemic circumstances of research, and the iterative work of biodiversity scientists, factors that contribute to the general concept.

As an empirical consequence of the formative power that measurements exercise on the measurand, this model does not posit a necessary covariance among the various measurements. Changes in the observed variables can lead to changes in the latent measurand. As far as the concept that is being measured goes, the semantic space associated with the concept will likely change depending on which indicators are used and which ones are dropped.

When making inferences about biodiversity *per se* from the plurality of measurements, the skeptic entails that the lack of consistency signals a failure of causality, which in turn suggests that there is no "entity" called biodiversity causing the measurement. But if one adopts a formative measurement model, namely if one thinks that biodiversity does not causally determine its various measurements but rather emerges from them, the skeptical worry simply disappears. There are some good reasons why we might want to think of measuring biodiversity according to a formative measurement model that goes beyond the *ad hoc* move to criticize the skeptic. Sets of formal criteria have been advanced to decide whether a reflective or formative models are more appropriate depending on the alleged nature of the study object, the analytical tools at one's disposal, and other formal desiderata.[10] In social science and healthcare, scholars often interpret measuring according to a formative model, as they help make sense of measurement practices around variables that are associated with multidimensional concepts such as socioeconomic status (Antonoplis, 2023) or quality of life (Alexandrova, 2017b). The concept of biodiversity, due to its normative connotation and complexity, parallels some of the notions for which a formative measurement model is often accepted. Another reason to resort to a formative model is to make sense of the sparse biodiversity data that seem to point in different dimensions, arguing that biodiversity is whatever construct emerges from a joint assessment of all these dimensions (how this might be done analytically is out of the scope of this paper).

I do not here mean to defend the adoption of a formative approach to measuring biodiversity, even though the arguments above might suggest so. My main point is to show an alternative to the unstated reflective model. A skeptical argument should not assume a reflective model but should argue for it. This model is not so obvious, as it is not so obvious that the variables usually associated with biodiversity must be caused by an underlying phenomenon.

Against the Realist Expectation. Entertaining the possibility of a formative model also shakes the corollaries to Assumption 1. Consider the scope of a realist expectation under a formative model: the reality of biodiversity as a measurand diverges from its existence pre-measurement. Biodiversity might be understood as an emergent phenomenon or property formed

[10]For instance, Coltman et al. (2008) offer theoretical and empirical criteria to decide whether, in specific measurement context, it is better to adopt a reflective or a formative approach. They argue that a formative measurement model best describes the quantification of concepts such as international business pressure and market orientation.

by a combination of its measurements, regardless of whether they covary.

This approach aligns with psychometric theory, which recognizes that a construct can emerge from the combinations of various measurements, each of which, individually or as a cluster, serves a specific purpose.For instance, in econometrics, socioeconomic status (SES) is often measured formatively by combining measurements of income, education, and financial security and individual perception (American Psychological Association). These indicators may not necessarily covary, but, when summed up, they merge into a very useful concept which can be mathematically rendered with a unique index. Likewise, biodiversity could emerge as an amalgamation of diverse measurements, such as species richness, genetic diversity, and ecosystem function. Despite being conceptually distinct and differently measured, all these aspects will contribute to the formation of the measurand. Thus, considering biodiversity under a formative approach would shift the focus from seeking measurement coherence as a sign of the preexistence of a measurand, to recognizing the construct as an outcome of its measurements. A formative model clearly aligns with ontological pluralism, which acknowledges that multiple ways of categorizing the world can be equally valid, even if they do not converge into a unified ontology.

A realist might complain that interpreting biodiversity as an emergent property based on the selection of different features in specific contexts challenges the desideratum that biodiversity be a natural kind or property, making it an entirely artificial construct whose existence is purely mind-dependent. This is not a disruptive criticism. Ontological pluralism is not virtually inconsistent with a minimal form of realism (for instance, "promiscuous realism" Dupré, 1993). Instead of committing to the existence of entities entirely independent of our minds, a minimal form of realism accepts that our scientific concepts can be both empirically grounded and determined by contingent factors such as values, interests, and context (Ludwig & Ruphy, 2021). If the commitment to biodiversity being a natural kind is still a major worry, the skeptic should acknowledge that, very often, the only condition for a concept to reflect a natural kind or property is that some of its dimensions can be clustered together and represent meaningful correlations in specific research contexts.[11] For instance, in psychometrics, constructs like intelligence or well-being are considered real in the sense some of their measurements can be effectively clustered to explain and predict certain policy-relevant patterns

[11]See also Bokulich (2014) for similar observations in the context of planetary science.

according to specific purposes (Serpico, 2018, Alexandrova, 2017b).

I am here neither advocating for the need to maintain some form of realism to tame measurement plurality, nor that biodiversity should be understood as a natural kind or property (as in Santana, 2018 and C. H. Lean, 2017). My goal here is to show that endorsing a formative model compromises the tenability of biodiversity skepticism. In principle, a formative model could be consistent with a realist attitude and a reworked account of biodiversity as a natural kind—all without the need to doubt the biodiversity concept.

Against a purely descriptive function of measurements. In Section 3, I linked the skeptical move to the descriptive function of measurements. However, under a formative model, measurement acquires a distinctive *constructive* function. Were biodiversity understood according to a formative approach, it would not be exclusively described using measurement, it would be made.

The constructive role of measurements in shaping study objects instead of just mirroring them is widely accepted in social sciences. Anthropologist Sally Engle Merry (2009) thoroughly documented how measurements developed to quantify Human Rights, Gender Violence, and Sex Trafficking by international organizations such as UNESCO ended up shaping how the phenomena are conceptualized today. Different measurement procedures stem from diverse ethical and pragmatic commitments, and remain a major tool for policy. Similarly, Ian Hacking (1999) documented the interplay of measurements and conceptual development in child abuse. These accounts do not endorse radical constructivist metaphysics, they are aware of the subtle iterativity between theoretical demands and empirical restraints when developing and validating measurements for a socially relevant concept. This is to say that while maintaining that measurement practices must be empirically grounded, they allow for theoretical consideration, such as concepts' definition, to be shaped by empirical practices.

The constructive function of measurements has even been acknowledged when measuring fundamental physical magnitudes. Hasok Chang (2005) has one of the most acclaimed analyses of how quantification—and more generally, measurement—alters the temperature concept from a purely qualitative bodily sensation to a sophisticated instrumental procedure. Similarly with length. Traditionally, the concept of "length" was applied to measure distances between meso-objects. However, as scientific exploration expanded into astronomy and theoretical physics, researchers sought to use the concept of length to express spatial relationships between astronomic objects or physical particles. This extension of the concept into new domains introduced friction but also drove scientific progress (Chang, 2005). In this process, "length" expanded its meaning, playing a crucial role in a new context. Importantly, this extension did not render the concept meaningless or empty; instead, it enriched its significance within the new domain. From Chang's analysis, I take the insight that even physical-magnitude concepts, rather than being a static precondition for measurement practices, are partially determined via measurement processes.

It is plausible to think of "biodiversity" in a comparable manner, as a general concept being shaped by the measurement practices developed as the concept adapts to novel research contexts. As biodiversity science explores different ecological contexts and theoretic conservation goals, it adapts and expands to accommodate the measurements developed for specific tasks and challenges posed by each domain.

Against Reference Failure. Reference failure is a corollary of the skeptical move that also needs to be rethought in light of the possibility of a formative model. "Biodiversity" does not need to be ontologically rooted in a mind-independent object to have referential power, once it is understood according to a formative model. Reference need not be ontologically prior to the measurement practices or a fit for a realist stance. The reference of "biodiversity", therefore, could be understood extensionally, as a list of all the actual and potential applications of the context in scientific practices.

Defining a concept extensionally often attracts criticism that by adding or removing an item from the list, the entire concept changes. However, I will argue that this is not a problem. In fact, scientific concepts frequently fluctuate in response to their usage or discontinuation.

Against the One-Purpose View. One of the unstated underpinnings of the skeptical view is the adoption of an adequacy for purpose view of measurements. The skeptic is right in endorsing an adequacy for purpose view of measurement, but their interpretation of what constitutes a purpose is too limited, especially within a formative approach. If we consider a formative approach to measuring biodiversity, the overarching purpose attributed to the general concept should be rethought, and the inference that the general concept of biodiversity lacks any specific purpose should be dismissed.

The skeptic should not be restricted to thinking that there must be a singular, superior purpose for the biodiversity concept. In analogy, think of the concept of health. Health is adequately measured in many ways (BMI or cholesterol levels are two ways) each serving a different purpose (such as stroke prevention or fitness monitoring). The multiplicity of measurements is no good reason for rejecting the general concept. While it is true that "biodiversity" was developed within the context of conservation, and one might argue that conservation has a special privilege in determining the fate of the concept, this view is too narrow. Conservation itself is fragmented, with various approaches such as preservation and restoration. The former prioritizes existing ecosystems, and often uses species richness metrics as best serving its purposes. The latter aims to recover degraded ecosystems and adopts ecosystem function measures to serve its purposes. These approaches stretch the purpose of the biodiversity concept and influence the most adequate candidates for its measurement. Fields like ecology and crop science should also have a say in defining the overarching purpose of "biodiversity" as they provide the basic theoretical scaffolding and empirical evidence that underpin conservation science. Within a formative framework, the concept of biodiversity can better accommodate these diverse concerns, as it acknowledges that the general concept emerges from integrating various measurements and purposes across different fields.

In conclusion, while I agree with skeptics that measurements should be assessed based on their adequacy for a specific purpose, it is not trivial to determine which of the many purposes should be prioritized as the benchmark for validating measurements. The concept serves a variety of purposes, and the diversity of measurements developed for those purposes should not diminish the concept's value.

Against Assumption 2: **Concepts Change in Response to Measurement Practices**

Assumption 2 was about the priority of definitional over measurement practices. This assumption conflicts with our best understanding of how measurement practices unfold and shape the theoretical landscape a concept occupies. Philosophers have long pointed out how, frequently, measuring is not a simple procedure of "Go ye forth and measure," (Kuhn, 1961, p.185). Measuring requires tunings and feedback loops among a measurand's various operationalizations, mathematical representations, and data practices (Bradburn, Cartwright, & Fuller, 2016). Measurement is a structured, iterative activity (Chang, 2005) that constantly shapes itself in the face of conceptual, empirical desiderata, theoretical constraints, and available instrumentation, and political and social considerations. Measurement practices are ever-changing due to the alignment of measurement tools to specific problem spaces and the expansion of biodiversity research to new domains. This is not neutral with respect to the development and change of a scientific concept.

Scientific concepts frequently change in response to new empirical tools and new theoretical frameworks (think of "gene", "homology", "temperature", and also Brigandt, 2020). "Biodiversity" is no different to be granted an exceptional treatment. The notion has already changed. Recent scholarship has insistently argued that diversity research should include measurements of socio-cultural diversity (Bocchi, 2024), which would significantly expand CBD's characterization. This proves biodiversity is a dynamic concept subject to re-adjustments and refinements in response of empirical and theoretical considerations, including new measurement practices.

The skeptic should abandon an understanding of the relationship between theoretical concepts and measurement practices as unidirectional, where the measurement practices should align and be guided by a theoretical concept preliminary clarified. A theoretical concept is not only a driver of research: it also expands and evolves as a result of the measurement procedures devised to fit various research scenarios.

Against the Skeptic view on concept utility. Finally, a concept's scientific utility is not undermined by the proliferation and heterogeneity of its measurements. For a skeptic, fruitfulness or utility boils down to prediction and explanation, when priority is given to the quantitative aspect of a concept. There are alternatives for assigning importance to scientific theoretical concepts beyond their roles in predictions and explanations—even though these activities are of utmost importance.

A convincing alternative to understanding the utility of complex concepts in science in relation to measurement practices has been offered by Novick and Haueis (2023). They interpret the role of theoretical concepts, especially complex concepts, in scientific research under the lenses of "patchwork theory" (Haueis, 2021). According to a patchwork approach, when a complex concept is utilized within a specific domain, it is adapted to a particular level of detail or scale. In addition, within this domain, specific techniques are employed to explore the distinct properties associated with the concept. Complex concepts are examined in different domains, each representing a distinct area, and these areas have both tangible and conceptual connections, but they are not identical, including their measures. As Novick and Haueis (2023) clarify, "[a]s concepts are extended to new domains, they naturally settle into local patches of use; this settling occurs under the auspices of a *reasoning strategy*"(2023, p.10, italics added). Reasoning strategies bind patches of concept use and confer unity to a theoretical concept, as well as reveal the utility of a concept as that of favoring the extension to new domains, doing so "while still respecting the operationalist insight that measurement techniques are partially constitutive of the meanings of scientific concepts" (2023, p.10-1). Additional work is required to determine if this approach captures existing measurement practices in biodiversity research. Nonetheless, it undeniably provides us with a valuable alternative for evaluating the utility of a concept beyond its effectiveness in prediction and explanation. By recognizing shared principles across domains, scientists can build on each other's work, hopefully leading toward better actionable knowledge.

By problematizing the assumptions underlying the skeptical argument, I hope that I have shown why the proliferation and heterogeneity of measurements do not necessarily render a concept useless. Eliminativist, deflationist, and revisionist solutions to the skeptical argument that rely on these assumptions should be reassessed. There might be other arguments for an eliminativist, deflationist, or revisionist strategy toward the concept of biodiversity—one of which I sketch in the conclusion.

5 Conclusion

The purpose of this paper was to contribute to the discussion on the interplay between the theoretical concept of biodiversity and the proliferation and heterogeneity of its measurements. I have shown that several eliminativist, deflationist, and revisionist accounts share an unwarranted inferential move. This move is based on the implicit assumption that biodiversity should be measured using a reflective approach and that its definition should precede and guide its measurement. By analyzing and criticizing these assumptions, I hope to have revamped a debate that has stagnated for a while. To further advance the discussion, we should compare the discourse on measuring biodiversity with the sophisticated debates on measurement and conceptual change occurring in other subfields of philosophy.

Let me disclose my sympathy for biodiversity revisionism. I believe that it is not the plurality of measurement practices but the value-laden nature of the biodiversity concept and its sociopolitical implications that warrant its revision. By "value-laden nature" I mean how the concept is co-determined by non-epistemic factors such as axiological theories on nature, scientific traditions, and worldviews. By sociopolitical impact, I want to stress the normative dimension of a concept in shaping policy action. When environmental policies ignore the normative and non-epistemic aspects included in the characterization of a concept, they risk implementing harmful and unjust strategies, such as evicting people or suppressing bio-cultural traditions. A revisionist approach should unpack the non-epistemic semantic aspects of a concept, looking for tools to assess which values are acceptable, to be able to weigh trade-offs, and to anticipate the consequences of specific characterizations.

To this end, philosophers of science hoping to revise "biodiversity" will benefit greatly from engaging with a subfield of philosophy that, although rarely dealing with quantifiable concepts, has developed tools for handling non-epistemic issues in concept formulation: conceptual engineering. Conceptual engineering provides methods for assessing the epistemic and non-epistemic values in characterizing a concept, how to negotiate tradeoffs, and even for determining whether the conversation has changed focus (Brigandt & Rosario, 2019). Biodiversity deserves such treatment given its sociopolitical dimension, including its role in the promotion of policies that can have adverse social effects (Bocchi, 2024). One can be a revisionist without being a skeptic, and adopting a formative approach to measurement is a good starting point for rethinking biodiversity as a construct emerging from measurements designed to meet both epistemic and non-epistemic purposes. Revisionism that only focuses on measurement plurality and does not prioritize non-epistemic considerations will fail to amend the concept effectively and miss the full potential of a revised biodiversity concept.

I cannot develop an exhaustive ameliorative account here. Therefore, I conclude with a reflection on how the special case of biodiversity presented in this paper can contribute to the advancement of general philosophy of science. I believe that the philosophy of measurement and conceptual engineering literature should engage in an open dialogue, as they each possess tools that can advance the other's agenda. First, I urge scholars working in philosophy of measurement to assimilate the work done in conceptual engineering. It is now commonly agreed that measurement practices are value-laden. Conceptual engineering can offer normative guidance on assessing the legitimacy of values, interests, and other non-epistemic considerations when validating the adequacy of certain measurements. In the case of biodiversity, if a revisionist approach is warranted by non-epistemic considerations and the impact of the biodiversity concept, conceptual engineering can assist in the assessment of which measurements should be included and the viability of an ameliorative project.

Second, I urge conceptual engineers to exploit the literature on measurement, especially considering how measurement practices can be utilized in revisionist, deflationist, or eliminativist projects. The notion of a formative approach to measurement is particularly relevant to conceptual engineering, as this approach allows the meaning of a concept to emerge from context-dependent measurement tailored to solve specific problems. Conceptual engineers almost exclusively focus on ethically-loaded terms, overlooking traditionally more quantitative concepts, perhaps under the assumption that measurement practices are purely empirical and less related to concept formation. However, as Chang (2004)'s historico-philosophical analysis of the concepts of length and temperature demonstrates, even physical magnitudes are subject to change and influenced by non-epistemic considerations. Therefore, purely quantitative concepts should not escape the analysis of conceptual engineers.

To conclude, I hope I have conveyed the message that the interplay of measurement proliferation and the utility of concepts transcends the debate around "biodiversity." There is potential for collaboration between philosophy of measurement and conceptual engineering, particularly in validating measurements and advancing the epistemic and non-epistemic goals of science.

References

Alexandrova, A. (2017a). *Is Well-Being Measurable?* (Vol. 1). Oxford University Press. Retrieved 2021-12-16, from https://oxford.universitypressscholarship.com/view/10.1093/oso/9780199300518.001.0001/oso-9780199300518-chapter-5 doi: 10.1093/oso/9780199300518.003.0005

Alexandrova, A. (2017b). *A Philosophy for the Science of Well-Being* (Vol. 1). Oxford University Press. Retrieved 2021-12-16, from https://oxford.universitypressscholarship.com/view/10.1093/oso/9780199300518.001.0001/oso-9780199300518 doi: 10.1093/oso/9780199300518.001.0001

Antonoplis, S. (2023). Studying socioeconomic status: Conceptual problems and an alternative path forward. *Perspectives on Psychological Science*, *18*(2), 275–292.

Appiah, K. A. (2006). How to decide if races exist. In *Proceedings of the aristotelian society* (Vol. 106, pp. 363–380).

Bocchi, F. (2024). Metrics in biodiversity conservation and the value-free ideal. *Synthese*, *203*(5), 145.

Bokulich, A. (2014). Pluto and the 'planet problem': folk concepts and natural kinds in astronomy. *Perspectives on Science*, *22*(4), 464–490.

Bokulich, A., & Parker, W. (2021, March). Data models, representation and adequacy-for-purpose. *European Journal for Philosophy of Science*, *11*(1), 31. Retrieved 2023-01-17, from http://link.springer.com/10.1007/s13194-020-00345-2 doi: 10.1007/s13194-020-00345-2

Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics.* Cambridge University Press.

Borsboom, D., Mellenbergh, G., & Heerden, J. (2003, April). The Theoretical Status of Latent Variables. *Psychological review*, *110*, 203–19. doi: 10.1037/0033-295X.110.2.203

Bradburn, N. M., Cartwright, N., & Fuller, J. (2016). A Theory of Measurement. , 1–20.

Bridgman, P. (1927). *The logic of modern physics.* (MacMillan ed.).

Brigandt, I. (2020). How are biology concepts used and transformed. *Philosophy of science for biologists*, *8*, 79–101.

Brigandt, I., & Rosario, E. (2019). Strategic conceptual engineering for epistemic and social aims.

Burch-Brown, J., & Archer, A. (2017, December). In defence of biodiversity. *Biology & Philosophy*, *32*(6), 969–997. Retrieved 2021-10-24, from http://link.springer.com/10.1007/s10539-017-9587-x doi: 10.1007/s10539-017-9587-x

Casetta, E., Silva, J. M. d., & La Vecchia, D. (Eds.). (2019). *From assessing to conserving biodiversity: conceptual and practical challenges* (No. Volume 24). Cham: SpringerOpen.

Chang, H. (2004). *Inventing temperature: Measurement and scientific progress.* Oxford University Press.

Chang, H. (2005). Inventing temperature: measurement and scientific progress. *Choice Reviews Online*, *43*(01), 43–0373–43–0373. Retrieved 2021-10-24, from http://choicereviews.org/review/10.5860/CHOICE.43-0373 doi: 10.5860/CHOICE.43-0373

Coltman, T., Devinney, T. M., Midgley, D. F., & Venaik, S. (2008). Formative versus reflective measurement models: Two applications of formative measurement. *Journal of Business research*, *61*(12), 1250–1262.

Conway, A. R., & Kovacs, K. (2015). New and emerging models of human intelligence. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*(5), 419–426.

Corns, J. (2016). Pain eliminativism: scientific and traditional. *Synthese*, *193*(9), 2949–2971.

Crupi, V., Nelson, J. D., Meder, B., Cevolani, G., & Tentori, K. (2018, July). Generalized Information Theory Meets Human Cognition: Introducing a Unified Framework to Model Uncertainty and Information Search. *Cognitive Science*, *42*(5), 1410–1456. Retrieved 2021-10-13, from https://onlinelibrary.wiley.com/doi/10.1111/cogs.12613 doi: 10.1111/cogs.12613

Daly, A., Baetens, J., & De Baets, B. (2018, July). Ecological Diversity: Measuring the Unmeasurable. *Mathematics*, *6*(7), 119. Retrieved 2021-10-13, from http://www.mdpi.com/2227-7390/6/7/119 doi: 10.3390/math6070119

Díaz, S. (2019). Why care about nature? a pluralistic agenda for biodiversity. *Biodiversity Revisited is an initiative of the Luc Hoffmann Institute, in collaboration with WWF, Fu-ture Earth, ETH Zürich Department of Environmental Systems Science, University of Cambridge Conservation Research Institute, and the Centre for Biodiversity and Environment Research at University College London.*, 62.

Duelli, P., & Obrist, M. K. (2003, September). Biodiversity indicators: the choice of values and measures. *Agriculture, Ecosystems & Environment*, *98*(1-3), 87–98. Retrieved 2022-09-22, from https://linkinghub.elsevier.com/retrieve/pii/S0167880903000720 doi: 10.1016/S0167-8809(03)00072-0

Dupré, J. (1993). *The disorder of things: Metaphysical foundations of the disunity of science*. Harvard University Press.

Ereshefsky, M. (1998). Species pluralism and anti-realism. *Philosophy of Science*, *65*(1), 103–120.

Faith, D. P. (1992). Conservation evaluation and phylogenetic diversity. *Biological Conservation*, *61*(1), 1–10. Retrieved 2021-10-13, from https://linkinghub.elsevier.com/retrieve/pii/0006320792912013 doi: 10.1016/0006-3207(92)91201-3

Faith, D. P. (2018). Avoiding paradigm drifts in ipbes: reconciling" nature's contributions to people," biodiversity, and ecosystem services. *Ecology & Society*, *23*(2).

Faith, D. P. (2021). Valuation and appreciation of biodiversity: The "maintenance of options" provided by the variety of life. *Frontiers in Ecology and Evolution*, *9*, 635670.

Fan, J.-x., Shen, S.-z., Erwin, D. H., Sadler, P. M., MacLeod, N., Cheng, Q.-m., ... others (2020). A high-resolution summary of cambrian to early triassic marine invertebrate biodiversity. *Science*, *367*(6475), 272–277.

Fischer, A., & Young, J. C. (2007, April). Understanding mental constructs of biodiversity: Implications for biodiversity management and conservation. *Biological Conservation*, *136*(2), 271–282. Retrieved 2022-09-22, from https://linkinghub.elsevier.com/retrieve/pii/S0006320706005313 doi: 10.1016/j.biocon.2006.11.024

Frigg, R., Thompson, E., & Werndl, C. (2015, December). Philosophy of Climate Science Part I: Observing Climate Change: Observing Climate Change. *Philosophy Compass*, *10*(12), 953–964. Retrieved 2022-03-15, from https://onlinelibrary.wiley.com/doi/10.1111/phc3.12294 doi: 10.1111/phc3.12294

Gerard, R. W. (1965, May). Intelligence, Information, and Education: Various systems—brains, computers, and institutions—are developed to handle new information in new ways. *Science*, *148*(3671), 762–765. Retrieved 2022-09-22, from https://www.science.org/doi/10.1126/science.148.3671.762 doi: 10.1126/science.148.3671.762

Glasgow, J., Haslanger, S., Jeffers, C., & Spencer, Q. (2019). *What is race?: Four philosophical views*. Oxford University Press.

Hacking, I. (1999). *The social construction of what?* Harvard university press.

Hardin, A. M., Chang, J. C.-J., Fuller, M. A., & Torkzadeh, G. (2011). Formative measurement and academic research: In search of measurement theory. *Educational and Psychological Measurement*, *71*(2), 281–305.

Haueis, P. (2021). A generalized patchwork approach to scientific concepts.

Hausman, D. M. (2015). *Valuing health: well-being, freedom, and suffering*. Retrieved 2022-03-15, from http://public.eblib.com/choice/publicfullrecord.aspx?p=3056472 (OCLC: 960447252)

Heberling, J. M., Miller, J. T., Noesgaard, D., Weingart, S. B., & Schigel, D. (2021, February). Data integration enables global biodiversity synthesis. *Proceedings of the National Academy of Sciences*, *118*(6), e2018093118. Retrieved 2023-03-28, from https://pnas.org/doi/full/10.1073/pnas.2018093118 doi: 10.1073/pnas.2018093118

Hurlbert, S. H. (1971, July). The Nonconcept of Species Diversity: A Critique and Alternative Parameters. *Ecology*, *52*(4), 577–586. Retrieved 2021-12-05, from http://doi.wiley.com/10.2307/1934145 doi: 10.2307/1934145

Justus, J. (2011). A Case Study in Concept Determination: Ecological Diversity. In *Handbook of The Philosophy of Science: Philosophy of Ecology* (pp. 147–168). San Diego: North Holland: Elsevier.

Kelling, S., Hochachka, W. M., Fink, D., Riedewald, M., Caruana, R., Ballard, G., & Hooker, G. (2009). Data-intensive science: a new paradigm for biodiversity studies. *BioScience*, *59*(7), 613–620.

Kuhn, T. S. (1961). The function of measurement in modern physical science. *Isis*, *52*(2), 161–193.

Lautenschlager, R. (1997). Biodiversity is dead. *Wildlife Society Bulletin*, 679–685.

Lean, C., & Maclaurin, J. (2016). The value of phylogenetic diversity. In *Biodiversity conservation and phylogenetic systematics* (pp. 19–37). Springer, Cham.

Lean, C. H. (2017, December). Biodiversity Realism: Preserving the tree of life. *Biology & Philosophy*, *32*(6), 1083–1103. Retrieved 2021-10-24, from http://link.springer.com/10.1007/s10539-017-9592-0 doi: 10.1007/s10539-017-9592-0

Lovejoy, T. E. (1994). The quantification of biodiversity: an esoteric quest or a vital component of sustainable development? *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *345*(1311), 81–87.

Ludwig, D., & Ruphy, S. (2021). Scientific Pluralism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2021 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2021/entries/scientific-pluralism/.

Mace, G. M., Norris, K., & Fitter, A. H. (2012). Biodiversity and ecosystem services: a multilayered relationship. *Trends in ecology & evolution*, *27*(1), 19–26.

Mace, G. M., & Purvis, A. (2008, January). Evolutionary biology and practical conservation: bridging a widening gap. *Molecular Ecology*, *17*(1), 9–19. Retrieved 2021-12-02, from https://onlinelibrary.wiley.com/doi/10.1111/j.1365-294X.2007.03455.x doi: 10.1111/j.1365-294X

.2007.03455.x
Maclaurin, J., & Sterelny, K. (2008a). *What is biodiversity?* Chicago: University of Chicago Press. (OCLC: ocn156874812)
Maclaurin, J., & Sterelny, K. (2008b). What is biodiversity? In *What is biodiversity?* University of Chicago Press.
Magurran, A. E., & McGill, B. J. (Eds.). (2011). *Biological diversity: frontiers in measurement and assessment*. Oxford ; New York: Oxford University Press. (OCLC: ocn642283168)
Maier, D. S. (2013). *What's so good about biodiversity? a call for better reasoning about nature's value* (First softcover print ed.) (No. 19). Dordrecht: Springer.
Meinard, Y., Coq, S., & Schmid, B. (2019). The vagueness of "biodiversity" and its implications in conservation practice. *From Assessing to Conserving Biodiversity: Conceptual and Practical Challenges*, 353–374.
Meincke, A. S. (2022). One or two? a process view of pregnancy. *Philosophical Studies*, *179*(5), 1495–1521.
Merry, S. E. (2009). *The seductions of quantification*. Chicago university press.
Milot, E., Béchet, A., & Maris, V. (2020, July). The dimensions of evolutionary potential in biological conservation. *Evolutionary Applications*, *13*(6), 1363–1379. Retrieved 2021-12-02, from https://onlinelibrary.wiley.com/doi/10.1111/eva.12995 doi: 10.1111/eva.12995
Morgan, M. S. (2007). An analytical history of measuring practices: The case of velocities of money. *Measurement in Economics: A Handbook*, 105–132.
Muller, J. Z. (2018). *The tyranny of metrics*. Princeton: Princeton University Press.
Novick, R., & Haueis, P. (2023). Patchworks and operations. *European Journal for Philosophy of Science*, *13*(1), 15.
Ohnesorge, M. (2022). The epistemic privilege of measurement: Motivating a functionalist account. *Philosophy of Science*, 1–16.
Parker, W. S. (2020, July). Model Evaluation: An Adequacy-for-Purpose View. *Philosophy of Science*, *87*(3), 457–477. Retrieved 2023-06-13, from https://www.cambridge.org/core/product/identifier/S0031824800015956/type/journal_article doi: 10.1086/708691
Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H. G., Scholes, R. J., ... Wegmann, M. (2013, January). Es-

sential Biodiversity Variables. *Science*, *339*(6117), 277–278. Retrieved 2021-10-13, from https://www.sciencemag.org/lookup/doi/10.1126/science.1229931 doi: 10.1126/science.1229931
Pine, K. H., & Liboiron, M. (2015). The politics of measurement and action. In *Proceedings of the 33rd annual acm conference on human factors in computing systems* (pp. 3147–3156).
Plutynski, A. (2018). *Explaining cancer: Finding order in disorder*. Oxford University Press.
Porter, T. M. (1996). Trust in numbers. In *Trust in numbers*. Princeton University Press.
Reydon, T. A. (2019). Are species good units for biodiversity studies and conservation efforts? In *From assessing to conserving biodiversity* (pp. 167–193). Springer, Cham.
Rochette, A.-J., Akpona, J. D. T., Akpona, H. A., Akouehou, G. S., Kwezi, B. M., Djagoun, C. A. M. S., … Vanhove, M. P. M. (2019, March). Developing policy-relevant biodiversity indicators: lessons learnt from case studies in Africa. *Environmental Research Letters*, *14*(3), 035002. Retrieved 2022-09-22, from https://iopscience.iop.org/article/10.1088/1748-9326/aaf495 doi: 10.1088/1748-9326/aaf495
Rousseau, R. (1998). Measuring Biodiversity. , 5.
Santana, C. (2014, November). Save the planet: eliminate biodiversity. *Biology & Philosophy*, *29*(6), 761–780. Retrieved 2021-10-13, from http://link.springer.com/10.1007/s10539-014-9426-2 doi: 10.1007/s10539-014-9426-2
Santana, C. (2018, April). Biodiversity is a chimera, and chimeras aren't real. *Biology & Philosophy*, *33*(1-2), 15. Retrieved 2021-10-24, from http://link.springer.com/10.1007/s10539-018-9626-2 doi: 10.1007/s10539-018-9626-2
Sarkar, S. (2002). Defining "biodiversity"; assessing biodiversity. *The Monist*, *85*(1), 131–155.
Sarkar, S. (2005). Biodiversity and Environmental Philosophy: An Introduction. , 278.
Schmeller, D. S., Mihoub, J.-B., Bowser, A., Arvanitidis, C., Costello, M. J., Fernandez, M., … Isaac, N. J. B. (2017, November). An operational definition of essential biodiversity variables. *Biodiversity and Conservation*, *26*(12), 2967–2972. Retrieved 2021-10-24, from http://link.springer.com/10.1007/s10531-017-1386-9 doi: 10.1007/s10531-017-1386-9

Secretariat, C. (1992). Convention on biological diversity. In *Convention on biological diversity.*
Serpico, D. (2018). What kind of kind is intelligence? *Philosophical Psychology*, *31*(2), 232–252.
Serpico, D. (2021, June). The Cyclical Return of the IQ Controversy: Revisiting the Lessons of the Resolution on Genetics, Race and Intelligence. *Journal of the History of Biology*, *54*(2), 199–228. Retrieved 2021-11-05, from https://link.springer.com/10.1007/s10739-021-09637-6 doi: 10.1007/s10739-021-09637-6
Shannon, C. E. (1948, July). A Mathematical Theory of Communication. *Bell System Technical Journal*, *27*(3), 379–423. Retrieved 2022-01-29, from https://ieeexplore.ieee.org/document/6773024 doi: 10.1002/j.1538-7305.1948.tb01338.x
Tal, E. (2015). Measurement in science.
Tal, E. (2016, March). Making Time: A Study in the Epistemology of Measurement. *The British Journal for the Philosophy of Science*, *67*(1), 297–335. Retrieved 2022-03-05, from https://www.journals.uchicago.edu/doi/10.1093/bjps/axu037 doi: 10.1093/bjps/axu037
Turnhout, E., Neves, K., & De Lijster, E. (2014). 'measurementality'in biodiversity governance: knowledge, transparency, and the intergovernmental science-policy platform on biodiversity and ecosystem services (ipbes). *Environment and Planning A*, *46*(3), 581–597.
Van Bork, R., Wijsen, L. D., & Rhemtulla, M. (2017). Toward a causal interpretation of the common factor model. *Disputatio*, *9*(47), 581–601.
Vellend, M., Baeten, L., Myers-Smith, I. H., Elmendorf, S. C., Beausejour, R., Brown, C. D., ... Wipf, S. (2013, November). Global meta-analysis reveals no net change in local-scale plant biodiversity over time. *Proceedings of the National Academy of Sciences*, *110*(48), 19456–19459. Retrieved 2021-10-13, from http://www.pnas.org/cgi/doi/10.1073/pnas.1312779110 doi: 10.1073/pnas.1312779110
Watkins, A., & Di Marco, M. (MS). Sex eliminativism.
Williamson, H. F., & Leonelli, S. (2022). Accelerating agriculture: Data-intensive plant breeding and the use of genetic gain as an indicator for agricultural research and development. *Studies in History and Philosophy of Science*, *95*, 167–176.
Zack, N. (2014). *Philosophy of science and race.* Routledge.

Zhao, K. (2023). Measuring the nonexistent: Validity before measurement. *Philosophy of Science*, *90*(2), 227–244.