

Golden spikes, scientific types, and the ma(r)king of deep time

Joeri Witteveen

Department of Science Education, Section for History and Philosophy of Science, University of Copenhagen, Denmark

ARTICLE INFO

Keywords:

Geology
Stratigraphy
Metrology
Taxonomy
Stratotype
GSSP
Measurement prototype
Kilogram standard
Type specimen
Holotype
Nomenclature
Classification
Standardization
Reference fixing
Natural classification

ABSTRACT

Chronostratigraphy is the subfield of geology that studies the relative age of rock strata and that aims at producing a hierarchical classification of (global) divisions of the historical time-rock record. The ‘golden spike’ or ‘GSSP’ approach is the cornerstone of contemporary chronostratigraphic methodology. It is also perplexing. Chronostratigraphers define each global time-rock boundary extremely locally, often by driving a gold-colored pin into an exposed rock section at a particular level. Moreover, they usually avoid rock sections that show any meaningful sign of paleontological disruption or geological discontinuity: the less obvious the boundary, the better. It has been argued that we can make sense of this practice of marking boundaries by comparing the status and function of golden spikes to that of other concrete, particular reference standards from other sciences: holotypes from biological taxonomy and measurement prototypes from the metrology of weight and measures. Alisa Bokulich (2020b) has argued that these ‘scientific types’ are in an important sense one of a kind: they have a common status and function. I will argue that this picture of high-level conceptual unity is mistaken and fails to consider the diversity of aims and purposes of standardization and classification across the sciences. I develop an alternative, disunified account of scientific types that shows how differences in ontological attitudes and epistemic aims inform scientists’ choices between different kinds of scientific types. This perspective on scientific types helps to make sense of an intriguing mid-twentieth-century debate among chronostratigraphers about the very nature of their enterprise. Should chronostratigraphers conventionally *make* boundaries by designating golden spikes, or should they attempt to *mark* pre-existing ‘natural’ boundaries with the help of a different kind of scientific type?

When citing this paper, please use the full journal title *Studies in History and Philosophy of Science*.

Data availability

No data was used for the research described in the article.

1. Introduction

Charting the Earth’s history by analyzing and classifying the layers of the rock record has long been one of the central endeavors of the Earth

sciences. A centerpiece of this effort is the production of an International Geologic Time Scale (GTS): a classification of global divisions of the historical time-rock record — the *chronostratigraphic* scale — combined with estimates of their elapsed durations — the *chronometric* scale. Both elements of the study of deep time are captured in the iconic International Chronostratigraphic Chart (Fig. 1), which is updated several times per year to incorporate improved estimates of the numerical ages of the stratigraphic subdivisions. Some of these updates also introduce yellow ‘pins’ at the boundaries of stages of the Phanerozoic era that previously lacked one. A yellow pin next to a stage name indicates that a so-called Global Stratotype Section and Point (GSSP) has been formally designated to precisely mark the lower boundary of the stage at a designated location on Earth. The inauguration of a GSSP is often accompanied by

E-mail address: jw@ind.ku.dk.

<https://doi.org/10.1016/j.shpsa.2024.02.004>

Received 7 November 2023; Accepted 9 February 2024

0039-3681/© 2024 The Author. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

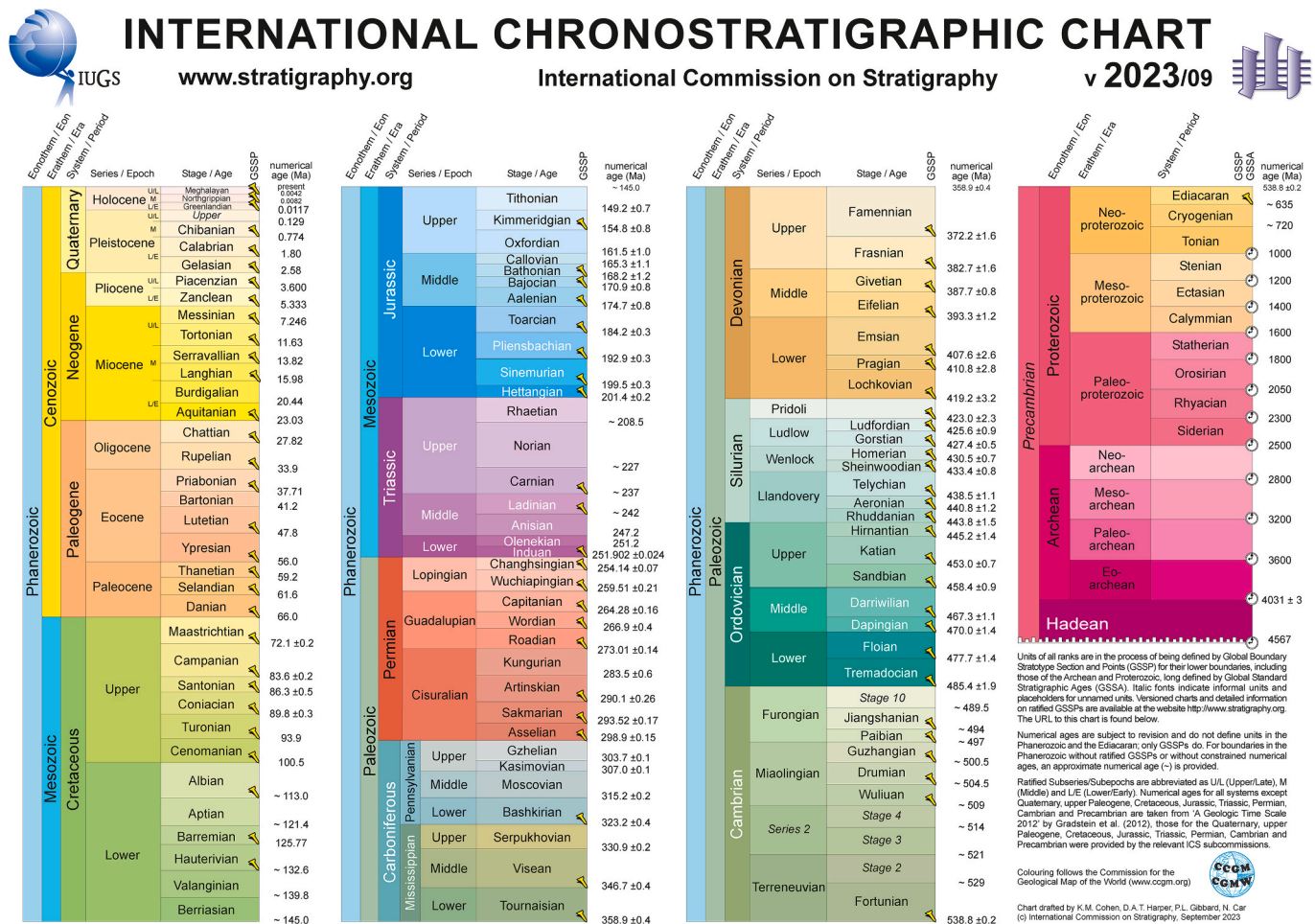


Fig. 1. The International Chronostratigraphic Chart presents the hierarchical classification of stages, series, systems, erathems, and eonothems and the numerical ages of their boundaries. The small yellow pins to the right of most stage names indicate that a GSSP has been ratified for its lower boundary. Some GSSPs for stage boundaries also mark boundaries of higher-level units. For example, the GSSP of the Danian also marks the boundary between the Cretaceous and the Paleogene, and between the Mesozoic and the Cenozoic. *Source:* Cohen et al. (2023).

driving a bronze- or gold-colored metal pin into an exposed part of a rock formation; hence their informal name: ‘golden spikes’.¹

These GSSPs serve as reference standards for global chronostratigraphic classification. At the site of a GSSP, it can be known with certainty that the boundary between two chronostratigraphic stages is located *exactly* at the tip of the golden spike. To determine the level of the boundary at any other location on Earth, chronostratigraphers need to trace the boundary laterally from the GSSP site across different regions and continents by comparing fossil contents, sedimentary rock characteristics, isotopic ratios, and changes in magnetic polarity, among other signals that may indicate isochronicity.

The practice of using local GSSPs to mark global boundaries may initially appear puzzling or even enigmatic. It seems particularly

¹ The practices of marking GSSPs in the field vary considerably. While traditionally GSSPs have been marked using metal pins, some are only marked by a plaque. Others remain unmarked in the field and are only indicated on a geological map, though these GSSPs should eventually become marked and conserved in the field (Finney & Hilaro, 2018; Gray, 2010). Some recently appointed GSSPs are not based on rock sequences and are marked differently. The base of the Holocene, for instance, is marked by a line in an ice core drilled from Greenland (Walker et al., 2008). The controversial proposal for an Anthropocene epoch has its candidate GSSP in a sediment core taken from Lake Crawford in Ontario, Canada (McCarthy et al., 2023).

perplexing that, as a general rule, GSSPs should *not* be positioned at stratigraphic levels that indicate breaks in the rock record or that display abrupt changes in fossil content. Indeed, the guidelines of the International Commission on Stratigraphy — the body that oversees the principles and procedures for assigning GSSPs — state that “an obvious boundary should be suspect” and specify that GSSPs should preferably be placed in geologically and paleontologically uneventful sections of (as nearly as possible) uniform sedimentation (Cowie et al., 1986, p. 6; Remane et al., 1996, p. 79). These formal guidelines resonate with what some of the most esteemed twentieth-century stratigraphers have argued. The British geologist Derek Ager often warned that “the most marked visible changes in their faunas or floras ... are likely to be some of the worst places for major boundaries ... the best level at which to place a boundary is, paradoxically, the level at which it is least obvious” (Ager, 1984, p. 8). Or in the memorable words of Digby J. McLaren: “Boundaries ... should be defined whenever possible in an area where ‘nothing happened’” (McLaren, 1970, p. 802). They should be “‘quiet’ boundaries, man-made by definition” (McLaren, 1978, p. 1).

In this article, I will trace the origins of this prima facie puzzling practice of designating GSSPs as a stepping stone for a broader philosophical inquiry into the conceptual, epistemic, and ontological dimensions of the use of *material reference standards* across various sciences. Following Alisa Bokulich (2020b), I will refer to these (token) reference standards as ‘scientific types’. Apart from global boundary

stratotypes, this category includes holotypes² from biological taxonomy and measurement prototypes from the international metrology of weights and measures — and perhaps others.³

There are some obvious practical and procedural correspondences in the use of these scientific types: all three involve the elevation of a tangible, material entity (an artefact or specimen) to the role of a formal reference standard in a classificatory practice by way of a stipulative, declarative speech act (“This material object *X* will hereby serve as the formal reference standard for unit *Y*”). It is a further question whether these surface correspondences are underpinned by shared conceptual and epistemic attributes. If so, identifying these attributes might allow us to articulate a general account of scientific types that is of philosophical, historical, and perhaps even practical interest. Bokulich has argued that such a shared basis of ‘typification’ exists: “There is a common focal function and status to holotypes, stratotypes, and measurement prototypes, that I argue unites all three under the common rubric of ‘scientific type’” (Bokulich, 2020b, p. 2). Accordingly, she has developed a general philosophical account of scientific types that aims to reveal interesting correspondences between the epistemologies and scientific practices of metrology, taxonomy, and geology.

I agree with Bokulich that the notion of a ‘scientific type’ picks out a category of objects and associated epistemic practices that are worthy of philosophical investigation, and I applaud her for putting this topic on the agenda of philosophy of science. That said, I will argue that her unified account of scientific types needs to be amended, as it wrongly parses the practices and purposes of standardization and classification as shared and aligned across the sciences. Consequently, I will argue that the main philosophical lessons that Bokulich draws from her account fail to hold up. As an alternative, I offer a *disunified* account of scientific types that recognizes how diverse ontological attitudes towards classification and diverse epistemic aims for standardization fit with the adoption of different *kinds* of scientific types in different contexts.

I will use the case of chronostratigraphy to illustrate the payoff of this alternative way of thinking about scientific types and referential practices. The disunified account of scientific types provides the philosophical tools and resources for evaluating an intriguing mid-twentieth-century debate about what the proper epistemic aims of chronostratigraphy should be, and which kind of scientific type would best support those aims. In short, my account of scientific types helps to understand why chronostratigraphers were — and to some extent still are — divided about whether theirs is a science of *marking* natural boundaries or of *making* conventional ones.

The structure of this article mirrors the three parts of its title. I start with a brief historical sketch of the conceptual and methodological background of the GSSP approach (Section 2). In the middle sections, I situate the notion of GSSPs in the context of scientific types more generally. First, I present Bokulich’s general account of scientific types and the philosophical lessons she derives from it (Section 3). Next, I show that this account is untenable since it rests on an erroneous interpretation of the (purported) shared function and status of scientific types (Section 4). In the final section of the middle part, I will show how the shortcomings of Bokulich’s account point to a fundamentally

² For the sake of simplicity, I will speak of ‘holotypes’ in an inclusive sense that encompasses other so-called ‘name-bearing types’, such as lectotypes and neotypes. I will return to the difference between these varieties of name-bearing types in Section 5.

³ In Section 6, we will see that there is at least one other kind of scientific type that has been described in the literature but that has not been adopted in scientific practice. Also, note that the shared ‘-type’ extension of the three examples mentioned here is largely a historical contingency. It is not required of a scientific type that it is called a ‘-type’ by scientists. For example, if an early-twentieth-century attempt at overhauling the jargon of biological taxonomy had succeeded, holotypes would today be known as ‘onomatophores’ (Simpson, 1940).

different, disunified account of scientific types: an account that acknowledges the differences in proximate standardization targets, broader epistemic aims, and underlying ontological attitudes that inform classification projects across the sciences (Section 5). In the final part of the paper, I return to the case of chronostratigraphy to demonstrate the dividends of this disunified account of scientific types for understanding a particular scientific debate. Distinguishing between kinds of types helps us make sense of a dispute among chronostratigraphers about the epistemic aims of their discipline and the ontological status of chronostratigraphic boundaries (Section 6).

2. Towards a global chronostratigraphic scale

How did chronostratigraphers end up hammering golden spikes into exposed rocks? To begin to understand how and why the GSSP approach was adopted, it helps to appreciate the problems it intends to address. This requires, first of all, a brief sketch of the historical context in which those problems arose.

2.1. A brief history of the (chrono)stratigraphic hierarchy

The geological rock record presents us with a jumbled archive of Earth history — an archive in dire need of an archivist.⁴ Stratigraphy is the science that aims to unravel the structural and temporal relations between layers of rock by analyzing their lithic, biotic, chemical, and magnetic properties on a regional or global scale. The origins of stratigraphy are usually traced to William Smith in England and Georges Cuvier and Alexandre Brongniart in France, around 1800.⁵ While working as a canal engineer, Smith observed that different ‘strata’ (a term he introduced) of rock contained characteristic fossils that could be used to laterally trace rocks across different outcrops (exposed parts of rock formation) in the countryside. The classification of fossils was already well-established at the time, but it took Smith’s efforts to marry the study of fossil differences with that of differences in sedimentation (Rudwick, 1996; Sepkoski, 2017), famously represented on his map of strata of England and Wales (Smith, 1815).

Inspired by Smith’s work, Cuvier and Brongniart began constructing schematic depictions of the strata in the Paris region. Whereas Smith focused on unraveling the structural order of strata, without making any inferences about chronology, Cuvier and Brongniart offered an interpretation of strata in geohistorical terms, as the record of a temporal sequence of geological and biotic events. Brongniart in particular was early to recognize the possibility of establishing correspondences in the order of deposition of strata across a region based on fossil content alone, regardless of their rock type (Berry, 1987). Roderick Murchison later adequately summarized the underlying principle that “the zoological contents of rocks, when coupled with their *order of superposition*, are the only safe criteria of their *age*” (Murchison, 1839, p. 9; italics in original).

This ‘principle of faunal succession’ soon became widely adopted to enrich (and eventually transform) the dominant practice of *geognosy*.⁶ Driven mainly by economic interests in mining mineral

⁴ On the metaphor of the ‘archive’ (or ‘record’) of the past, also see Currie (2023), Sepkoski (2017), and Turner (2019).

⁵ Some would want to identify deeper roots of stratigraphy in the work of Nicolaus Steno on superposition and that of 18th-century ‘geognosts’ in documenting the structural order of rock sequences. However, their work did not straightforwardly pave the way for that of Smith, Brongniart, and Cuvier. Smith, for example, had probably never heard of Steno (Hancock, 1977, p. 3). See Vai (2007) for an even deeper ‘prehistory’ of (chrono)stratigraphy going back to Leonardo da Vinci around 1500.

⁶ For further historical and philosophical discussion of this principle and of the practice of *biostratigraphy* more generally, see the work of Max Dresow (2021, 2023).

resources, geognosts had already distinguished between ‘Primary’, ‘Secondary’, ‘Tertiary’, and ‘Quaternary’ formations based on differences in petrological content. By directing their attention to distinctive fossils, the new stratigraphers were able to discern smaller units within these broad divisions (see Rudwick, 1996). Geologists across Europe named and identified *systems* such as the Cretaceous and the Jurassic within the Secondary division. By the mid-19th century, almost all current names of the stratigraphic systems had already gained widespread informal acceptance among geologists. Within the Tertiary division, Charles Lyell named what we today recognize as units below the system level: the *series* of the Eocene, Miocene, Pliocene, and Pleistocene (Lyell, 1833, 1839). Around the same time, William Smith’s nephew John Phillips put the overarching division between Primary, Secondary, and Tertiary rocks on a paleo(bio) logical footing by introducing the *erathems* of the Paleozoic, Mesozoic, and Cenozoic, based on a globally invariant sequence of drops in the history of life on earth (Phillips, 1840). Finally, the Frenchman Alcide d’Orbigny made a further downward extension of the stratigraphic hierarchy by introducing the lowest-level global chronostratigraphic rank that is recognized today: the *stage*. D’Orbigny argued that correspondence in faunal assemblages was key to classifying segments of the stratigraphic column below the system and series levels on a global scale (d’Orbigny, 1842).

By 1900, there was a broad consensus that erathems, systems, series, and stages were to be the nested ranks for the units of global stratigraphic classification (Vai, 2007). However, this agreement about ranks still left plenty of room for disagreement about which particular units to recognize at any given rank. The ‘great Devonian controversy’ is a well-known example of one such dispute, about the recognition and delineation of particular systems in the Paleozoic (Rudwick, 1985). At the stage level, however, the disagreements were even more pronounced and pervasive.

2.2. Unit stratotypes and the alignment problem

Problems began soon after d’Orbigny introduced the stage concept. Based on extensive field studies in France, combined with a thorough reading of the stratigraphic literature, d’Orbigny had argued that the Jurassic system was comprised of ten stages which “nature has delineated with bold strokes across the whole earth” (d’Orbigny, 1842, p. 603). This puzzled many foreign stratigraphers with closer knowledge of the fossil record in their respective regions. Unable to identify the same set of “bold strokes” as d’Orbigny had discerned, they began suggesting alternative ways of carving up the Jurassic that were more consistent with their local fossil record. Before long, the number of suggested stage names exploded. By the 1930s, the Jurassic expert William J. Arkell documented the use of nearly 130 different stage names for crosscutting segments of the Jurassic, amounting to “a meaningless complex of overlapping stages” (Arkell, 1933, p. 12). The practice of stratigraphy was headed for a veritable “stratigraphical Babel” that affected the delimitation of stages in general, not just of the Jurassic (Arkell, 1956, p. 461).

A proximate cause of much of the confusion lay in the method by which stages were recognized and new stage names were introduced. d’Orbigny pioneered the use of ‘typical sections’ as standards for naming and delimiting stages (Torrens, 2002). He named each stage after the locale at which “the best type (*le meilleur type*) is found, a deposit that I regard as a standard (*étalon*), that is to say as one that can always serve as a point of comparison” (d’Orbigny, 1842, p. 604). For example, d’Orbigny used a characteristic outcrop near the French town of Semur as the basis for recognizing and naming the *Sinemurian* stage of the Jurassic. This outcrop being “a true yardstick (*un point réellement étalon*) for the stage,” it could be used as a standard for tracing its limits in other rock sections across different regions and continents (d’Orbigny, 1852, p. 434). But as stratigraphers from different countries began to recognize additional ‘typical sections’ or ‘unit stratotypes’ for strata from their own region that were

unaccounted for, a patchwork of crosscutting conceptions of stages emerged (‘stage concepts’, for short).

Part of the solution was straightforward: new stage names that were synonymous with an already accepted one (i.e., referring to the same stage concept) were redundant and could be discarded. For example, some of the 100+ names for stages of the Jurassic and the Cretaceous that were introduced after d’Orbigny were obscure synonyms that could be ignored without cost. The more stubborn problem concerned the use of crosscutting stage concepts, based on the recognition of different unit stratotypes from different regions. A well-formed and complete chronostratigraphic scale would have to cover the whole of the geologic time scale and should not allow for any rocks to belong to more than one time-rock bin. This meant that the boundaries of unit stratotypes for ‘adjacent’ stages on the geologic time scale would need to be perfectly aligned, without leaving gaps or introducing overlaps between stages. However, the method of using unit stratotypes to delimit stages made this hard to achieve in practice.

One challenge was that many unit stratotypes were *unconformity-bounded*. Unconformities are surfaces of contact between strata that include a hiatus in the geological record. An unconformity-bounded unit stratotype is a rock section that is delimited by a hiatus at its base and/or top. The problem with these unconformities was that they often turned out to be local geological anomalies: they transitioned laterally into conformable contacts between strata. This raised the question of how to classify those strata that had formed during the interval that was missing from the unit stratotype. One option would be to assign them to a new, intermediate stage, but this would risk relapsing into the chaos of proliferating stage names and concepts that stratigraphers were trying to address. Another option would be to move the stage boundary beyond the limits of the unit stratotype. However, this would entail that stratigraphers could no longer rely on unit stratotypes to define stage boundaries, which would in turn aggravate the risk of creating overlaps with adjacent stages.

A closely related challenge pertained to the weak *correlation potential* of most unit stratotypes. ‘Correlation’, as understood by stratigraphers, refers to the practice of establishing correspondence in the chronostratigraphic position of rocks and rock contents from different locations around the world. To ascertain that the upper boundary of a given unit stratotype coincides with the lower boundary of the unit stratotype for the overlying stage, it was necessary to precisely match and align signals — biotic, chemical, magnetic, or other — from the edges of both unit stratotypes. This required unit stratotypes to be rich in fossil contents or other rock characters around their edges. In practice, many unit stratotypes failed to live up to this requirement, meaning that they could not be reliably correlated. Once again, this introduced a risk of introducing gaps or overlaps in the process of assembling the chronostratigraphic scale.

In theory, both problems could be addressed effectively by replacing most existing unit stratotypes with new ones that did not suffer from unconformities around their edges and that had high correlation potential. Around the mid-twentieth century, several stratigraphers advocated for this solution. But when they started looking for good candidates for alternative unit stratotypes, they soon discovered that for the majority of stages, no complete and highly correlatable sections of exposed rock could be found. The rock record resisted an easy methodological solution.

2.3. Boundary stratotypes and ‘golden spikes’

In the 1950s and ‘60s, the international stratigraphic community converged on another rather ingenious solution that promised to solve the problem of gaps and overlaps in one fell swoop. Instead of trying to align the upper boundary of one stage (determined in one location) with the lower boundary of its overlying stage (determined in another location) using unit stratotypes, it was proposed to define each boundary in one location only, using a conventionally designated

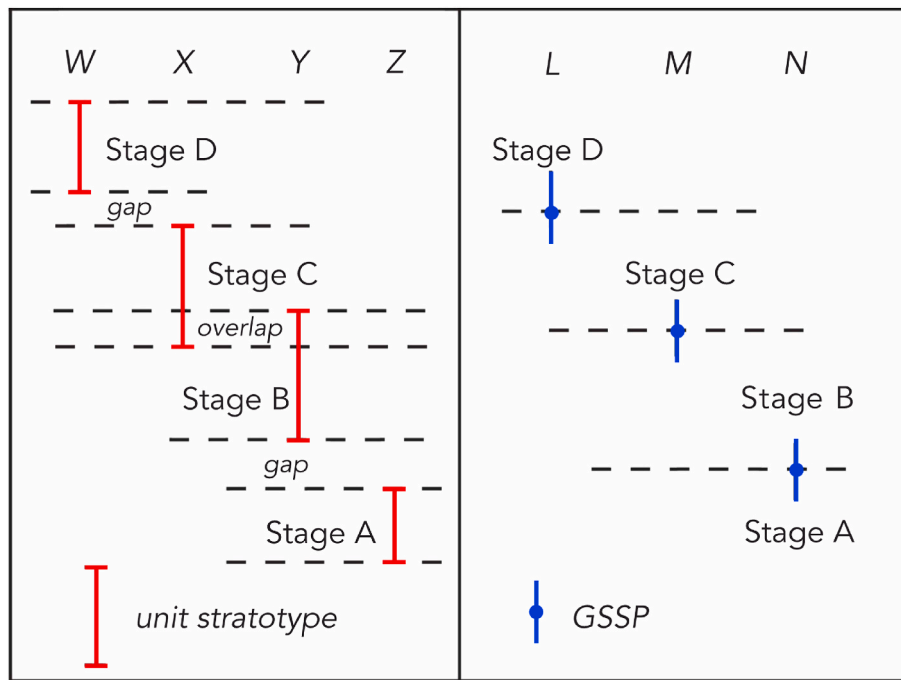


Fig. 2. A comparison of the delimitation of stage boundaries using unit stratotypes (left) and GSSPs (right). The unit stratotypes from locations W, X, Y, and Z cannot be used to define stage boundaries without leaving gaps or introducing overlaps. This problem can be addressed by assigning GSSPs in sections of continuous deposition in locations L, M, and N. (Drawn after Fig. 14 from Salvador (1994), with modifications.)

boundary marker. Thus, the GSSP approach was born.⁷ If a single golden spike were to be appointed for each stage boundary, there would no longer be any uncertainty or ambiguity about precisely which boundary point stratigraphers were trying to correlate along an isochronous plane across the Earth (as illustrated in Fig. 2). Moreover, the GSSP approach promised to facilitate the empirical work of global correlation by addressing the practical problem of low correlation potential that had hampered old the unit stratotypes. To be useful in practice, GSSPs would need to be placed in sections of as nearly as possible continuous deposition, containing as many different signals of age-significant information as possible, such as rapidly evolving fossils of different species and distinctive physico-chemical signals (Hedberg, 1976; Salvador, 1994; Smith et al., 2015). It would take considerable effort to find outcrops that satisfied these demands, but relative to the failed attempt to find outcrops that covered entire stages, the task of finding outcrops that were only continuous around a single boundary seemed feasible.

However, to satisfy the requirement of defining boundaries in a section of continuous deposition, many boundaries would have to be ‘moved’ to slightly different horizons. Hence, the practical advantages of the GSSP approach would come at the cost of needing to revise many recognized stage concepts. For example, the GSSP for the Silurian/Devonian boundary — the first GSSP ever ratified, at the appropriately named site of ‘Klonk’ in the Czech Republic — departed from the then dominant conception of the top of the Silurian and the base of the Devonian. However, this slight conceptual revision was expected to bring major practical advantages going forward. With its combined presence of deep-water graptolite fossils and shallow-water brachiopod-trilobite fossils of different species, the newly-minted boundary

promised to finally enable reliable correlation with a wide range of different rock sections across the world (Chlupáč et al., 1972).

2.4. Another kind of ‘type’?

Early on in the development of the GSSP approach, some researchers noted that the role that GSSPs could serve in chronostratigraphy resembled that which other concrete, material reference standards were already playing in other sciences. The American geologist Hollis D. Hedberg, who served as chairman of the International Subcommittee of the Stratigraphic Classification, pointed out that GSSPs are “as essential to stratigraphic classification as types are to biologic classification, or as Bureau of Standards references are to physical measures” (Hedberg, 1959, p. 676). Hedberg stopped short of spelling out the nature of the correspondence between these different material reference standards, but soon others jumped in to discuss the analogy — and to dispute it.

Some concurred that GSSPs and holotypes served kindred roles but disagreed about the nature of the correspondence between these reference standards (Holland et al., 2003; Melchin et al., 2004; Sylvester-Bradley, 1967). Others claimed that GSSPs could not be fruitfully compared to holotypes (Schindewolf, 1960, 1970; Schoch, 1989), or argued that only the comparison with standard units of measure in physics was meaningful and instructive (Harland, 1992; Remane, 2000; Remane et al., 1996). Still others considered this metrological analogy to be unrevealing or even misleading (Aubry & Berggren, 2000; Bell, 1959). Finally, there were those who viewed all analogies between GSSPs and reference standards from other sciences with deep suspicion (Walsh, 2005; Walsh et al., 2004).

While some of these discussions about the (dis)analogies between GSSPs, holotypes, and measurement prototypes became rather heated, they remained philosophically superficial. They failed to ascend to a level of abstraction from where one could begin to draw more general philosophical lessons about the landscape of reference-fixing practices across the sciences. This changed when Alisa Bokulich recently stepped into this space and articulated the first general philosophical account of ‘scientific types’ (Bokulich, 2020b).

⁷ The British initially referred to GSSPs as ‘marker points’ (Ager, 1963; Sylvester-Bradley, 1967), whereas the Americans and the International Subcommittee on Stratigraphic Classification pushed the term ‘boundary stratotypes’ (ACSN, 1965; Hedberg, 1976). The terms later coalesced into the notion of ‘golden spikes’ (Holland, 1986) and the more formal ‘Global Boundary Stratotype Section and Points’, or GSSPs (Cowie et al., 1986).

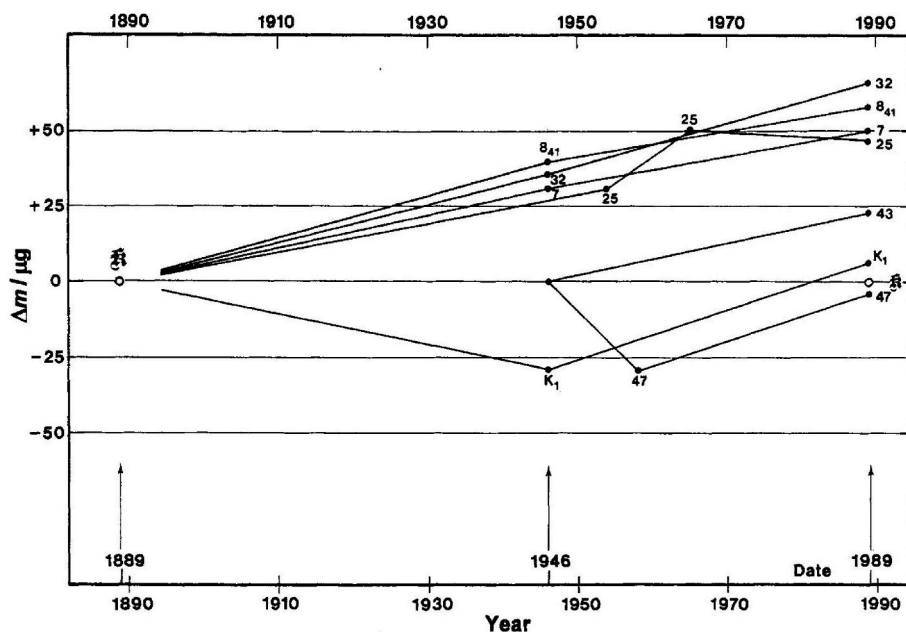


Fig. 3. The change in mass of several of the sister copies of the IPK with respect to the mass of the IPK itself (marked with the K-symbol), as determined in periodic verifications in 1889, 1946, and 1989. Source: Girard (1994), Fig. 5.

3. Scientific types: the very idea

Bokulich argues that despite their many discipline-specific differences, GSSPs, measurement prototypes, and holotypes can be united under the rubric of ‘scientific types’ because of their “common focal function and status” (Bokulich, 2020b, p. 11). In this section, I will present her account and distill three key theses from it: one about the shared function of scientific types, one about their shared status, and a third about the implications of their status and function for our more general philosophical understanding of standardization practices in the sciences.

3.1. The function of scientific types

According to Bokulich, the common functional profile of scientific types pertains to their role in bridging definitions with their realizations. This functional profile is captured in the general definition of scientific types that she offers:

A scientific type is a concrete individual object that serves as a standard of reference for, and realization of, the definition or taxon category that it names. (Bokulich, 2020b, p. 2)

To understand what Bokulich means by scientific types serving a standardizing role in the realization of definitions, it helps to consider the definition-realization distinction in the context of metrology, where it originates.

In the vocabulary of metrology, ‘realization’ refers to the concrete materials and procedures for bringing the definition of a quantity into practical use (JCGM, 2012). To understand the role of a realization in relation to a prototype-based definition, consider the definition for the unit of mass: the kilogram. Until recently, ‘1 kg’ was defined as follows:

The kilogram is the unit of mass. It is equal to the mass of the international prototype of the kilogram. (BIPM, 1901)

This definition is a linguistic entity. It is a statement about what the term ‘1 kg’ refers to by convention. The international prototype of the kilogram (IPK) features in this definition as the reference standard for ‘1 kg’. In the definition, the IPK only features in the abstract, as a concept.

The IPK qua concrete, material object is a key component of the realization of the kilogram definition (sometimes called its ‘primary realization’). The IPK’s material form makes it possible to put the definition into practice by enabling the calibration of other measurement instruments (Riordan, 2019). Apart from the IPK, the realization of the kilogram definition includes the balances that are used for comparing masses, maintenance and cleaning procedures, and theoretical models for analyzing and correcting the results of comparisons (Riordan, 2019; Tal, 2017).

According to Bokulich’s general definition of scientific types, the role of the IPK in mediating between a definition and its (primary) realization is shared by other scientific types. Let us capture this functional role more precisely in the following thesis about the shared function of scientific types:

Functional Unity Thesis: a scientific type helps to realize the definition of a unit of measurement or classification in scientific practice by serving as the formal, concrete, material reference standard for the use of the unit term.

As such, the Functional Unity Thesis does not specify any (minimal) criteria that a material object ought to satisfy to exercise its functional role. In any given scientific context further demands will be placed on the constitution and/or preservation of the relevant scientific types. For example, since it is vitally important that the realization of unit of measure can be reproduced with as little variation as possible, the artefact that served as the IPK was fashioned (back in the 1880s) out of a highly durable platinum-iridium alloy and has been kept under a set of glass jars to make sure that it didn’t collect any dust. To keep wear and tear to a minimum, it was rarely removed from its enclosure. However, even this extreme care did not prevent the IPK from changing its constitution, by losing an ever-so-slight part of its mass through the decades. This possibility for scientific types to change has implications that bring us to their second shared feature on Bokulich’s account.

3.2. The status of scientific types

First, let us consider the case of the instability of the IPK in a bit more detail. Since it is impossible to weigh the IPK against itself at an earlier

time, determining that its mass had changed was not straightforward. It was inferred from periodic comparisons to its ‘sister copies’: metal cylinders that were manufactured to have as nearly as possible the same mass as the IPK.⁸ On their own, these periodic verifications could at most demonstrate that the masses of the prototypes were drifting relative to each other (Quinn, 1991). The verifications underdetermined whether the sister copies had been gaining mass with respect to the IPK, or if the IPK was losing mass with respect to its copies (Fig. 3). Based on further studies and background knowledge, metrologists eventually concluded that the likeliest interpretation was that the mass of the IPK itself was unstable: its mass had been falling with respect to the mean of the ensemble of its sister copies (Davis, 2003).⁹

Bokulich argues that this case illustrates how a philosophically interesting tension can arise between the definition of a unit and the scientific type that helps to realize it. She claims that while the definition of the kilogram stipulated that the mass of the IPK was exactly one 1 kg, the periodic verifications established that, in fact, it weighed less than 1 kg. Thus, she concludes, the IPK itself was found to fail to belong to the class of objects of exactly 1 kg. This failure of the IPK to match the kilogram definition rendered that definition inadequate, spurring metrologists to come up with a new definition. This resulted in the formal redefinition (in 2019) of the kilogram in terms of a fundamental physical constant: the Planck constant.¹⁰

Bokulich acknowledges that this need to redefine the kilogram may initially appear puzzling, since we intuitively think of definitions as infallible and absolutely accurate: “If the kilogram was defined to be whatever the IPK weighed, then how could it be judged inadequate?” (Bokulich, 2020b, p. 15). She argues that the answer to this question lies in recognizing what Eran Tal has called the “myth of absolute accuracy” of measurement standards. Tal points out that once we appreciate the distinction between a definition and its realizations, it becomes easy to see that any realized measurement standard is necessarily always somewhat inaccurate (Tal, 2011, 2017). Bokulich claims that with respect to a measurement prototype such as the IPK, this leads to the insight that it too can be inaccurate, in the more specific sense of failing to belong to the class for which it serves as the definitional reference standard. According to Bokulich, this is exactly what metrologists diagnosed. It was “precisely the metrologists’ discovery [that the IPK failed to weigh exactly 1 kg] that prompted them to make the stipulative redefinition, moving from ... the IPK artefact definition of the kilogram to ... the new Planck-constant-tied definition of the kilogram”. Indeed, Bokulich argues that unless we concede that the IPK failed to weigh exactly 1 kg, “it is not clear how one can make sense of the decision to undertake the difficult redefinition” (Bokulich, 2020b, p. 17).

Importantly, Bokulich thinks that this lesson does not hold solely for the IPK or other measurement prototypes but can be extended to all scientific types. The myth of absolute accuracy holds sway over our thinking about concrete, material reference standards across the board and needs to be debunked for the entire category: “In all three cases [of

measurement prototypes, holotypes, and GSSPs], a detailed examination of scientific practice shows that the supposed infallibility or absolute accuracy of types in belonging to the taxon that they name is a myth” (Bokulich, 2020b, p. 26). This implies that taxonomists can discover that a holotype fails to belong to the taxon for which it serves as the name-bearer, and that chronostratigraphers can find out that a GSSP is not actually situated at the boundary for which it was designated as the reference standard. Bokulich summarizes this general point about the fallibility of scientific types as follows:

There is a tension ... between a scientific type serving as part of a stipulative definition, hence conventionally defined to be infallible, and a scientific type serving the purpose of securing a stability and coherence of scientific practice ... Although types are *taken by convention* to be infallible, scientists are fully aware that *in practice* types are in fact fallible” (Bokulich, 2020b pp. 11, 14; italics in original).

Let us capture this component of Bokulich’s account of scientific types with the following thesis:

Fallibility Thesis: any scientific type can fail to belong to the class or unit for which it, by definition, serves as the formal reference standard.

The Fallibility Thesis leaves open how it is determined that a scientific type fails to conform to its formal, definitional role. As with the Functional Unity Thesis, the practical details will differ from one scientific context to the next. However, Bokulich does think that they diagnoses of failure follow a general pattern. She explains that a failure of a scientific type to belong to its class or unit is established by appeal to independent ‘common sense’ standards. For example:

In the case of prototypes, the independent standard by which to judge whether the mass of the IPK was exactly 1.00 kg was, first, the common sense background knowledge that all concrete physical objects take up contamination from their environment and lose mass through cleaning; and second, the “communal knowledge” of the IPK’s stability through intercomparison projects (coherence testing) with other kilogram standards both at BIPM and around the world. (Bokulich, 2020b, pp. 25–26)

Bokulich’s more general point here is that a ‘common sense’ standard can in exceptional circumstances overrule the formal standard. In the absence of such an independent standard, “there could never be a determination that the IPK is anything other than 1.00 kg” and we would not be able to make claims to the effect that “the mass of the platinum-iridium artefact in Sèvres belongs, for example, to the class of things that are 0.9999995 kg” (Bokulich, 2020b, p. 15). The same applies to other scientific types: we can only establish their failure to belong to the class they are defined to belong to by appealing to independent standards of communal knowledge or practice. For any scientific type, there are “independent standards for judging which taxons [*sic*] these scientific types belong to” (Bokulich, 2020b, p. 25). Let us recognize this as a third thesis that is closely tied to the Fallibility Thesis:

Independent Standards Thesis: To determine that a type fails in the sense specified by the Fallibility Thesis one needs to appeal to an independent standard of ‘common sense’ (i.e., communal knowledge or practice).

This thesis encapsulates Bokulich’s overarching aim of articulating a “Duhemian philosophy of scientific types” (Bokulich, 2020b, p. 22). Without entering into the details of Pierre Duhem’s philosophy of science, we can understand Bokulich’s aim as that of applying and extending an important distinction that Duhem introduced in his critique of conventionalism (Duhem, 1954).

⁸ Among these sister copies there are six ‘official copies’ that are stored together with the IPK, several dozen ‘national prototypes’ that are used by bureaus of standards all over the world, and a few ‘working prototypes’ that are used by the BIPM to calibrate the national prototypes.

⁹ That such a decrease could occur was no surprise, since another series of measurements had already shown that the prototypes lost mass in the cleaning and washing procedures that were carried out prior to the verifications (Davis et al., 2016; Girard, 1994).

¹⁰ The kilogram is currently defined “by taking the fixed numerical value of the Planck constant h to be $6.62607015 \times 10^{-34}$ when expressed in the unit J·s, which is equal to $\text{kg}\cdot\text{m}^2\cdot\text{s}^{-1}$. The IPK followed the fate of the prototype meter, which was abandoned in 1963 in favor of a new definition of the unit of length based on the distance traveled by light in a vacuum in a fraction of a second.

Duhem famously disagreed with the conventionalist position that "certain fundamental hypotheses of physical theory cannot be contradicted by any experiment because they constitute in reality *definitions*, and because certain expressions ... take their meaning only through them" (Duhem, 1954, p. 209; as cited in Bokulich, 2020b, p. 24). For example, if common sense tells us that an object is in free fall, but it is experimentally shown that it actually has a slightly variable acceleration, the conventionalist would conclude that the common sense view has been proven wrong. Duhem, in contrast, argued that a situation of this kind presented physicists with a genuine choice: they could either stick with the theoretical, symbolic meaning of 'free fall' as 'uniformly accelerated motion' or they could prioritize the common sense meaning and construct "a mechanics in which the words 'free fall' no longer signify 'uniformly accelerated motion,' but 'fall whose acceleration varies according to a certain law'" (Duhem, 1954, p. 210).

Bokulich's ultimate aim is to adapt Duhem's argument to the context of typification, by arguing that scientists can (and sometimes do) appeal to common sense in challenging a 'symbolic' system of conventionally stipulated scientific types. She writes: "Scientists do not cede common sense, and when scientific types falter in their ability to secure a stability and coherence of practice, that external coherence itself becomes an independent standard by which to judge the accuracy and adequacy of the type. When this happens, scientists can choose to locate the errors in the scientific types and revise them accordingly" (Bokulich, 2020b, p. 26). If Bokulich is right about this, her account of scientific types has an important philosophical payoff. It would demonstrate that Duhem's critique can be applied and extended to scientific domains and practices for which its import has not previously been appreciated. However, I will argue that we should not accept her account at face value. In the next section, I will show that on closer scrutiny all three theses turn out to be untenable.

4. Reconsidering the philosophy of scientific types

I will enter my critical evaluation of the three theses that I distilled from Bokulich's account of scientific types by zooming in on the "myth of absolute accuracy". As we saw in the previous section, Bokulich claims that we should recognize and dispel this myth for all scientific types by accepting the Fallibility Thesis. In Section 4.1, I will argue that this attempt at generalizing the myth of absolute accuracy is rooted in a misinterpretation of that very myth. The upshot of this error is that the Fallibility Thesis fails to hold. In Section 4.2, I will argue that apart from being mistaken about the status of scientific types as fallible entities, Bokulich also errs in attributing a common function to them. Hence, the Functional Unity Thesis also fails to withstand scrutiny. I conclude this section by briefly considering the implications for the Independent Standards Thesis. Due to its entanglement with the Fallibility Thesis, the Independent Standards Thesis falls with it, but its falsity invites further reflections on the goals of scientific systems of typification that will feed into Section 5.

4.1. The myth of absolute accuracy and the status of scientific types

The myth of absolute accuracy originates in the intuition that measurement standards necessarily provide us with absolutely accurate results. We have seen that this truly is a myth since the realization of a unit definition is never completely exact and fully replicable. Therefore, a definition's realization cannot provide complete certainty about the quantity being measured.

The impossibility of obtaining absolute accuracy is most evident for measurement standards that are defined in terms of a physical process (Tal, 2011). The 'second', for example, is defined as the time it takes to cycle through 9,192,631,770 oscillations of the radiation that unperturbed cesium-133 atoms emit and absorb when they switch between certain states at a temperature of absolute zero. Since this definition specifies an ideal state that is experimentally inaccessible, no realization

of the second can possibly be an absolutely accurate realization of the definition.¹¹

Prototype-based definitions are different in an important respect: they do not appeal to an ideal physical process or state but to an actual object. In case of the kilogram definition from before 2019, this is the mass of the IPK. Its mass, it seems, really *is* exactly 1 kg since this is stipulated by the kilogram definition. Nevertheless, the full realization of the IPK introduces inaccuracies related to the use of balances, cleaning procedures, etc. Thus, "like any set of physical procedures, the replication of the kilogram is not completely exact and therefore involves some uncertainty" (Tal, 2017, p. 244).

As we saw in the previous section, Bokulich argues that the instability of the IPK introduces a further wrinkle into this analysis. She reasons that since it was determined that the mass of the IPK had changed, the inaccuracy of the realization of the kilogram had to be partially attributed to the IPK itself: it failed to weigh exactly 1 kg. On the one hand, this line of reasoning seems impeccable: since the IPK weighed exactly 1 kg when it was designated as the ultimate reference standard for '1 kg', and since its mass changed later, it follows that the IPK no longer weighed exactly 1 kg. On the other hand, if '1 kg' was *defined* in terms of the mass of the IPK, it follows that for as long as the IPK served as the reference standard of '1 kg', it must have weighed *exactly* 1 kg. And so, we arrive at an apparent contradiction: the IPK both does and does not weigh exactly 1 kg — this cannot be right.

4.1.1. Reweighing the case of the kilogram

We can avoid this contradiction by distinguishing between the IPK's *material instability* and its role in securing the *referential stability* of '1 kg'. The act of designating an object as the measurement prototype for a unit term bestows referential stability (or fixity) upon that term up to the point that the prototype is replaced or discarded. Accordingly, the unit term '1 kg' was referentially stable over the period that started with the inauguration of the IPK as the kilogram prototype in 1889 and ended with its redefinition in 2019. Referential stability fails to hold across kilogram definitions, but the stability in the meaning of '1 kg' is absolute under each definition separately, including under the IPK-based definition. Material stability, on the other hand, is never absolute: any measurement prototype will be subject to contamination and/or degradation over time. Whereas the lack of absolute material stability of the IPK rendered it inevitable that its constitution would change *over time*, it continued to secure absolute referential stability of the unit term '1 kg' kilogram *at any given time* within the period in which it served as the kilogram prototype.

This interpretation of the relation between referential and material stability is supported by the metrological literature. For example, Georges Girard, the metrologist who meticulously documented the instability of the IPK in the third verification procedure from the late 1980s, maintained that "[b]ased on the value of the international prototype (always 1 kg exactly) one may deduce mass values for the others" (Girard, 1994, p. 320). The only way the IPK can be "always 1 kg exactly" despite the evidence from the verifications, is by distinguishing its material instability from its role in securing referential stability. Likewise, the U.S. National Institute of Standards and Technology ex-

¹¹ In practice, metrologists use several atomic clocks (maintained by national metrology laboratories around the globe) to multiply realize the definition of the second. This presents them with the task of "forg[ing] a unified second out of disparately ticking clocks ... [by] continually evaluating the accuracy of each realization relative to the ideal cesium transition frequency and correcting its results accordingly. But the ideal frequency is experimentally inaccessible, and primary standards have no higher standard against which they can be compared" (Tal, 2011, pp. 1087–1088).

plains on its website that whereas the masses of the IPK's sister copies could change over time, "by definition, of course, the IPK mass could not actually change: Because it was the official kilogram, its mass was always exactly 1 kg, even if it actually gained or lost mass!"¹² Finally, metrologist Richard S. Davis noted that under the prototype-based definition of the kilogram, " $\{m(\text{IPK})\} = 1$ by definition, a value that has no uncertainty even though $m(\text{IPK})$ can be unstable with respect to a fundamentally constant mass, such as the rest mass of the electron" (Davis, 2011, p. 3978).¹³ As Davis points out here, it is possible to conceive of the material instability of the IPK by comparison to (what we expect to be) a more stable or even fundamentally invariable mass. But Davis is careful not to infer from such a comparison that the IPK failed to weigh exactly 1 kg when it was the ruling kilogram standard, since this would violate the referential stability condition: $\{m(\text{IPK})\} = 1$. Like the other metrologists, he recognizes that the IPK infallibly weighed exactly 1 kg during its reign as the official kilogram standard.

We can conclude, then, that the Fallibility Thesis does not hold for measurement prototypes such as the IPK. If we designate an object (such as a metal cylinder) as the measurement prototype for a unit term (such as '1 kg'), the unit term is given referential stability due to the infallibility of the object in instantiating the measurement unit for as long as it serves as that unit's measurement prototype.

4.1.2. Infallibility and defeasibility

The fact that the IPK was eventually abandoned as the kilogram standard tells us that the infallibility of measurement prototypes does not imply their *indefeasibility*. The redefinition of the kilogram in 2019 demonstrates that the IPK was defeasible qua reference standard. Appreciating this distinction between infallibility and defeasibility helps to reinforce the point that the redefinition of the kilogram cannot possibly have been *precipitated* by the IPK's failure to weigh 1 kg (as Bokulich argued). As metrologists pointed out long before the redefinition was enacted, upon a future redefinition "the mass of the international prototype would no longer be known exactly but would have to be determined by experiment" (Mills et al., 2005, p. 71).¹⁴

The real motivation for the redefinition, then, was not the IPK's failure to weigh 1 kg but its failure to afford a primary realization that allowed '1 kg' to be reproduced with the desired operational measurement accuracy. When the level of measurement accuracy and precision that was sought could no longer be guaranteed using a realization that relied on a measurement prototype, the time had come to consider adopting a definition that was not based on a measurement prototype.¹⁵ This lesson is consonant with the general message that Tal tried to drive home when debunking the myth of absolute accuracy: since no realization is absolutely accurate, there is always room to improve accuracy. If there is a practical need for it and if the theoretical and practical means are available, scientists from time to time decide to adopt a new definition that enables higher accuracy realizations. This applies to the case of the kilogram. The redefinition of the kilogram was prompted by the prospect of obtaining a higher-accuracy primary realization than a

definition based on a (infallible but defeasible) measurement prototype could provide for.¹⁶

These lessons about the infallibility but defeasibility of measurement prototypes can be extended to other scientific types, such as GSSPs. There are close conceptual and epistemic parallels between the practice of using a GSSP to correlate a boundary and that of using the IPK to measure the mass of an object. Both varieties of scientific types function alike in coupling a stipulative definition to its necessarily inaccurate realization. (Although chronostratigraphers do not use the term 'realization', we can conceive of the epistemic practice of correlating a chronostratigraphic boundary on a global scale as broadly analogous to it.)

Consider, for example, that extending a boundary from its GSSP-defined level to other locales is inevitably somewhat inaccurate and uncertain, e.g., due to the imperfect chronostratigraphic resolution of the guiding criteria (fossils, chemical signals, etc.) that are available at the GSSP site. For example, the GSSP that marks the boundary between the Devonian and Carboniferous (located in La Serre, in southern France) was placed in an outcrop that was later found to contain many 'reworked' fossils — fossils that have eroded from their original location and that have been redeposited in a different geological layer. Moreover, the primary signal for correlating this boundary — the lowest occurrence of the conodont marker fossil *Siphonodella sulcata* — turns out to occur at a stratigraphic level significantly below that of the GSSP. These complications make it challenging to accurately determine the level of the Devonian-Carboniferous boundary outside of the GSSP section (Kaiser, 2009). The GSSP approach further resembles the case of the IPK in its reliance on 'sister copies' in the form of auxiliary stratotypes that support correlation in different paleogeographic regions, but that do not have the same infallible status as the primary stratotype (Cowie et al., 1986). Furthermore, there is an analog of the problem of material instability for GSSPs. In theory, it is possible for a GSSP itself to become unstable and 'drift' from the level it was originally hammered into due to erosion, slope failure, overgrowth, or other problems arising from poor preservation (Finney & Hilario, 2018). But, as with the IPK, a drifting golden pin continues to formally define a stage boundary for as long as it serves as the GSSP for a stage — the boundary definition is referentially stable.

The role of GSSPs in securing referential stability is acknowledged in formal guidelines for establishing chronostratigraphic boundaries. If it becomes impossible to realize a boundary with the required level of accuracy (whether due to the lack of good guiding criteria or due to the instability of the GSSP itself), the guidelines allow for a boundary to be redefined by appointing a new GSSP: "A GSSP ... can be changed if a strong demand arises out of research subsequent to its establishment. But in the meantime it will give a stable point of reference" (Remane et al., 1996, p. 80). In other words, the guidelines tell us that whilst GSSPs are defeasible, they provide referential stability until their status as reference standards is revoked. This is illustrated by the case of the Devonian/Carboniferous boundary (DCB) mentioned above. The problems with the correlation potential of the GSSP for the base of the Carboniferous have led to the creation of a task group charged with redesignating this GSSP. However, the stratigraphers involved in this effort duly recognize that until a new GSSP has been formally approved, the current GSSP continues to formally define the base of the Carboniferous. They point out that "to avoid any stratigraphic chaos and

¹² <https://www.nist.gov/si-redefinition/kilogram-present>.

¹³ In the notation used, $\{Q\}$ is the numerical value of physical quantity Q .

¹⁴ Likewise, Davis (2011) pointed out that before the redefinition the relative uncertainty of $m(\text{IPK})$ was zero, but that immediately after the redefinition this relative uncertainty took a positive value that would have to be determined experimentally.

¹⁵ In this context, precision ("the closeness of agreement among measured values obtained by repeated measurements" (Tal, 2011)) can be considered an aspect of accuracy. It is worth noting in this context that Bokulich (2020a) provides an insightful discussion of the difference between precision and accuracy in geochronology that is not susceptible to my criticism of her discussion of accuracy in the chronostratigraphic context.

¹⁶ As Sally Riordan has pointed out, in the case of the kilogram another motivation played a role: "If the drift [to the mass of the IPK] had not been apparent, the desire to replace the IPK would remain. Indeed, the desire to replace an artefact mass standard with something 'more fundamental' existed long before the third — and even the second — verification took place" (Riordan, 2019, p. 161). However, this (ontological) desire for a 'more fundamental' standard is closely related to an (epistemic) preference for a more invariant primary realization.

ambiguity where and how the DCB should be placed in the light of the current ongoing discussions, it has to be stressed that the GSSP at La Serre is still valid and our current reference” (Aretz & Corradini, 2021, p. 289).

4.2. Scientific types of a different kind?

So far, I have argued that the Fallibility Thesis fails to hold for measurement prototypes (such as the IPK) and GSSPs, but I have not yet considered the case of holotypes. The reason for this is not that holotypes are the exception — I will end up arguing that the Fallibility Thesis needs to be rejected for *all* scientific types. Instead, holotypes merit a separate discussion because they point us to an additional problem that we need to address before we can evaluate their status. This problem concerns the Functional Unity Thesis.

4.2.1. Realization and application

The assumption that scientific types have a shared functional profile is baked into Bokulich’s definition of scientific types as concrete objects that “serve as a standard of reference for, and realization of, the definition or taxon category that it names” (Bokulich, 2020b, p. 2). The problem with this definition is that holotypes do not fit it. Holotypes aren’t conduits for the *realization* of definitions; they only mediate in the *application* of names.

To appreciate the difference between these functional roles, let us start by considering (counterfactually) how holotypes would be deployed if, analogous to GSSPs and measurement prototypes, their job was to help ‘realize’ a taxon description or definition. It would mean that taxonomists used each holotype as a yardstick for arbitrating on the inclusion of other specimens in the same taxon. This would in turn require each holotype to be a perfectly representative element of its taxon. However, this is not what holotypes are. They are typically rather atypical, and sometimes outright aberrant members of their taxa. This should come as no surprise since holotypes are often designated before the full range of variation in their taxa is known (Daston, 2004; Witteveen, 2018). Fortunately, this unrepresentativeness is not necessarily a problem, since holotypes do not play any privileged, standardizing role in the practice of delimiting taxa. This is made explicit in the taxonomic codes of nomenclature that specify the role of holotypes. For example, the International Code of Zoological Nomenclature states in its ‘Principles’ section that “nomenclature does not determine the inclusiveness or exclusiveness of any taxon, nor the rank to be accorded to any assemblage of animals, but, rather, provides the name that is to be used for a taxon whatever taxonomic limits and rank are given to it ... The device of name-bearing types allows names to be applied to taxa without infringing upon taxonomic judgment” (ICZN, 1999).

Another way of articulating this difference in the functional roles of holotypes vis-à-vis GSSPs or measurement prototypes is to appeal to the classical distinction between *fixing the reference* and *giving the meaning* of a term or expression (Kripke, 1980). When a specimen α is designated as the holotype for the species name ‘Xus yus’, the referent of this name is fixed to the taxon that includes α , without making any firm commitments about the hypothesized limits or boundaries of this taxon. What those boundaries are is a question that is left open for further empirical research. In contrast, when we designate a GSSP for the lower boundary of the (imaginary) Flinstonian stage and its overlying Jetsonian stage, we have not merely fixed the reference of the term ‘Flinstonian stage’ but we have also given it meaning by specifying its limits. The same holds, *mutatis mutandis*, for measurement prototypes: they do not merely help to coordinate the application of unit terms, but they also help to

specify the meaning of those unit terms.¹⁷ In sum, there is a genuine difference in epistemic aims between using scientific types to delimit units and using scientific types to apply names to units that should be delimited by other means.

This difference in functional roles among scientific kinds refutes the Functional Unity Thesis. It is not true that all scientific types serve to assist in the realization of definitions. Holotypes don’t; they only mediate in the application of names. Interestingly, Bokulich appreciates this distinction at some level, e.g., when she acknowledges that “a holotype does not define a taxon in the sense of determining how a given species is delimited” (Bokulich, 2020b, p. 5). However, she fails to recognize that this is a distinction that ought to make a difference to her account of scientific types. Consider, for example, that if holotypes are not in the business of realizing definitions, the reason for replacing a holotype cannot be that the realization that it is part of is insufficiently accurate. Accuracy is a term that characterizes realizations, not applications. A user of a taxon name can apply it correctly (to the taxon that includes the holotype) or incorrectly (to the taxon that fails to include it), but not with insufficient accuracy.¹⁸ What reason, then, can there be for replacing a functional holotype?¹⁹ And what does this tell us about their (in)fallibility? To answer this question, we need to zoom in on a further distinction regarding the application of names to biological taxa: the distinction between a name’s *valid designation* and its *prevailing usage*.

4.2.2. Validity, usage, and revision

The valid designation of a name is determined by a baptism that turns a specimen into a taxon’s holotype. This baptismal act tells us, for instance, that “The taxon *Xus yus* is the taxon that minimally includes this specimen α as the official name-bearer for ‘Xus yus’”. However, a stipulative act of this kind cannot prevent that, in practice, a taxon name sometimes becomes used to refer to a taxon that fails to include the holotype for that name. In other words, the actual usage of a name can fail to be aligned with the name’s official, valid designation.²⁰ To see how a situation of this kind can arise, and how it can motivate the replacement of a holotype, let us take an example.

In the early 2000s, taxonomists became convinced that what had previously been considered two subspecies of the species *C. tenuimanus* — commonly known as marron, a freshwater crayfish — were actually two distinct species (Austin & Ryan, 2002). One of these newly recognized species has a broad distribution that includes populations in Australia, Chile, China, South Africa, and the U.S.A. The other species is

¹⁷ Interestingly, Kripke himself failed to apply his distinction correctly to a metrological example he discussed. Kripke argued that the “standard meter stick *S*” — i.e., the measurement prototype for ‘1 m’ — fixes the reference of ‘1 m’ without giving its meaning. More specifically, he asserted that we can use *S* at t_0 to determine the reference of the phrase ‘1 m’, but that we cannot say that *S* is necessarily 1 m long at t_0 , because “if heat had been applied to this stick *S* at t_0 , then at t_0 stick *S* would not have been 1 m long” (Kripke, 1980, p. 55). This is incorrect, as Eric Loomis (1999) has convincingly argued. Briefly: if *S* had been heated, *S* would still have been exactly 1 m long since *S* serves as the reference standard for ‘1 m’ in that situation. The case is analogous to the actual case of the IPK. It changed materially but continued to serve as the ultimate reference standard for the kilogram.

¹⁸ It is of course possible for a holotype to degrade so much that it is no longer possible to determine to which taxon the name it carries should be applied. But this is neither a problem of inaccuracy of the relevant kind, nor (as I will show) a problem with the putative fallibility/inaccuracy of holotypes that Bokulich is concerned with.

¹⁹ I speak of ‘functional holotypes’ here to exclude cases of holotypes that went missing, were damaged, or perished too much for them to be reliably attributed to a particular taxon. We can ignore such cases since they do not present instances of what Bokulich recognizes as holotypes failing to belong to their taxa.

²⁰ Zoologists speak of a taxon’s ‘valid name’, botanists use the term ‘correct name’. I will follow the zoological terminology.

geographically restricted to the upper reaches of the Margaret River in Western Australia. Given that the holotype for *C. tenuimanus* had originally been collected (back in 1911) from this small river population, the name ‘*C. tenuimanus*’ would from now on apply to this geographically restricted species. A new name, ‘*C. cainii*’, was introduced for the marron species with the almost global presence.

Taxonomic revisions of this kind are routine. They call for a readjustment of how an existing name is used that follows the name’s valid designation, as specified by its holotype. But sometimes changing existing name usage can be hard to achieve, for example because the existing usage is entrenched in non-taxonomic contexts, such as in legislation for conservation or in commerce. This was also the case with ‘*C. tenuimanus*’: recreational fishers, scientists, and the aquaculture industry had all grown accustomed to using this name for the marron species with a global distribution. It would be a challenge to ask these non-taxonomists to start using the name ‘*C. cainii*’ instead. Recognizing this challenge, a group of researchers at the Department of Fisheries in Western Australia made a formal request to the International Commission on Zoological Nomenclature (‘the Commission’) to revoke the name-bearing status of the current holotypes for ‘*C. tenuimanus*’ and ‘*C. cainii*’ and to select a new name-bearing specimen (a lectotype) for ‘*C. tenuimanus*’ from within the range of the broad-ranging marron species (Molony et al., 2006).²¹ This proposed change in designations would dissolve the nomenclatural confusion by realigning the valid designation of ‘*C. tenuimanus*’ with its long-standing name usage, instead of requiring name usage to be changed.

This case teaches us two lessons that are relevant in reflecting on Bokulich’s account. First, it shows that, like measurement prototypes and GSSPs, a specimen that serves as a holotype provides referential stability in the application of a name until it is stripped of its role as name-bearer. Upon taxonomic revision of the marron species, the holotype continued to belong to the species for which it had already served as name-bearer, even if this species now turned out to be much smaller. The researchers who appealed to the Commission duly recognized this: it was *because* the holotype for ‘*C. tenuimanus*’ belonged to the geographically restricted species, that they made their request. Bokulich gets this wrong when she claims that requests of this kind are prompted by situations in which “a holotype carries the name designating one taxon, but in fact belongs to another” (Bokulich, 2020b, p. 14). In reality, these requests are prompted by a name being *prevalingly used* for one taxon, but in fact belonging to (in the sense of validly designating) another.²²

The second lesson this case teaches us is that the grounds for replacing a holotype are importantly different from the grounds for replacing a measurement prototype or a GSSP. We saw in Section 4.1 that scientific types of the latter sort may need to be replaced when the accuracy of the realizations that they are part of is deemed inadequate for users’ demands. In contrast, the case of the marron species shows that users can call for a holotype to be replaced due to an inadequacy *on part of the users themselves*. The case shows that when users of a taxon name prevalingly apply it incorrectly, the favored way of restoring coherent scientific practice might be to ‘reset’ the valid designation of the name by selecting a lectotype from the species for which that name is

being prevalingly used. This point about restoring coherence leads me to briefly consider the final thesis that I distilled from Bokulich’s account: the Independent Standards Thesis.

4.2.3. External coherence and independent standards

As we saw in Section 3.2, the Independent Standards Thesis holds that an appeal to an independent ‘common sense’ standard is required to judge that a scientific type fails in the sense specified by the Fallibility Thesis. Since we have already seen that the Fallibility Thesis must be rejected, the Independent Standards Thesis may not seem to warrant further consideration. Its dependence on the soundness of the Fallibility Thesis tells us that it too must be false.

Nevertheless, I think it is worth considering whether Bokulich’s general point about ‘common sense’ being able to serve as an ‘independent standard’ could be salvaged if we cleave it from the Fallibility Thesis. The question then becomes whether scientists, driven by a concern “to secure a stability and coherence of practice” rely on certain standards of communal knowledge or practice when judging that a scientific type needs to be replaced (Bokulich, 2020b, p. 26). I believe that the answer is negative. To show why, let me return to the case of the marron species once more.

It goes without saying that the request to the Commission to replace the holotype for ‘*C. tenuimanus*’ was motivated by a desire to maintain a stable and coherent community practice of name usage. However, this by no means implies that community practice (or common knowledge) can function as a *standard* that has the normative power to overrule the ruling, type-based standard.

Indeed, if community practice had acted as a genuinely overruling standard, we would expect it to have led to the designation of a lectotype for ‘*C. tenuimanus*’, as requested in the petition. But this is not what happened. Upon considering the petition, the Commission ruled that the misalignment of prevailing community usage of ‘*C. tenuimanus*’ with the name’s valid designation was not a strong enough reason to discard the holotype and select a lectotype. As one of the Commissioners who voted against the proposal put it, the change in the meaning of ‘*C. tenuimanus*’ following the taxonomic revision was “simply a matter of getting used to,” as that revision had been “completely in accordance with the Code” (ICZN, 2008, p. 321). In other words, the request to preserve external coherence of practice was not granted and the appeal by the scientific community did not lead to a revision of the ruling scientific type.

This case is not an anomaly. Sometimes the Commission rules in favor of requests to replace a holotype, sometimes it doesn’t. The Commissioners make informed decisions and do not blindly follow the *communis opinio*. If they did, they would indeed risk subverting the normative force of scientific types and thereby undermining any type-based system of reference standards. If, for example, any appeal to ‘prevailing usage’ in zoological nomenclature would automatically lead to the replacement of the corresponding holotype, this would threaten to undercut the distinction between ‘valid’ and ‘invalid’ usage and with it undermine the very purpose of assigning holotypes for names. To avoid this erosion of normativity while still taking seriously concerns of maintaining stability and coherence in usage, the International Code of Zoological Nomenclature and other similar codes of nomenclature offer a middle-road solution of allowing for requests to replace scientific types on a case-by-case basis, subject to discussion and voting by one or more bodies of experts.

This lesson also applies to other domains in which scientific types are used, often with even stricter requirements for considering their replacement.²³ For example, the guidelines on the establishment of

²¹ Requests of this sort typically appeal to Article 75.6 of the International Code of Zoological Nomenclature, which reads: “When an author discovers that the existing name-bearing type of a nominal species-group taxon is not in taxonomic accord with the prevailing usage of names and stability or universality is threatened thereby, he or she should maintain prevailing usage and request the Commission to set aside under its plenary power the existing name-bearing type and designate a neotype” (ICZN, 1999 Article 75.6).

²² In other work, I argue that the distinction between a taxon name’s valid designation and its actual usage is a special case of the distinction between ‘semantic reference’ and ‘speaker’s reference’ that is familiar to philosophers of language (Witteveen, 2015, 2021; cf. Brzozowski, 2020).

²³ In botany, a request to appoint a neotype to preserve prevailing usage of a name needs to receive a qualified majority approval (60% votes in favor) from two committees and a congress before it is adopted. In chronostratigraphy, any proposal for the replacement of a GSSP likewise must survive a tiered system of voting by different bodies. The replacement of measurement standards requires a majority vote from the General Conference on Weights and Measures.

Table 1

A comparison of scientific types showing the key differences between scientific types that function as naming standards and scientific types that function as definition standards.

	holotypes (including lecto- and neotypes)	GSSPs (i.e., boundary stratotypes)	measurement prototypes
<i>kind of scientific standard</i>	naming standard	definition standard	definition standard
<i>proximate standardization target</i>	application of names to taxa	realization of (chronostratigraphic) boundaries	realization of (measurement) units
<i>epistemic aim</i>	globally uniform usage of taxon names	globally uniform fixation of the extension of (chronostratigraphic) unit terms	globally uniform fixation of the extension of (metric) unit terms
<i>ontological attitude</i>	realism and/or pluralism	conventionalism	conventionalism

GSSPs emphasize the caution that needs to be exercised in proposing to replace a GSSP by requiring that any GSSP should normally remain in place “for a minimum period of ten years” before a request to replace it should be considered (Remane et al., 1996). After those ten years, a GSSP should only be replaced in exceptional circumstances, lest “we generate a situation where boundaries are repeatedly redefined,” the result of which “will be chaos” (Holland et al., 2003, p. 69). This underscores that revising a scientific type is walking a tightrope. It is not a question of using one standard to overrule another, but of cautiously applying the means for revision that are offered as part of setting the ruling standard.

5. The disunity of scientific types

The previous section has shown that Bokulich’s general philosophical account of scientific types is untenable. Contrary to what Bokulich argued, not all scientific types serve the function of realizing definitions, nor is any scientific type fallible qua reference standard. Should we conclude from this that the concept of a scientific type needs to be abandoned?

In this section, I will argue that the notion of scientific types can be retained in a useful and meaningful way, but only as part of an alternative, disunified account of scientific types. I will sketch the outlines of this alternative perspective on scientific types by pulling together some of the lessons from the previous section that point to important differences in the epistemology and ontology of different typification practices. The disunified account of scientific types emerges from recognizing a fundamental divide between two kinds of material reference standards: naming standards and definition standards. Naming standards provide a material link between names and their referents, definition standards establish a material link between definitions and their realizations. Holotypes are an example of naming standards; measurement prototypes and GSSPs fall into the category of definition standards (see Table 1).

It is tempting to regard naming standards simply as stripped-down versions of definition standards since standards of the latter kind *also* provide the names (or terms) for the units they help to define and realize. (The IPK, for instance, specifies the value of the mass quantity in the unit term ‘kilogram’.) This view is correct insofar as we focus on the proximate goals of standardization. Whereas naming standards promote stability and uniformity in the use of names, definition standards provide this service for the names of units *and* the named units themselves. But when we turn to the underlying epistemic aims of the two varieties of scientific types, this perspective on the difference between naming standards and definition standards becomes too limiting. Definition standards do not just standardize *more*, but they standardize *differently*, to different epistemic ends.

We have seen that in the context of chronostratigraphy and metrology, the aim of standardization using scientific types is to facilitate the delimitation of classificatory units with as little variation and

ambiguity as practically possible. But in biological taxonomy, standardization serves virtually the opposite purpose: allowing practitioners to adopt incongruent and varying conceptions of their classificatory units of interest. By letting holotypes specify which name applies to which taxon, regardless of how taxon boundaries are drawn, taxonomists are handed a common standard for identifying and labeling taxa without requiring them to agree on matters of classification with their peers or predecessors. In other words, holotypes facilitate scientific discussion and development concerning the delimitation and classification of taxa by providing an independent means for coordinating how we talk about them.

An early advocate of the use of holotypes in taxonomy noted that this coordinating role of holotypes resonated with a widely shared ontological commitment of biological taxonomists. By appointing holotypes, taxonomists would “have a designation ready for the final entity, but also available for any number of approximating concepts which may follow each other with no unnecessary confusion” (Cook, 1898). Contemporary taxonomists need not sign onto this (monistic) realist outlook on classification — as slowly converging on a final, supremely ‘natural’ classification — to appreciate the advantages of this approach to standardization. If one expects disputes and disagreements about how to delimit species or genera to be deep and lasting rather than superficial and fleeting, it makes even more sense to only standardize the relations between names and taxa and to leave room for different conceptions of those taxa.

The metrologists’ choice of definition standards over mere naming standards is likewise informed by a particular ontological attitude. It would be misguided for a metrologist to adopt a realist or pluralist attitude toward the definition of a base unit of measure such as the kilogram. A quest for the ‘real’, nature-given value of the kilogram would be elusive, and chaos would result from having to differentiate and translate between a plurality of incommensurable definitions. What counts as a kilogram is a (scaling) convention (Tal, 2018).²⁴ Notably, this conventionalism about units and scales of fundamental quantities by no means renders metrology a straightforward practice of formulating and applying definitions. As we saw in the discussion of the myth of absolute accuracy, realizing and replicating metric units is a complex endeavor that involves empirical research to increase accuracy.

Contemporary chronostratigraphy resembles the metrology of weights and measures in its ontological attitude of conventionalism. The chronostratigrapher Hollis Hedberg, whom we briefly encountered in Section 2.3 as a trailblazer for the GSSP approach, presented and defended this conventional stance toward chronostratigraphic classification with vigor. In his view, the GSSP approach offered a purely

²⁴ This does not exclude the possibility that the new conventional unit definition of the kilogram, anchored in the Planck constant, is in some sense more fundamental than the old IPK-based definition, though even this point is far from straightforward — it happens to be heavily contested (Riordan, 2019; Tal, 2018).

practical, conventional solution to “our inability to comprehend and handle in their entirety variations in rock characters, rock properties, rock attributes, without breaking up rock strata into more or less arbitrary units ... [J]ust as we have to cut our meat into bite-size pieces before we can swallow it, so practically, we have to break the properties of our rock sequences into comprehensible and conveniently useable categories and units before we can handle them” (Hedberg, 1958, p. 1882). Since the most useful units would be those that “can be used unequivocally and with the same sense and scope by everyone” the adoption of GSSPs to conventionally fix the extension of those units looked like the obvious solution (Hedberg, 1965, p. 102). Or was there an alternative?

6. Making or marking chronostratigraphic boundaries

The disunified account of scientific types helps to appreciate that the standardization strategy of contemporary chronostratigraphy was, in fact, not self-evident. Unlike in the cases of metrology and biological taxonomy, the question of what kind of scientific type would suit chronostratigraphy did not have a straightforward answer.

We have seen that the choice of approach to standardization depends on what one considers to be its proper epistemic aim, which is in turn often informed by a particular ontological attitude toward the delimited units. As it turns out, mid-twentieth century chronostratigraphers were divided on what constituted the appropriate aim and attitude for their discipline. While the majority sided with Hedberg’s chronostratigraphic conventionalism, a vocal minority made a case for chronostratigraphic realism. This minority maintained that chronostratigraphy should not be in the business of *making* artificial, conventional boundaries, but should rather be a science of *marking* “natural” boundaries that exist mind-independently. The latter view is at odds with the conventionalist orientation of the GSSP approach (and with the use of definition standards more generally), but it does allow for the standardized coordination of chronostratigraphic names using naming standards. This, we will see, is what the proponents of natural chronostratigraphy suggested would be the right approach to take.

6.1. Chronostratigraphic realism?

Before considering in more detail this alternative proposal for the use of scientific types in chronostratigraphy, let us first take a closer look at the ontological thesis of chronostratigraphic realism that motivated it. Many stratigraphers from the U.S.S.R. and a minority of mostly German, French, British, and American stratigraphers signed onto a form of chronostratigraphic realism. The American geologist and paleontologist Norman D. Newell was perhaps its most prominent exponent. In a lecture on the mission of chronostratigraphy, he posed the question: “Should it [chronostratigraphy] reflect the most widespread events in earth history or is it an arbitrary device for measuring time?” His own response: “To the first question, I answer yes. To the second, no ... Conspicuous world events ... provide natural datums for the division of chapters of earth history and should be stressed in standard stratigraphic classification” (Newell, 1966, p. 80). The natural datums he was mainly thinking of were the mass extinctions of marine genera that he had been studying for over a decade (Sepkoski, 2020). In Newell’s view, these were good candidates for indicating the boundaries of chronostratigraphic systems. The division between the Triassic and Jurassic Systems, for instance, ought to be regarded as “a real, tangible, boundary based on an important biological event” (Newell, 1966, p. 78). He likewise believed that it should be possible to identify natural boundaries below the level of Systems, using changes in the evolutionary sequence of lifetimes “at a lower taxonomic level” than those that signaled the end of, say, the Permian, Triassic, or Cretaceous (ibid., p. 77).

Newell’s conception of changes in fossil assemblages as datums for recognizing natural boundaries was not the only contender. Others pointed to deformations of the earth’s crust (diastrophism), large-scale

up- or downward movements of the seas, changes in climatic history, or combinations of these as indicators of ‘natural events’ of chronostratigraphic significance (Dunbar & Rodgers, 1957, p. 302ff.). In Hedberg’s view, this diversity of interpretations was exactly why the project of natural chronostratigraphy was doomed: “To doubt that ‘natural breaks’ delimit our present systems and series, it is only necessary to examine the controversies which have arisen between specialists over almost every one of these boundaries, and to note the lack of agreements which still remains” (Hedberg, 1948, pp. 452–453). It was precisely this perennial state of disagreement that signaled a practical need to end equivocations in the usage of terms by conventionally fixing boundaries (Hedberg, 1961).

Newell’s assessment of this situation was completely different. Far from considering disagreements to be pointless, he argued that they were at the heart of what made chronostratigraphy an empirical science. “Stratigraphy is still in the exploratory phase and growth of knowledge in this, as in other branches of science, is naturally characterized by controversy and instability” (Newell, 1966, p. 79). Attempting to eliminate controversy by decree, as proposed by the International Subcommittee on Stratigraphic Classification that Hedberg chaired, would be antithetical to its epistemic aims: “[C]ontroversy is an essential part of the method of science and I favor trial and error as a means of discovering the facts in stratigraphy.” Hence, the classification of chronostratigraphic stages should “remain flexible” and fixing their boundaries conventionally would be “premature and undesirable” (ibid., p. 73). Others went further and warned that the proposal to conventionally define chronostratigraphic boundaries would “lead stratigraphy *ad absurdum*” (Schindewolf, 1960, p. 24) and would “completely invert the logic of stratigraphical classification” (Krassilov, 1978, p. 97). In short, the proponents of chronostratigraphic realism considered it a gross mistake to try to mitigate the disorder in the use of chronostratigraphic unit terms by artificially and arbitrarily ‘freezing’ their boundaries using GSSPs.

Many others who did not take as firm a stance in this debate still recognized that the global stratigraphical community faced a genuine choice in deciding how to work towards a global chronostratigraphic scale. For example, the co-authors of a widely read textbook on the principles of stratigraphy noted that the question of the right “philosophical attitude” towards the basis for the subdivision of the stratigraphic record was “at present ... a controversial problem and it would certainly be premature to anticipate a final judgment” (Dunbar & Rodgers, 1957, p. 302).

6.2. The taxonomic solution

Those who favored a natural basis for chronostratigraphic classification did not deny that standardization and regulation of some sort would be desirable. W. J. Arkell, for instance, noted even before the installation of the Subcommittee that Hedberg would preside over that while “there will be general agreement that this is the body which should eventually act ... stratigraphers should consider very carefully what it is that they want to ask it to do” (Arkell, 1946). To his mind, it would be “highly undesirable” for such a commission to produce a “stereotyped scheme” of classification. Instead, he suggested that the best way to address the stratigraphical Babel was to take a cue from the zoologists’ practice of instituting rules that only governed nomenclature:

If, however, the Congress were to restrict itself to promulgating a Code of Rules, all future research, by objective application of the Rules, would lead to a progressively closer approximation to the ideal scheme desired, as has undoubtedly been the case in zoological nomenclature. (Arkell, 1946, p. 2)

Others were independently converging on the same conclusion. For example, the British paleontologist P. C. Sylvester-Bradley argued that it would be a key advantage to “leave the question of boundaries to be

decided by each stratigrapher in his own way, just as, in zoology, determination of the boundaries of a taxon is left to the subjective opinion of each taxonomist” (Sylvester-Bradley, 1967, p. 52).

In line with this suggestion, several stratigraphers started pointing out that a stable scheme of applying names to stratigraphic units required a stratigraphic equivalent of holotypes. Moreover, some of them noted that there were obvious candidates for the role: the unit stratotypes that had originally been used to describe and delimit stages could be reinstated as purely nomenclatural devices: “Type sections in stratigraphic classification should have no more significance than [as] name bearers” (Wilson, 1959, p. 770). “Apart from their function as name bearers, type localities are conceded no special importance” (Scott, 1960, p. 580). Newell also noted that by using unit stratotypes merely as “nomenclatural devices (name-bearers) ... [they] would not rigidly fix universal temporal limits” (Newell, 1966, p. 79). A reference fixing system based on such “nominote stratotypes” (Newell, 1972) or “nominotypes” (Krassilov, 1978) would be able to accommodate different viewpoints and changes in chronostratigraphic classification while steering clear of nomenclatural chaos.

The success that botanists, zoologists, and paleontologists were having with this approach spoke in its favor. Some argued that it was high time for “stratigraphic taxonomy [to] catch up with biologic taxonomy” (Wilson, 1959). One commentator even speculated that the concept of a unit stratotype had itself probably been “derived by analogy from biologic taxonomy. What was not fully realized by geologists was that rules relating to priority and to types in biology, from the very beginning stabilized only names — and did not stabilize (freeze) the concepts for which the names stood” (Bell, 1959, p. 2864). In any case, it was clear that biological taxonomists would never tolerate a system in which a select few would by decree ‘freeze’ concepts for the others: “Few modern biologists (or paleontologists) would tolerate the idea that the *limits* of a species, genus, or family be fixed by an arbitrarily chosen ‘type’” (Newell, 1966, p. 73). Why, then, would modern stratigraphers put up with a system of this kind?

6.3. The road not taken

Even after the GSSP approach was formally adopted by the International Commission on Stratigraphy as the official method for standardizing chronostratigraphic classification (Hedberg, 1976), the alternative approach of designating naming standards continued to have its outspoken supporters. For example, in the textbook *Stratigraphy: Principles and Methods*, Robert Schoch noted his discontentment with the direction that standardization in stratigraphy had taken (Schoch, 1989). He characterized the GSSP approach as “perhaps unscientific” since it implied that “any individual or small group of investigators attempting to arrive at some sort of natural stratigraphic classification will inherently be at a disadvantage if forced not only to pursue science but also to compete with an arbitrary classification” (Schoch, 1989, p. 228). Following Newell and others before him, Schoch took the view that international rules and regulations should merely govern nomenclature. This could be done by designating specific “type points” at particular locations as name anchors.²⁵

An original author would be free to define the boundaries of the unit as he or she saw fit, but later workers would be free to revise such boundaries as they thought necessary. However, when it came to naming units, any particular unit would take the oldest name among any type points contained within the recognized unit. If no type point were contained within the recognized unit, then it would be

necessary to propose a new name (and type point) for the unit.” (Schoch, 1989, p. 229).

Schoch presented the most elaborate proposal to date for a nomenclatural system modeled on taxonomy, but it never resulted in the development of an alternative standard for chronostratigraphic classification.

A more detailed historical account of why the stratigraphical community ended up favoring the adoption of one kind of scientific type over another will have to wait for another day. My objective here has been to demonstrate that in order to even begin asking these historical questions, we need to recognize that scientific types come in different kinds. Furthermore, the disunified account of scientific types that I have presented helps to appreciate how a choice of scientific type not only needs to be responsive to the epistemic aims and conceptual attitudes of a scientific field but can actually end up co-determining those aims and attitudes.

With the adoption and implementation of the GSSP approach, the room for realist attitudes toward assembling a global chronostratigraphic scale was steadily marginalized. The very idea of trying to identify natural boundaries became ridiculed as a pointless search for the “One True Boundary” (Walsh, 2004, p. 145) of every chronostratigraphic unit, or as a misguided “quest for the golden horizon” ... the assumption that the magic moment that was the beginning of [say] the Devonian was ordained by God or Marx long before man started his investigation” (Ager, 1994, p. 106). The GSSP approach, on the other hand, has become synonymous with a scientific attitude in chronostratigraphy, since “a discipline working with units of measure which are not rigorously defined cannot claim to be scientific” (Remane, 2003, p. 11). This is in stark contrast with the diagnosis from Dunbar and Rodgers several decades earlier that it was an open question which “philosophical attitude” towards the delimitation of chronostratigraphic boundaries would be the most appropriate (Dunbar & Rodgers, 1957, p. 302).

However, the debate over whether chronostratigraphic boundaries are human-made or should only be human-marked has not been entirely put to rest. The American stratigrapher and paleontologist Spencer Lucas has recently taken up the gauntlet again by arguing that the GSSP approach continues to be “fraught with problems of philosophy and methodology”. He maintains that earlier critiques of it were unduly brushed aside and need to be reconsidered (Lucas, 2018, p. 15; cf. Henderson, 2019). Much like Newell before him, Lucas and some others take the firm stance that fixing boundaries by committee, at levels of non-events is “inherently not scientific” (Lucas, 2018, p. 2). A truly scientific chronostratigraphy should be empirically open-ended rather than closed by convention (Lucas, 2018, 2019, 2020, 2023; also see Davydov, 2020). “The way forward with chronostratigraphy is to return to the concept of natural chronostratigraphy, with improvements based on modern techniques” (Lucas, 2018, p. 15).

7. Conclusion

I began this article by drawing attention to a puzzle about golden spikes or GSSPs. What motivates chronostratigraphers to mark global boundaries locally, in sections where ‘nothing happened’? In Section 2, I gave a first-pass answer: the GSSP approach addressed the persistent problem of gaps and overlaps in the construction of a global chronostratigraphic scale. Using golden spikes to conventionally ‘make’ boundaries appeared to be the silver bullet solution to the problems that chronostratigraphers had been grappling with. Fast forward to the end of the paper, and we saw that this approach toward fixing boundaries by convention was contested — and to some degree continues to be. According to some, boundaries should not be man-made, but only man-marked; the mission of chronostratigraphy should be to discover natural time-rock boundaries.

In between these brief excursions into the history of

²⁵ Using specific (non-dimensional) points within unit stratotypes as name-bearers would incidentally solve a problem that Walsh (2005) later pointed to: if someone hypothesizes that a stage boundary runs through a stratotype, which stage is it the name-bearer of?

chronostratigraphy, I developed a philosophical account of scientific types that helps put these debates into perspective. Bokulich (2020b) helpfully introduced the notion of a ‘scientific type’ to draw attention to the use of concrete, material reference standards across a range of classificatory and measurement practices in the sciences. However, I have argued that the philosophical account of scientific types that she developed fails to accurately represent their status and function(s). Scientific types are neither fallible, nor do they share the function of realizing definitions or categories. Scientific types are *infallible* but defeasible reference standards, and they are functionally diverse. Some scientific types serve to link names to their referents, others help to bridge definitions (of conventional boundaries or units) and their realizations. Recognizing the infallibility, defeasibility, and functional disunity of scientific types yields a more complex and nuanced picture of the similarities and differences between reference-fixing practices across the sciences.

Declaration of competing interest

None.

Data availability

No data was used for the research described in the article.

Acknowledgements

Earlier versions of this paper were presented in the symposium ‘Making Geologic Time’ at the Society for Philosophy of Science in Practice meeting in Ghent and in the Research Seminar of the Section for History and Philosophy of Science in Copenhagen. I thank the participants in these meetings for their encouraging feedback. I would also like to thank Hernan Bobabilla, Federica Bocchi, Alisa Bokulich, Max Dresow, Katie Kendig, Spencer Lucas, Jesper Lundsryd Rasmussen, Thomas Reydon, David Sepkoski, Lara Sabatier, Richard Staley, Aja Watkins, and two anonymous reviewers for helpful discussion and feedback.

References

- ACSN. (1965). Note 32: Definition of geologic systems. *Bulletin of the American Association of Petroleum Geologists*, 49(10), 1694–1703. <https://doi.org/10.1306/a66337b6-16c0-11d7-8645000102c1865d>
- Ager, D. V. (1963). Jurassic stages. *Nature*, 198(4885), 1045–1046. <https://doi.org/10.1038/1981045a0>
- Ager, D. V. (1984). The Stratigraphic Code and what it implies. In W. A. Berggren, & J. A. V. Couvring (Eds.), *Catastrophes and Earth History: The New Uniformitarianism* (pp. 91–100). <https://doi.org/10.1515/9781400853281.91>
- Ager, D. V. (1994). *The Nature of the Stratigraphic Record* (3rd ed.). John Wiley & Sons.
- Aretz, M., & Corradini, C. (2021). Global review of the Devonian-Carboniferous boundary: An introduction. *Palaeobiodiversity and Palaeoenvironments*, 101(2), 285–293. <https://doi.org/10.1007/s12549-021-00499-8>
- Arkell, W. J. (1933). *The Jurassic System in Great Britain*. Oxford: Clarendon Press.
- Arkell, W. J. (1946). Standard of the European Jurassic. *GSA Bulletin*, 57(1), 1–34.
- Arkell, W. J. (1956). Comments on stratigraphic procedure and terminology. *American Journal of Science*, 254(8), 457–467. <https://doi.org/10.2475/ajs.254.8.457>
- Aubry, M.-P., & Berggren, W. A. (2000). The homeless GSSP: The dilemma of the Paleocene/Eocene boundary. *Tertiary Research*, 20(1/4), 107–112.
- Austin, C. M., & Ryan, S. G. (2002). Allozyme evidence for a new species of freshwater crayfish of the genus *Cherax* Erichson (Decapoda: Parastacidae) from the south-west of Western Australia. *Invertebrate Systematics*, 16(3), 357–367. <https://doi.org/10.1071/it01010>
- Bell, W. C. (1959). Uniformitarianism — or uniformity. *AAPG Bulletin*, 43, 2862–2865. <https://doi.org/10.1306/Obda5f69-16bd-11d7-8645000102c1865d>
- Berry, W. B. N. (1987). *Growth of a Prehistoric Time Scale* (Revised Edition). Blackwell Scientific Publications.
- BIPM. (1901). *Comptes Rendus des Séances de la Troisième Conférence Générale des Poids et Mesures*. Gauthier-Villars.
- Bokulich, A. (2020a). Calibration, coherence, and consilience in radiometric measures of geologic time. *Philosophy of Science*, 87(3), 425–456. <https://doi.org/10.1086/708690>
- Bokulich, A. (2020b). Understanding scientific types: Holotypes, stratotypes, and measurement prototypes. *Biology and Philosophy*, 35(54), 1–28. <https://doi.org/10.1007/s10539-020-09771-1>
- Brzozowski, J. A. (2020). Biological taxon names are descriptive names. *History & Philosophy of the Life Sciences*, 42, 1–25.
- Chlupáč, I., Jaeger, H., & Zikmundova, J. (1972). *The Silurian-Devonian boundary in the Barrandian* (Vol. 20, pp. 104–174). <https://doi.org/10.35767/gscpgbull.20.1.104.1>
- Cohen, K. M., Harper, D. A. T., & Gibbard, P. L. (2023). *ICS International Chronostratigraphic Chart 2023/09*. International Commission on Stratigraphy, IUGS. www.stratigraphy.org
- Cook, O. F. (1898). The method of types. *Science*, 8(198), 513–516.
- Cowie, J. W., Ziegler, W., Boucot, A. J., Bassett, M. G., & Remane, J. (1986). Guidelines and statutes of the International Commission on Stratigraphy. *Courier Forschungsinstitut Senckenberg*, 83, 1–14.
- Currie, A. (2023). Of records and ruins: Metaphors about the deep past. *Journal of the Philosophy of History*, 17(1), 154–175. <https://doi.org/10.1163/1872263612341493>
- Daston, L. (2004). Type specimens and scientific memory. *Critical Inquiry*, 31(1), 153–182.
- Davis, R. S. (2003). The SI unit of mass. *Metrologia*, 40(6), 299–305. <https://doi.org/10.1088/0026-1394/40/6/001>
- Davis, R. S. (2011). The role of the international prototype of the kilogram after redefinition of the International System of Units. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, 369(1953), 3975–3992. <https://doi.org/10.1098/rsta.2011.0181>
- Davis, R. S., Barat, P., & Stock, M. (2016). A brief history of the unit of mass: Continuity of successive definitions of the kilogram. *Metrologia*, 53(5), A12–A18. <https://doi.org/10.1088/0026-1394/53/5/a12>
- Davydov, V. I. (2020). Shift in the paradigm for GSSP boundary definition. *Gondwana Research*, 86, 266–286. <https://doi.org/10.1016/j.gr.2020.06.005>
- d’Orbigny, A. (1842). *Paléontologie Française; Terrains Jurassiques*. Paris: Victor Masson.
- d’Orbigny, A. (1852). *Cours Élémentaire de Paléontologie et de Géologie*. Paris: Victor Masson.
- Dresow, M. (2021). Measuring time with fossils: A start-up problem in scientific practice. *Philosophy of Science*, 88(5), 940–950. <https://doi.org/10.1086/714855>
- Dresow, M. (2023). Biased, spasmodic, and ridiculously incomplete: Sequence stratigraphy and the emergence of a new approach to stratigraphic complexity in paleobiology, 1973–1995. *Journal of the History of Biology*, 56(3), 419–454. <https://doi.org/10.1007/s10739-023-09720-0>
- Duhem, P. (1954). *The Aim and Structure of Physical Theory* (P. P. Wiener, trans.). Princeton University Press.
- Dunbar, C. O., & Rodgers, J. (1957). *Principles of Stratigraphy*. John Wiley & Sons.
- Finney, S. C., & Hilario, A. (2018). GSSPs as international geostandards and as global geoheritage. In E. Reynard, & J. Briha (Eds.), *Geoheritage: Assessment, Protection and Management*. <https://doi.org/10.1016/b978-0-12-809531-7.00010-1>
- Girard, G. (1994). The third periodic verification of national prototypes of the kilogram (1988–1992). *Metrologia*, 31(4), 317–336. <https://doi.org/10.1088/0026-1394/31/4/007>
- Gray, M. (2010). GSSPs: The case for a third, internationally recognised, geoconservation network. *Geoheritage*, 3(2), 83–88. <https://doi.org/10.1007/s12371-010-0028-3>
- Hancock, J. M. (1977). The historic development of biostratigraphic correlation. In E. G. Kauffman, & J. E. Hazel (Eds.), *Concepts and Methods of Biostratigraphy* (pp. 3–22). Dowden, Hutchinson and Ross Inc.
- Harland, W. B. (1992). Stratigraphic regulation and guidance: A critique of current tendencies in stratigraphic codes and guides. *The Geological Society of America Bulletin*, 104(10), 1231–1235.
- Hedberg, H. D. (1948). Time-stratigraphic classification of sedimentary rocks. *The Geological Society of America Bulletin*, 59(5), 447.
- Hedberg, H. D. (1958). Stratigraphic classification and terminology. *AAPG Bulletin*, 42(8), 1881–1836.
- Hedberg, H. D. (1959). Towards harmony in stratigraphic classification. *American Journal of Science*, 257(10), 674–683. <https://doi.org/10.2475/ajs.257.10.674>
- Hedberg, H. D. (1961). The stratigraphic panorama. *The Geological Society of America Bulletin*, 72(4), 419. [https://doi.org/10.1130/0016-7606\(1961\)72\[419:rsp\]2.0.co;2](https://doi.org/10.1130/0016-7606(1961)72[419:rsp]2.0.co;2)
- Hedberg, H. D. (1965). Earth history and the record in the rocks. *American Philosophical Society*, 109(2), 99–104.
- Hedberg, H. D. (Ed.). (1976). *International Stratigraphic Guide*. John Wiley & Sons.
- Henderson, C. M. (2019). Henderson’s harangue #5. *Permophiles*, 67, 6–9.
- Holland, C. H. (1986). Does the golden spike still glitter? *Journal of the Geological Society*, 143(1), 3–21. <https://doi.org/10.1144/gsjgs.143.1.0003>
- Holland, C. H., Bassett, M. G., & Rickards, R. B. (2003). Stability in stratigraphy. *Lethaia*, 36(1), 69–70. <https://doi.org/10.1111/j.1502-3931.2003.tb00102.x>
- ICZN. (1999). *The International Code of Zoological Nomenclature* (4th ed.). London: International Trust for Zoological Nomenclature.
- ICZN. (2008). Opinion 2212 (case 3267) *cherax tenuimanus* Smith, 1912 (Crustacea, Decapoda, parastacidae): Proposed designation of neotype not accepted and usage not conserved. *Bulletin of Zoological Nomenclature*, 65(4), 320–322. <https://doi.org/10.21805/bzn.v65i4.a17>
- JCGM. (2012). *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM)* (3rd ed., p. 200). JCGM <http://www.bipm.org/vim>
- Kaiser, S. I. (2009). The Devonian/Carboniferous boundary stratotype section (La Serre, France) revisited. *Newsletters on Stratigraphy*, 43(2), 195–205. <https://doi.org/10.1127/0078-0421/2009/0043-0195>
- Krassilov, V. (1978). Organic evolution and natural stratigraphical classification. *Lethaia*, 11(2), 93–104. <https://doi.org/10.1111/j.1502-3931.1978.tb01292.x>
- Kripke, S. A. (1980). *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Loomis, E. (1999). Necessity, the a priori, and the standard meter. *Synthese*, 121(3), 291–307. <https://doi.org/10.1023/a:1005229326821>

- Lucas, S. G. (2018). The GSSP method of chronostratigraphy: A critical review. *Frontiers in Earth Science*, 6, 191. <https://doi.org/10.3389/feart.2018.00191>
- Lucas, S. G. (2019). Phanerozoic chronostratigraphy: Top-down instead of bottom-up boundary definitions. *The 2nd International Electronic Conference on Geosciences*, 26. <https://doi.org/10.3390/iecg2019-06199>
- Lucas, S. G. (2020). GSSP-Based chronostratigraphy: Should boundaries be defined by arbitrarily chosen non-events? *Pemphiles*, 68, 9–11.
- Lucas, S. G. (2023). Imprecision and instability of the phanerozoic chronostratigraphic scale: A solution. *Evolving Earth*, 1, Article 100011. <https://doi.org/10.1016/j.eve.2023.100011>
- Lyell, C. (1833). *Principles of Geology, Being an Attempt to Explain the Former Changes of the Earth's Surface by Reference to Causes Now in Operation* (Vol. III). London: John Murray.
- Lyell, C. (1839). *Nouveaux Éléments de Géologie*. Paris: Pitois-Levrant.
- McCarthy, F. M., Patterson, R. T., Head, M. J., Riddick, N. L., Cumming, B. F., Hamilton, P. B., Pisaric, M. F., Gushulak, A. C., Leavitt, P. R., Lafond, K. M., Llew-Williams, B., Marshall, M., Heyde, A., Pilkington, P. M., Moraal, J., Boyce, J. I., Nasser, N. A., Walsh, C., Garvie, M., ... McAndrews, J. H. (2023). The varved succession of Crawford Lake, Milton, Ontario, Canada as a candidate Global boundary Stratotype Section and Point for the Anthropocene series. *The Anthropocene Review*, 10(1), 146–176. <https://doi.org/10.1177/20530196221149281>
- McLaren, D. J. (1970). Time, life, and boundaries. *Journal of Paleontology*, 44(5), 801–815.
- McLaren, D. J. (1978). Dating and correlation, a review. In G. V. Cohee, M. F. Glaessner, & H. D. Hedberg (Eds.), *Contributions to the Geological Time Scale* (pp. 1–7). American Association for Petroleum Geologists. <https://doi.org/10.1306/st6398c1>
- Melchin, M. J., Jiayu, R., Gradstein, F., Koren, T. N., & Finney, S. C. (2004). Stability in stratigraphy. *Lethaia*, 37(1), 124–125. <https://doi.org/10.1080/00241160410005025>
- Mills, I. M., Mohr, P. J., Quinn, T. J., Taylor, B. N., & Williams, E. R. (2005). Redefinition of the kilogram: A decision whose time has come. *Metrologia*, 42(2), 71–80. <https://doi.org/10.1088/0026-1394/42/2/001>
- Molony, B. W., Jones, B., Lawrence, C. S., & Gouteff, V. A. (2006). Case 3267: Cherax tenuimanus Smith, 1912 (Crustacea, Decapoda, parastacidea): Proposed conservation of usage of the specific name. *Bulletin of Zoological Nomenclature*, 63(4), 231–235.
- Murchison, R. I. (1839). *The Silurian System, Founded on Geological Researches in the Counties of Salop, Hereford, Radnor, Montgomery, Caermarthen, Brecon, Pembroke, Monmouth, Gloucester, Worcester, and Stafford: With Descriptions of the Coalfields and Overlying Formations, Part 1*. London: John Murray.
- Newell, N. D. (1966). Problems of geochronology. *Proceedings of the Academy of Natural Sciences of Philadelphia*, 118, 63–89.
- Newell, N. D. (1972). Stratigraphic gaps and chronostratigraphy. In *Proceedings of the 24th International Geological Congress, Section 7: Paleontology* (pp. 198–204). Montreal, Quebec.
- Phillips, J. (1840). Palaeozoic series. In G. Long (Ed.), *The Penny Cyclopaedia of the Society for the Diffusion of Useful Knowledge* (Vol. 17, pp. 153–154). Charles Knight.
- Quinn, T. J. (1991). The kilogram: The present state of our knowledge. *IEEE Transactions on Instrumentation and Measurement*, 40(2), 81–85. <https://doi.org/10.1109/tim.1990.1032888>
- Remane, J. (2000). Should the golden spike glitter? — Comments to the paper of M.-P. Aubry et al. *Episodes*, 23(3), 211–213.
- Remane, J. (2003). Chronostratigraphic correlations: Their importance for the definition of geochronologic units. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 196(1–2), 7–18. [https://doi.org/10.1016/s0031-0182\(03\)00310-9](https://doi.org/10.1016/s0031-0182(03)00310-9)
- Remane, J., Bassett, M. G., Cowie, J. W., Gohbrandt, K. H., Lane, H. R., Michelsen, O., & Naiwen, W. (1996). Revised guidelines for the establishment of global chronostratigraphic standards by the International Commission on Stratigraphy (ICS). *Episodes*, 19(3), 77–81.
- Riordan, S. (2019). And how experiments begin: The International Prototype Kilogram and the Planck constant. In N. Courtenay, O. Darrigol, & O. Schlaudt (Eds.), *The Reform of the International System of Units (SI): Philosophical, Historical and Sociological Issues* (pp. 150–179). London: Routledge. <https://doi.org/10.4324/9781351048989>
- Rudwick, M. J. S. (1985). *The Great Devonian Controversy: The Shaping of Scientific Knowledge among Gentlemanly Specialists*. Chicago: The University of Chicago Press. <https://doi.org/10.7208/chicago/9780226731001.001.0001>
- Rudwick, M. J. S. (1996). Cuvier and Brongniart, William Smith, and the reconstruction of geohistory. *Earth Sciences History*, 15(1), 25–36. <https://doi.org/10.17704/eshi.15.1.h25k254u61374h00>
- Salvador, A. (Ed.). (1994). *International Stratigraphic Guide* (2nd ed.). Geological Society of America.
- Schindewolf, O. H. (1960). Stratigraphische Methodik und Terminologie. *Geologische Rundschau*, 49(1), 1–35. <https://doi.org/10.1007/bf01802389>
- Schindewolf, O. H. (1970). *Stratigraphie und Stratotypus*. Mainz: Verlag der Akademie der Wissenschaften und der Literatur.
- Schoch, R. M. (1989). *Stratigraphy: Principles and Methods*. New York: Van Nostrand Reinhold.
- Scott, G. H. (1960). The type locality concept in time-stratigraphy. *New Zealand Journal of Geology and Geophysics*, 3(4), 580–584. <https://doi.org/10.1080/00288306.1960.10420146>
- Sepkoski, D. (2017). The Earth as Archive: Contingency, narrative, and the history of life. In L. Daston (Ed.), *Science in the archives* (pp. 53–84). Chicago: The University of Chicago Press. <https://doi.org/10.7208/chicago/9780226432533.003.0002>
- Sepkoski, D. (2020). *Catastrophic Thinking*. Chicago: The University of Chicago Press.
- Simpson, G. G. (1940). Types in modern taxonomy. *American Journal of Science*, 238(6), 413–431.
- Smith, W. (1815). *A Delineation of the Strata of England and Wales, with Part of Scotland*. London: John Cary.
- Smith, A. G., Barry, T., Bown, P., Cope, J., Gale, A., Gibbard, P., Gregory, J., Hounslow, M., Kemp, D., Knox, R., Marshall, J., Oates, M., Rawson, P., Powell, J., & Waters, C. (2015). GSSPs, global stratigraphy and correlation. *Geological Society, London, Special Publications*, 404(1), 37–67. <https://doi.org/10.1144/sp404.8>
- Sylvester-Bradley, P. C. (1967). Towards an International Code of Stratigraphic Nomenclature. In C. Teichert, & E. L. Yochelson (Eds.), *Essays in Paleontology and Stratigraphy* (pp. 49–56). Lawrence: The University of Kansas Press.
- Tal, E. (2011). How accurate is the standard second? *Philosophy of Science*, 78(5), 1082–1096. <https://doi.org/10.1086/662268>
- Tal, E. (2017). A model-based epistemology of measurement. In N. Möbner, & A. Nordmann (Eds.), *Reasoning in Measurement* (pp. 233–253). London: Routledge. <https://doi.org/10.4324/9781781448717>
- Tal, E. (2018). Naturalness and convention in the international system of units. *Measurement*, 116, 631–643. <https://doi.org/10.1016/j.measurement.2017.09.008>
- Torrens, H. S. (2002). From d'Orbigny to the Devonian: Some thoughts on the history of the stratotype concept. *Comptes Rendus Palevol*, 1(6), 335–345. [https://doi.org/10.1016/s1631-0683\(02\)00064-7](https://doi.org/10.1016/s1631-0683(02)00064-7)
- Turner, D. (2019). *Paleoaesthetics and the Practice of Paleontology*. Cambridge University Press.
- Vai, G. B. (2007). A history of chronostratigraphy. *Stratigraphy*, 4(2/3), 83–97.
- Walker, M., Johnsen, S., Rasmussen, S. O., Steffensen, J.-P., Popp, T., Gibbard, P., Hoek, W., Lowe, J., Andrews, J., Björck, S., Cwynar, L., Hughen, K., Kershaw, P., Kromer, B., Litt, T., Lowe, D. J., Nakagawa, T., Newnham, R., & Schwander, J. (2008). The Global Stratotype Section and Point (GSSP) for the base of the Holocene Series/Epoch (Quaternary System/Period) in the NGRIP ice core. *Episodes*, 31(2), 264–267. <https://doi.org/10.18814/epiugs/2008/v31i2/016>
- Walsh, S. L. (2004). Solutions in chronostratigraphy: the Paleocene/Eocene boundary debate, and Aubry vs. Hedberg on chronostratigraphic principles. *Earth-Science Reviews*, 64(1–2), 119–155. [https://doi.org/10.1016/s0012-8252\(03\)00040-0](https://doi.org/10.1016/s0012-8252(03)00040-0)
- Walsh, S. L. (2005). The role of stratotypes in stratigraphy. Part 1: Stratotype functions. *Earth-Science Reviews*, 69(3–4), 307–332. <https://doi.org/10.1016/j.earscirev.2004.11.002>
- Walsh, S. L., Gradstein, F., & Ogg, J. (2004). History, philosophy, and application of the Global Stratotype Section and Point (GSSP). *Lethaia*, 37(2), 201–218. <https://doi.org/10.1080/00241160410006500>
- Wilson, J. A. (1959). Stratigraphic concepts in vertebrate paleontology. *American Journal of Science*, 257(10), 770–778. <https://doi.org/10.2475/ajs.257.10.770>
- Witteveen, J. (2015). Naming and contingency: The type method of biological taxonomy. *Biology and Philosophy*, 30(4), 569–586. <https://doi.org/10.1007/s10539-014-9459-6>
- Witteveen, J. (2018). Objectivity, historicity, taxonomy. *Erkenntnis*, 83(3), 445–463. <https://doi.org/10.1007/s10670-017-9897-z>
- Witteveen, J. (2021). Taxon names and varieties of reference. *History & Philosophy of the Life Sciences*, 43(2), 78. <https://doi.org/10.1007/s40656-021-00432-4>