

Forthcoming in *Biology & Philosophy*

Understanding Scientific Types: Holotypes, Stratotypes, and Measurement Prototypes

Alisa Bokulich
Department of Philosophy
Boston University
abokulic@bu.edu
<https://orcid.org/0000-0002-9406-3904>

Abstract:

At the intersection of taxonomy and nomenclature lies the scientific practice of typification. This practice occurs in biology with the use of holotypes (type specimens), in geology with the use of stratotypes, and in metrology with the use of measurement prototypes. In this paper I develop the first general definition of a scientific type and outline a new philosophical theory of types inspired by Pierre Duhem. I use this general framework to resolve the necessity-contingency debate about type specimens in philosophy of biology, to advance the debate over the myth of the absolute accuracy of standards in metrology, and to address the definition-correlation debate in geology. I conclude that just as there has been a productive synergy between philosophical accounts of natural kinds and scientific taxonomic practices, so too there is much to be gained from developing a deeper understanding of the practices and philosophy of scientific types.

Keywords:

Nomenclature; taxonomy; type specimen; holotype; stratotype; measurement prototype; Pierre Duhem; typification; natural kinds; classification; kilogram redefinition; species; geology; metrology

1. Introduction

When it comes to scientific taxonomy and nomenclature, the predominant philosophical focus has been on the issue of natural kinds. While this focus on natural kinds has been philosophically productive, there is another component to some taxonomic and nomenclatural practices that has been largely overlooked by philosophers of science, and that is what we might call 'scientific types' and the practice of typification. Scientific types appear across the sciences, though they are only a part of some classificatory practices. Scientific types are found, for example, in biology in the context of holotypes, in the geosciences as stratotypes, and in physics in the form of measurement prototypes (standards). It is worth underscoring at the outset that scientific types are not kinds—they are instead concrete particulars: a particular biological specimen, a particular section of rock, a particular hunk of machined metal. My aim in this paper is to draw together these different cases, identify a common core in their characterization

and use, and abstract from that common core a new philosophical account of scientific types.¹ I will do this through a detailed examination of scientific practice: examining the use of holotypes in biology in section 2.1, the use of stratotypes in geology in section 2.2, and the use of measurement prototypes in section 2.3. In section 2.4, I will extract from these three cases a general characterization of scientific types and propose the following general definition: *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*. I will then demonstrate the philosophical value of this new general conception of scientific types by using it to advance three debates in the philosophy of science: the necessity-contingency debate about holotypes, the absolute accuracy debate over measurement standards, and the definition-correlation debate about stratotypes.

One of the few contexts in which scientific types have received some philosophical attention is in the philosophy of biology, where there has been a flurry of papers over an apparent paradox concerning holotypes (e.g. Levine 2001): specifically, the debate has been over whether a holotype (i.e., type specimen) belongs contingently or necessarily to the species it names. In section 3 I briefly review this necessity-contingency debate, arguing that it can be profitably advanced by placing holotypes within the broader category of scientific types that I define.

In section 4.1, I turn to the philosophical literature on measurement prototypes, and review the debate in metrology over the absolute accuracy of measurement standards. I argue that what Eran Tal (2011) calls the "myth of the absolute accuracy of standards" should be rejected not just for measurement prototypes, but also for other scientific types. This insight also provides a foundation for addressing the debate in geology over whether definition precedes or follows correlation (section 4.3). I examine the conditions under which the scientific institutions governing the use of these scientific types (viz., the International Bureau of Weights and Measures, the International Commission on Zoological Nomenclature, and the International Commission on Stratigraphy) allow for their revision. A common thread among these various provisos suggests what I call a Duhemian approach to scientific types, which I outline in the final section 5. As we will see, the key insight from Duhem is his distinction between a stipulative "symbolic" definition and a "common sense" definition that allows him to reject the conventionalist argument for the "absolute accuracy" of stipulative definitions. I will show how a similar move can be adapted to the context of scientific types.

2. Scientific Types

The terms 'type' and 'typification' arise from the scientific discourse in these subfields, and hence are the terms I adopt here, even though they can lead to an initial confusion in philosophical circles, where the term 'type' usually takes on a different—and indeed opposite—meaning. I have tried to flag the technical meaning in this context by using the phrase 'scientific type.' As it

¹ The general philosophical attitude that guides my approach to scientific types is similar to a particular strain of philosophical work on natural kinds, namely the practice-oriented approach exemplified by Kendig's (2015) work on *kinding* (which focuses on how natural kind categories are identified, constructed, maintained, and even revised); and the view that natural kinds are neither simply a matter of discovery, nor simply a matter of conventional stipulation, but rather involve elements of both (e.g., LaPorte 2004, Bokulich 2014). However, as my focus in this paper is on scientific types and not natural kinds, I will not discuss this literature further here.

is used here, a scientific type is neither synonymous with 'kind' or 'category,' nor is it to be contrasted with token, as in the familiar type-token distinction. Indeed scientific types are a kind of token. Although the term 'typification' is used by itself, 'type' usually appears as either a suffix (as in holotype, stratotype, prototype), or as an adjective (as in type specimen). Because the discussion of scientific types has hitherto taken place only within one of these specific scientific fields, there has not yet, to my knowledge, been a general definition offered for scientific types more broadly. The definition I develop here is the following: *A scientific type is a concrete individual object that serves as an objective standard of reference for, and realization of, the definition or taxon category it names.* As we will see, this definition captures what is common among the prima facie diverse notions of holotypes, stratotypes, and measurement prototypes.

2.1 Holotypes

The use of holotypes is a long-standing practice in biology that is governed by the International Commission on Zoological Nomenclature (ICZN). According to the ICZN's code (the *International Code of Zoological Nomenclature*), a holotype is defined as "the single specimen upon which a new nominal species group taxon is based in the original publication" (Article 73). When a new species is discovered, that specimen is described in detail in a publication and named, thereby leading to the recognition of a new taxonomic category. That individual specimen is then designated the holotype for that species and is stored in a museum in perpetuity, so that it is available to subsequent researchers for comparison and consultation. If one goes behind the scenes at a museum such as the Museum of Comparative Zoology at Harvard University, one will typically find many such holotype specimens, conventionally marked with a red ribbon. Thus, in principle, for every named species known to biology, there should be somewhere in the world an individual marked specimen that serves as that species' holotype.

The specimen that is designated a holotype need not be a complete organism: it can be a part of an animal (such as a femur or tooth), a fossil (either a natural replacement or a natural impression), or even a microscope "type slide," depending on the species in question. The idea behind holotypes is to ground the name of a new taxon in a concrete object that serves as an objective standard, though there is no requirement that the specimen in question be "typical," in the sense of representing an average range of variation, a particular sex, or even a particular life stage of a species. Given the limitations of a single specimen holotype, additional "paratypes" can also be designated, for example representing a different sex or life-stage than the holotype. These paratypes are conventionally designated by a blue ribbon in museum collections.

To better understand the practice of holotypes it is helpful to highlight some key features of the ICZN code's *principle of typification*:

The fixation of the name-bearing type of a nominal taxon provides the *objective standard of reference* for the application of the name it bears. No matter how the boundaries of a taxonomic taxon may vary in the opinion of zoologists the valid name of such a taxon is determined . . . from the name-bearing type(s) considered to belong within those boundaries. . . . Once fixed, *name-bearing types are stable and provide objective continuity in the application of names.* Thus the name-bearing type of any nominal taxon, once fixed in conformity with the provisions of the Code, *is not subject to change except* in the case of nominal genus-group taxa as provided in *Article 70.3.2*, of nominal species-group taxa as provided in *Articles 74 and 75*, and by use of the plenary power of the Commission [*Art. 81*]. (ICZN Code, Article 61.1; emphases added).

There are three features of this principle of typification that I want to highlight and have indicated by the italics above. The first is that the central function of a holotype is to serve as an objective standard of reference, and in particular, a standard of reference for the application of the name it bears. This role of being an objective standard of reference is one that I argue characterizes all scientific types.

A second key point in the ICZN's principle of typification is that name-bearing types (holotypes) function to provide a stability and coherence across the scientific community. For effective communication, it is important that scientists are using terms in the same way, and know when they are referring to the same taxon. As has been emphasized in the scientific literature on holotypes,

species names provide the most consistent anchor to which all taxonomic, ecological, molecular, conservation, and other biologically relevant data are attached. Legal protection and policy are also linked with names on the assumption that the groups indicated by the names are consistent through time and among places. . . . Medical and veterinary implementation requires communication about unambiguous identifications. (Pyle and Michel 2008, p. 40)

This is a vast community to coordinate, and scientific types are an effective tool by which to secure the needed consistency and stability. This does not, of course, imply that our knowledge about these species or other taxonomic categories is fixed. Scientific knowledge is continually changing, but in order to track that change there has to be stability and coherence in how names and terms are applied. This function of providing stability and coherence within an evolving scientific practice is, as we will see, a general feature of all scientific types.

Finally, a third feature of the principle of typification—and the one most relevant to the necessity-contingency debate in the philosophy of biology that I will discuss below in section 3—is the clause that says a holotype is not subject to change *except* in cases of Articles 70.3.2, 74, 75, and 81. In other words, a holotype is taken conventionally to be infallible in picking out the taxon that it names—except in a variety of special circumstances. I will return to discuss what those special circumstances are and their philosophical implications in section 4.2.

It is important to emphasize that a holotype does not define a taxon in the sense of determining how a given species is delimited. That is a substantive empirical question whose answer evolves as scientific research progresses, and opinions about "lumping" versus "splitting" evolve. As former ICZN executive secretary Ellinor Michel and Richard Pyle note, the "process of typification allows the name to be tied to a physical standard (and hence provides an objective basis for identifications), but leaves room for taxonomy to change" (Pyle and Michel 2008, p. 41). They offer the following diagram (Fig. 1) to illustrate the way in which type specimens function as a bridge between taxonomy and nomenclature, having a foot in each.

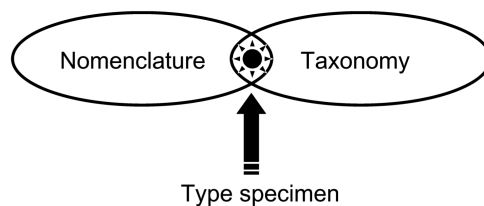


Figure 1: Type specimens as lying at the intersection of taxonomy and nomenclature. Reproduced from Pyle and Michel 2008, Figure 2.

To summarize this subsection, then, we have seen that holotypes have three key features that can be found in the ICZN's principle of typification: First, its function is to serve as an objective standard of reference for the application of the name it bears; second, its purpose is to secure a stability and coherence of practice across a community of users; and third, a type specimen's role as standard is not typically subject to revision, but can in fact be revised in exceptional (though not necessarily rare) circumstances. As I will show next, these three features can be found in the guidelines governing stratotypes in geology as well.

2.2. Stratotypes

Within geology, chronostratigraphy is the study and organization of the stratigraphic (rock) record in relation to geologic time. Geologic time, which encompasses the some 4.5 billion years of Earth's history, is itself organized into a hierarchical set of named units known as the Geological Time Scale (GTS), the ongoing construction of which is a vast interdisciplinary effort. There is a duality inherent in the GTS in that the named divisions refer both to *chronostratigraphic* ("time-rock") units, that is, all and only those rocks or strata formed during a certain interval of time, and to *geochronologic* units, which are purely temporal ("time") units. So, for example, the 'Jurassic' refers both to the geochronologic ("time") unit, the Jurassic Period, and to the chronostratigraphic ("time-rock") unit, the Jurassic System (the set of all rocks formed during the Jurassic Period). This duality is preserved through all taxonomic levels of the GTS hierarchy. Much like the biological taxonomic hierarchy of species-genus-family-order-class, the chronostratigraphic ("time-rock") hierarchy includes stage-series-system-erathem-eonothem, which parallels the geochronologic (time) hierarchy of age-epoch-period-era-eon. Although somewhat awkward, this duality is arguably important insofar as the stratigraphic record (and the fossils or other materials it contains) is the material evidential base from which geologic time is inferred.

While it has long been recognized that at any given location, deeper stratigraphic layers correspond to periods further back in time, precisely which periods of geologic time are represented in the strata can vary dramatically from place to place. Depending on local factors such as rates of erosion and uplift, the stratigraphic layers exposed at any given location can represent sediments deposited 50 years ago (in which case those rocks would be members of the Quaternary System), 150 million years ago when the stegosaurus dinosaurs roamed (making them part of the Jurassic System), or even 500 million years ago when the trilobites flourished (hence members of the Cambrian System). Thus, the project of chronostratigraphy is to sort the various stratigraphic rocks around the world into the correct taxonomic "time-rock" bin, thereby providing a complete chronostratigraphic classification.

In order for such a classification to be scientifically useful, there needs to be a clear and univocal nomenclature and understanding of precisely where the boundaries, marking the end of one chronostratigraphic unit and the beginning of another, fall. For example, the infamous K-Pg boundary, which is the particular stratigraphic layer of rock recording the Chicxulub impact that killed off the nonavian dinosaurs, marks the end of the chronostratigraphic unit known as the Cretaceous System and the beginning of the next chronostratigraphic unit, the Paleogene

System.² It is rare, however, to find such a clearly marked, relatively isochronous, and globally present event; hence in most cases, the precise boundaries between chronostratigraphic units need to be drawn at a particular stratigraphic layer somewhat conventionally. The method that has been adopted by stratigraphers is the “GSSP” or “golden spike” method.

GSSP stands for Global Boundary Stratotype Section and Point, and it is a kind of boundary stratotype. The *Guidelines and Statutes of the International Commission on Stratigraphy (ICS)* defines GSSPs as follows:

This Boundary Stratotype Section and Point is the *designated type* of a stratigraphic boundary identified in published form and marked in the section *as a specific point in a specific sequence of rock strata* and constituting the *standard for the definition and recognition* of the stratigraphic boundary between two named global standard stratigraphic (chronostratigraphic) units. (Cowie et al. 1986, p. 5; emphases added)

As we see in this definition, a GSSP is a concrete individual (i.e., a specific point in a particular series of rock strata at a particular locale, and typically marked and labeled with a metal "spike") that serves as a standard for the definition and recognition of the beginning of a named chronostratigraphic unit. The decision to have the type designate a chronostratigraphic *boundary* (such as the Permian-Triassic boundary), rather than a whole chronostratigraphic unit (e.g., the Triassic System) arises from the desire to have a system of chronostratigraphic units that completely covers the Geologic Time Scale without any gaps or overlapping units. So the GSSP at the Permian-Triassic boundary (which is marked at a particular point in the base of Bed 27c of Meishan section D, in Southern China) marks the beginning of the Triassic chronostratigraphic unit and the end of the Triassic is marked by the boundary stratotype defining the beginning of the next chronostratigraphic unit, the Jurassic (which is a GSSP "golden spike" placed in the Kuhjoch section of the Karwendel mountains in the Northern Calcareous Alps, Austria).³

As the ICS guidelines clarify, the GSSP defines not just the chronostratigraphic ("time-rock") unit, but is also taken to define a moment of time: "Insistence on a Boundary Stratotype Point is in order to define without a doubt an instant of geologic time" (Cowie et al. 1986, p. 5). In other words, the GSSP also defines the moment of time at which the corresponding geochronologic ("time") unit begins. So, for example, the GSSP in Meishan, Southern China defines not just the beginning of the Triassic System but also the start of the Triassic Period. Importantly, this moment of time is not defined in terms of a numerical "absolute" time age, but rather is defined stratigraphically as a precise moment in "relative" time. If there is suitable material at the boundary to be radiometrically dated, then an absolute age may be associated with the boundary, but it is not taken to define the moment in time. Indeed the absolute ages associated with boundaries like the Permo-Triassic are fallible, and are often revised as radiometric dating methods are iteratively improved (for a philosophical discussion see Bokulich 2020).

² It thus by default also marks the boundary of any smaller chronostratigraphic units, in this case, the end of the Upper Cretaceous Series and beginning of the Paleocene Series, and the end of the Maastrichtian Stage and beginning of the Danian Stage.

³ A single GSSP is placed for each chronostratigraphic boundary, but there is no requirement that successive GSSPs be anywhere near each other. Hence, as discussed below, the scientific challenge is to use empirical evidence to extend that point into an isochronous horizon around the world.

Stratotype points are thus the only places where geoscientists know (by definition) that time-rock and time coincide. The scientific challenge then is to extend this chronostratigraphic boundary around the world. Although there is some degree of arbitrariness in where exactly a GSSP is placed, GSSP locations are chosen with two guiding criteria: first, the GSSP should be placed in "continuous" sections with no breaks or unconformities in the sedimentary record, which would represent "missing" time of unknown duration, and hence not uniquely define a moment of time. And, second, GSSPs should be placed where there are clear markers that facilitate global correlatability. Biostratigraphic events, like the lowest (first) occurrence or highest (last) occurrence of a fossil species are the most common "primary markers" for chronostratigraphic boundaries in the Phanerozoic, but correlatability is improved by having many "secondary markers," which can be other biostratigraphic events, magnetostratigraphic reversals (records of flips in Earth's magnetic north and south poles), stable isotope excursions, etc. The stratotype for the beginning of the Triassic, for example, has as its primary marker the lowest occurrence (first appearance datum) of the conodont fossil *Hindeodus parvus*, which is an extinct eel-like species. And the end of the Triassic (beginning of Jurassic) has as its primary marker the first appearance of the smooth *Psiloceras spela* group Ammonite. These biologic markers are used to identify and globally extend, but not define, the stratotype. Nonetheless, this epistemic dependence of most stratotypes on biological index fossils makes them subject to many traditional philosophical issues in biology, such as the species problem (for a discussion see McGowran 2005).

As we have seen, stratotypes (GSSPs) provide a material standard of reference for the definition and recognition of the base of a named chronostratigraphic (and corresponding geochronologic) unit. The stratotype thus fixes the reference of that named chronostratigraphic unit, providing a taxonomic class into which stratigraphic rocks all over the world can be sorted (e.g., the class of all rocks formed during the Cambrian Period). Although designating a stratotype is an act of stipulation, determining which other rocks belong in that taxon is of course a substantive empirical project. Like biological holotypes, stratotype 'specimens' must be in publically accessible localities for scientific consultation, and a ratified scientific type requires not just the physical specimen, but also its detailed description in an official scientific publication.

More broadly, the purpose of stratotypes is to secure a stability and coherence of scientific nomenclature and practice across a diverse community of scientists. As many geoscientists lamented, prior to the institution of stratotypes, "names of the supposedly 'standard' periods, epochs, and ages were used in significantly different ways by geologists in different parts of the world" (Walsh et al. 2004, p. 202). By providing an objective anchor for a taxonomic name in a material standard governed by an international scientific organization, scientific types are an effective tool for securing the necessary coherence. Insofar as stratotypes, are conventionally taken to be permanent, they also function to provide stability. However, as a recent paper from UK's Stratigraphy Commission notes, the situation can be a little more complicated in practice.

GSSPs, once ratified, are generally regarded as fixed. There are in fact a few conditions which permit revision of an existing GSSP such as its having been destroyed or become inaccessible. A GSSP can also 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.' (Smith et al. 2015, p. 40, with quotation from Remane et al. 1996, p. 80)

Thus, although stratotypes are taken to be permanent, fixing the reference of a named chronostratigraphic unit base and defining the corresponding instant of geologic time, the guidelines of the relevant international organization (here the ICS) provide for various circumstances under which the scientific type can be revised, as we also saw in the case of holotypes. I will return to discuss what those special circumstances are for stratotypes in section 4.3.

Once again, we see in the case of stratotypes that a scientific type's function is to, first, serve as an objective standard of reference for the application of the name it bears; second, secure a stability and coherence of practice across a community of users; and third, that a scientific type's role as standard is not typically subject to revision, but can in fact be revised in exceptional circumstances. Before elaborating this general characterization, however, I will examine a third and final example of scientific types: measurement prototypes.

2.3 Prototypes: Measurement Standards

Although we often take for granted the units by which we measure world, such as the meter and the kilogram, there is a highly complex field known as metrology that is devoted to making sure that in meaning and practice, we are all in agreement about these units. The international organization responsible for governing these various units of measurement is the International Bureau of Weights and Measures, or BIPM (after the French Bureau International des Poids et Mesures). BIPM describes its mandate as "providing the basis for a single, coherent system of measurements to be used throughout the world" (BIPM 2019b). One of the primary ways that BIPM has traditionally secured this stability and coherence is through the use of physical measurement standards known as prototypes. Within metrology, a prototype is defined as a concrete object or artefact that serves as the basis of the definitive definition or realization of its unit of measure.⁴

One example of a prototype is the *International Prototype Meter*, which is a particular metal bar formed out of a platinum and iridium alloy and conserved at the BIPM in Sèvres, France. This particular bar served as the basis of the definition of the meter from 1889 until 1960. Recognizing the influence of temperature, pressure, and gravity on this metal artefact, the definition of the meter was more precisely formulated in 1927 as follows:⁵

The unit of length is the metre, defined by the distance, at 0°, between the axes of the two central lines marked on the bar of platinum–iridium kept at the Bureau International des Poids et Mesures and declared Prototype of the metre by the 1st Conférence Générale des Poids et Mesures, this bar being subject to standard atmospheric pressure and supported on two cylinders of at least one centimetre diameter, symmetrically placed in the same horizontal plane at a distance of 571 mm from each other. (BIPM 2019a, *SI*, 9th ed., p. 159).

This prototype meter could then be compared to secondary standards of similarly constructed bars, which were distributed around the world for the standardization of measurements. Only the original prototype bar stored in Sèvres, however, served as the basis of the definition of a meter.

⁴ By 'realization' one means the conversion of a theoretical quantity to reality.

⁵ Even the original 1889 definition of the meter specified that it was the "prototype, at the temperature of melting ice, [that] shall henceforth represent the metric unit of length" (<https://www.bipm.org/en/CGPM/db/1/1/>).

Whatever the length of this scientific type was (under the above conditions), it was by definition exactly one meter.

Another example of a measurement prototype is the *International Prototype Kilogram* (known as IPK or the 'Big K'). The IPK is a concrete artefact made of a platinum-iridium alloy, machined into a right-circular cylinder, and stored in the BIPM headquarters in Sèvres, France. Also established in 1889, the IPK served as the definition and basis of the realization of the kilogram:

The kilogram is the unit of mass; it is equal to the mass of the international prototype of the kilogram. . . . It follows that the mass of the international prototype of the kilogram is always 1 kilogram exactly." (BIPM 2006, *SI* 8th ed., p. 112)

Although the stipulative definition takes the mass of the IPK to always be exactly 1 kilogram, the real world is little messier (literally). After giving the above definition of the kilogram in terms of the mass of the prototype, they immediately go on to note the following:

However, due to the inevitable accumulation of contaminants on surfaces, the international prototype is subject to reversible surface contamination that approaches 1 µg per year in mass. For this reason, the CIPM [Comité International des Poids et Mesures] declared that, pending further research, the reference mass of the international prototype is that immediately after cleaning and washing by a specified method." (BIPM 2006, *SI* 8th ed., p. 112)

This (cleaned) IPK thus served as the realization of the definition of the kilogram and the reference standard against which other, secondary national kilogram standards could be calibrated.

Both the International Prototype Meter and the International Prototype Kilogram are examples of scientific types that serve as an objective standard of reference for, and realization of, the definition of their kind. Their purpose, as we saw, is to secure a coherence and stability of practice across a community of users. Finally, a prototype is intended to have a permanence and not be subject to revision. Like holotypes and stratotypes, however, they can in fact be revised in exceptional circumstances. The prototype meter's status as a type definition was revised in 1960 and the IPK's status as the type definition of the kilogram was revised in 2019. I will return to discuss the circumstances that prompted these revisions in section 4.1.

2.4 Understanding Scientific Types and Their Functions

Although occurring in different fields and involving subtle differences, there is a common focal function and status to holotypes, stratotypes, and measurement prototypes, that I argue unites all three under the rubric of 'scientific type'. To capture this commonality, I propose the following general definition: *A scientific type is a concrete individual object that serves as an objective standard of reference for, and realization of, the definition or taxon category it names.* This general definition is derived from, and supported by, the detailed analyses of the three cases given in the preceding subsections.

As we have seen, all three types involve tying a theoretical kind category to a single physical object that has a privileged status as the definitive anchor for the name or definition of that kind of thing. All three have as their focal function serving as a publicly accessible standard of reference for their kind. Furthermore, all three are tightly governed by the relevant international scientific organization and are taken by stipulation to be unrevisable, except by the

provisos of these organizations (as will be discussed in more detail in section 4). More broadly, all three of these scientific types serve the purpose of securing a coherence and stability of practice across their relevant scientific communities, despite the recognition of an ever-evolving body of scientific knowledge. There is a tension that should be noted here 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. This tension motivates the Duhemian approach to scientific types I develop in section 5.

The general definition and account of scientific types I develop here is valuable not just for its own sake, but also because it provides a new framework within which to advance a number of recent debates in the philosophy of science and scientific practice. I will show its fertility more concretely by using it to address the following three debates: the necessity-contingency debate in philosophy of biology, the absolute accuracy of standards debate in philosophy of metrology, and the definition-correlation debate in stratigraphy. With this deeper understanding of scientific types in hand, let us begin by first examining the philosophical debate that has arisen over biological holotypes and whether they belong necessarily or contingently to the species that they name, and see whether by placing holotypes within this broader class of scientific types, this debate can be further advanced.

3. Holotypes and the Necessity-Contingency Debate

One context in which the scientific practice of typification has attracted some philosophical attention is the necessity-contingency debate over biological holotypes. The debate began with a paper by Alex Levine (2001) in which he identified a paradox or "contradiction" about type specimens. Levine introduces this paradox in the context of another, well-known debate over whether species should be understood as *kinds* or *individuals* (although the paradox arguably arises regardless of which position one takes on that species debate). In advocating the species-as-individuals view, David Hull had noted that the practice of fixing species names through type specimens (holotypes) suggests that they function as rigid designators: species names "cannot change their reference, although we can find out that we are mistaken about what we thought their reference was" (Hull 1982, p. 492). Since the nomenclatural principle of priority (ICZN Code, Article 23) means that the name follows the type specimen regardless of which species it is placed in, it looks like the relation between a type specimen and the species it typifies is necessary. However it also seems that a type specimen, like any other organism, belongs only contingently to its species. Levine summarizes the paradox as follows:

We can only conclude that, *qua organism*, the type specimen belongs to its respective species contingently, while *qua type specimen*, it belongs necessarily. But this statement merely codifies the contradiction without resolving it. (Levine 2001, p. 334)

Like any good paradox, sorting out whether holotypes belong necessarily or contingently to their species turned out to be surprisingly subtle question, giving rise to a philosophical debate.

The first proposed solution came from Joseph LaPorte, who argues that the paradox stems from a failure to distinguish *de dicto* and *de re* readings of necessity. He explains,

The *de dicto* sentence 'Necessarily, any species with a type specimen contains its type specimen' is true if and only if . . . 'Any species with a type specimen contains its type specimen'. This does seem to be true in any possible world. . . The *de re* sentence 'Any species with a type specimen necessarily contains its type specimen' is true if and only if

any species with a type specimen in the *actual* world contains that very type specimen in every *possible* world. This reading does not seem to be true. (LaPorte 2003, p. 587)

While he concludes that a type specimen belongs only contingently to its species, he goes on to argue that it can in fact be known a priori that it belongs to it, hence providing an example of contingent a priori knowledge. Using the example of Rover as the holotype for *Canis familiaris*, he continues, "Although it is contingent that *Canis familiaris* contains Rover, it is a priori certain that *Canis familiaris* does contain Rover since Rover is, in the actual world, the type specimen for *Canis familiaris*" (LaPorte 2003, p. 587). Of course what is known a priori is known with certainty, but that does not mean it is necessary. Hence that a type specimen belongs to its species is, he argues, an example of contingent a priori knowledge.

In a follow-up article, Matthew Haber agrees that type specimens belong only contingently to their species, but denies that it can be known a priori. More specifically, he rejects the view that even *de dicto* necessity holds. He argues, "*de dicto* necessity fails, and it fails in the actual world. This happens every time a type specimen is misidentified—something that should not be possible if type specimens belong of *de dicto* necessity to their species" (Haber 2012, p. 774). He supports this view by appealing to an actual case from scientific practice, where a type specimen was declared not to belong to the species that it named. Very briefly, the case involved two subspecies of the garter snake species known as *Thamnophis sirtalis* (*T.s.*): One was the common California Red-Sided Garter Snake, *T.s. infernalis*, and the other was the endangered San Francisco Garter Snake, *T.s. tetraenia*. In the mid-1990s researchers discovered that the holotype specimen for the common *T.s. infernalis* was in fact a member of the endangered *T.s. tetraenia*. While the nomenclatural principle of priority (Article 23) prescribes that the name *T.s. infernalis* should in such cases be transferred to the endangered *T.s. tetraenia*, researchers successfully petitioned (using Article 75.6) that the names be conserved for each, and that a new holotype (i.e., a neotype) be designated *T.s. infernalis*. The upshot, as Haber notes, is that we cannot know a priori that a type specimen belongs to the species it names, since in this actual case, the holotype for the common *T.s. infernalis* was not in fact a member of that species. I will come back in a moment to discuss the ICZN rules that permit these sorts of revisions to holotypes, but before doing so, there is one more twist in this debate that must be examined.

While the previous two papers seemed to be converging on the view that holotypes belong only contingently to their species, the most recent paper in this debate by Joeri Witteveen argues that is a mistake. More specifically, he argues *pace* LaPorte and Haber that there is *no* sense in which type specimens belong contingently to the species they name; rather they belong necessarily. While Witteveen raises a number of criticisms, it is his critique of Haber that I want to specifically focus on. He writes,

To falsify *de dicto* necessity, it is not sufficient for a type specimen to fail to belong to the taxon for which it *formerly* served as name-bearer. Instead, it must be possible for a type specimen to fail to belong to the taxon for which it *actually* serves as name-bearer. (Witteveen 2015, p. 580).

According to Witteveen, such a situation is not possible. He offers the following helpful diagram, which outlines three scenarios.

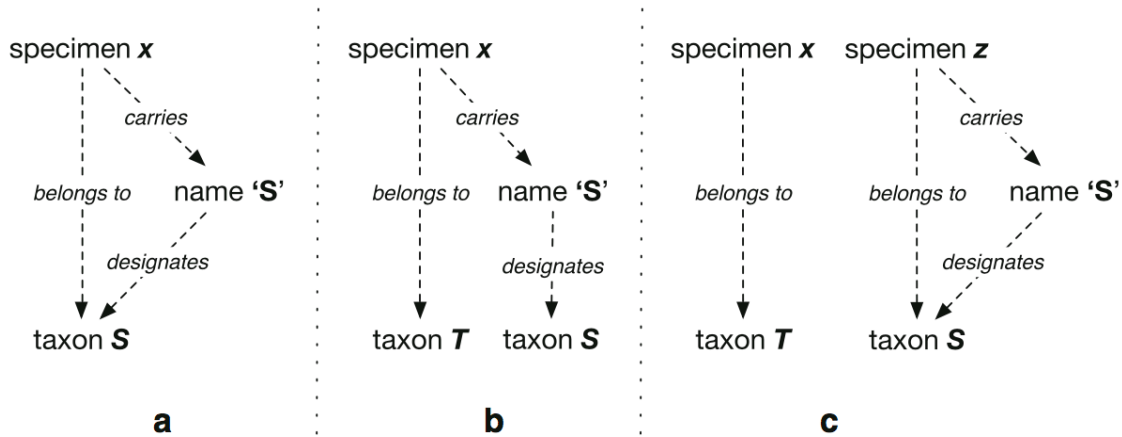


Figure 2: Figure from Witteveen (2015, p. 581) describing three different scenarios regarding the relationship between a specimen *x*, its name 'S', the taxon that name designates, and the taxon to which the specimen belongs.

In the first scenario, *a*, the type specimen works as intended: specimen *x* carries the name 'S' that designates taxon *S*, and *x* in fact belongs to taxon *S*. The second situation, *b*, is the situation that Haber's example of the California garter snake was intended to describe: specimen *x* carries the name 'S' (*T.s. infernalis*) that designates taxon *S* (*T.s. infernalis*), but the specimen *x* in fact belongs to taxon *T* (*T.s. tetraenia*). The third scenario, *c*, describes the situation after the old holotype was set aside and a neotype was designated for *T.s. infernalis*. With this diagram in hand, Witteveen then notes

This leaves us with the question whether it is possible *in principle* to arrive at the situation depicted in Fig. [b]. . . . [Is it] possible for a type specimen to belong to one taxon while carrying the name for another taxon? The answer has to be 'no.' (Witteveen 2015, p. 581)

According to Witteveen, the type method is infallible: it is impossible for the name of a type specimen to designate a different taxon than that to which it belongs. Thus he concludes that *de dicto* necessity holds.

Although Witteveen provides an insightful analysis of the necessity-contingency debate and method of typification, I will argue that he is mistaken. In particular, he is wrong to think that it is impossible to arrive at the situation described in scenario *b*, where a type specimen carries the name designating one taxon, but in fact belongs to another. More generally, I will argue that his argument falls prey to a version of what has been called the "myth of the absolute accuracy of standards." As will become clear in section 5, Witteveen is essentially making the same sort of mistake that Duhem criticized the conventionalists for making.

4. Myth of the Absolute Accuracy of Standards

Is the method of typification infallible as Witteveen argues? Although types are *taken by convention* to be infallible, scientists are fully aware that *in practice* types are in fact fallible. Although Witteveen is correct that the type method is designed to be typically insulated from such errors in taxonomic practice, I will argue that it is not in fact immune. A full defense of this will require two parts: first, an understanding of the various conditions under which scientific types can be revised or rejected; and second, an account of how scientists are able to arrive at the

sort of situation described in scenario *b* above, that Witteveen claims is impossible. I begin by reviewing some of the conditions and provisos that allow a scientific type to be recognized as inadequate, which will lay the foundation for the Duhemian approach to scientific types that I outline in section 5. In the context of measurement prototypes, these conditions of inadequacy have been used to argue against what Eran Tal calls the "myth of the absolute accuracy of measurement standards." I will argue that we should extend the recognition and rejection of this myth to all scientific types.

4.1 Revising Prototypes

In the context of his work in the philosophy of metrology, Eran Tal has identified what he calls the *myth of the absolute accuracy of standards*. He writes,

A common philosophical myth states that the meter bar in Paris is exactly 1 meter long. . . . One variant of the myth comes from Wittgenstein. . . . Kripke . . . develops a variant of the same myth by stating that the length of the bar at a specified time is rigidly designated by the phrase 'one meter.' Neither of these pronouncements is easily reconciled with the 1960 declaration of the General Conference on Weights and Measures, according to which 'the International Prototype does not define the metre with an accuracy adequate for the present needs of metrology'. (Tal 2011, pp. 1082-83)⁶

Tal's point is that scientific practice in metrology has routinely been able to make sense of the fact that a prototype standard can fail to be a sufficiently accurate realization of the unit it is used to define. Tal's analysis in the above paper turns to the example of the temporal standard for the second, showing how the myth also fails in that context, but here I will instead illustrate the point with the example of the prototype kilogram.

The International Prototype Kilogram (IPK) is a particularly instructive case because it is an example of a prototype that was recently judged to provide an insufficiently accurate basis for the definition of its unit of measure, leading to the formal redefinition of the kilogram in May of 2019.⁷ If the kilogram was defined to be whatever the IPK weighed, then how could it be judged inadequate? To what independent standard could one appeal to make such a judgment? If there were such an independent standard, then it would be possible to be in a situation analogous to Witteveen's scenario *b*: The mass of the platinum-iridium artefact in Sèvres carries the name '1.00 kg' which then designates the class of things having the property of a one kilogram mass. However, if there were an independent standard, then we would be able to say that the mass of the platinum-iridium artefact in Sèvres belongs, for example, to the class of things that are 0.99999995 kg. On Witteveen's reading of scientific types, there can be no such independent standard, and hence there could never be a determination that the IPK is anything other than 1.00 kg.

⁶ Tal notes that while Wittgenstein and Kripke helped spread this myth, their own positions were arguably more subtle.

⁷ As of May 20th, 2019 the kilogram is now "defined by taking the fixed numerical value of the Planck constant h to be $6.626\,070\,15 \times 10^{-34}$ when expressed in the unit J s, which is equal to $\text{kg m}^2 \text{s}^{-1}$, where the metre and the second are defined in terms of c and $\Delta\nu\text{Cs}$ " (BIPM 2019a, p. 131).

Although the IPK’s mass is stipulated to be exactly one kilogram, metrologists have never had a problem making sense of the fact that as a material artefact in the real world, its mass varies over time. Even before the redefinition of the kilogram, while the IPK was taken by definition to be exactly 1 kg with zero error, metrologists knew on the basis of background knowledge, experience, and theory that the IPK absorbs contamination from the atmosphere (even when stored under three bell jars). Hence, in the years between 1939 and 1946 they developed the BIPM cleaning method for the IPK and its sisters (Girard 1990). It was known that cleaning would remove between $5\mu\text{g}$ – $60\mu\text{g}$, and that it would begin to gain mass again right afterwards. As one source explains,

“After cleaning—even when they are stored under their bell jars—the IPK and its replicas immediately began gaining mass again. The BIPM even developed a model of this gain and concluded that it averaged $1.1\mu\text{g}$ per month for the first 3 months after cleaning and then decreased to an average of $1\mu\text{g}$ per year thereafter.” (IPK, Wikipedia; see also Girard 1990).

This result is depicted in Figure 3 below, which shows the change of mass of the IPK (designated \mathfrak{K}) and its “identical” sister copies.

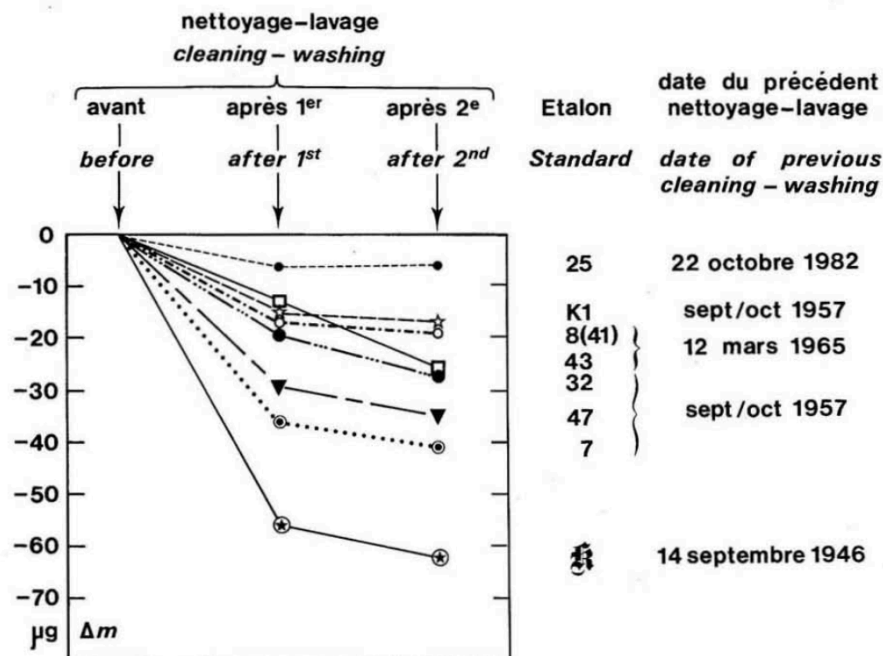


Fig. 3: The change of mass Δm of the International Prototype Kilogram (\mathfrak{K}) compared to its sister copies after a first and second cleaning. From Girard 1990, Figure 15.

As we see in this figure, even in 1990 when the IPK definition of the kilogram was in full sway, metrologists could make sense of the changes in the IPK’s mass under different conditions.

One source, or component, of an independent standard is precisely this network of background knowledge, experience, and theory that recognizes the physical causes of these fluctuations in the IPK’s mass. I argue that this ability to recognize that the material standard does not exactly belong to the class that it designates is actually a crucially important feature that

allows metrologists to help the IPK *better* realize its ideal status as the unit prototype. Indeed if metrologists could only theorize the mass of the IPK as exactly 1 kg during its reign, then they would not have been able to succeed in constructing as accurate and stable a system of measurement as they did.

A second, related component of the independent standard by which the IPK's mass could be judged is through the periodic verifications of prototypes. Three such periodic verifications have been carried out, in 1889, 1948, and 1989. This involved a careful comparison of the mass of the IPK and its six official copies, as well as a comparison of two of those copies with the platinum-iridium national prototypes from 34 countries. The results of these three periodic verifications revealed that the IPK, its six sister copies, and other kilogram prototypes have been steadily diverging in mass from each other, as can be seen in Figure 4:

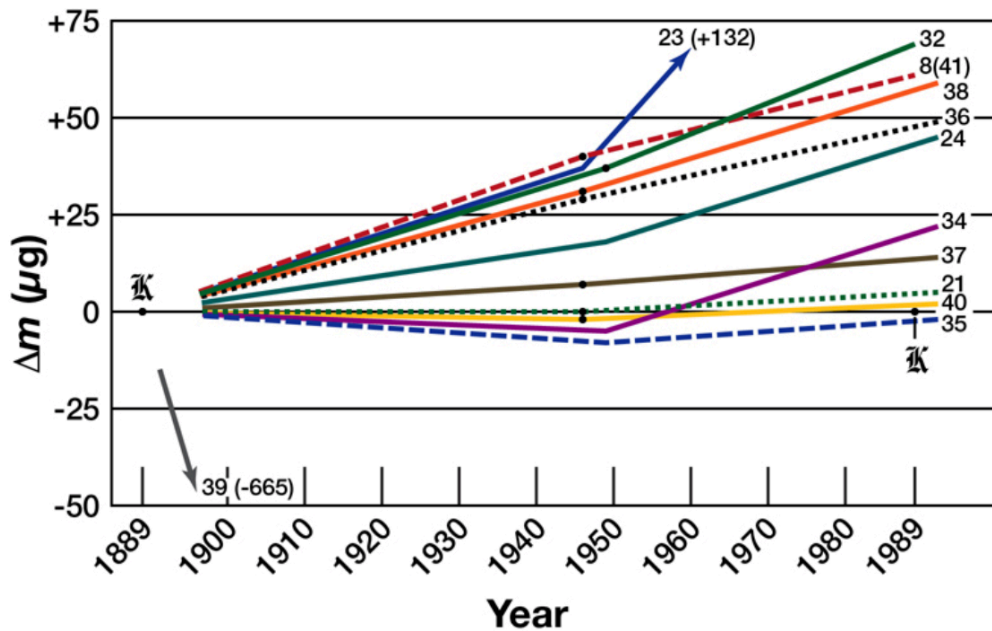


Fig. 4: Mass drift of national prototype kilograms and two IPK sister prototype kilograms relative to the IPK, which is depicted in accordance with its definition as being exactly 1 kilogram, and hence not having had any change of mass over the 100 year period of comparisons. (Source: Wikimedia commons)

Although metrologists may stipulate that the error in the mass of the IPK is exactly zero, they use the results of intercomparisons such as this to try to assess its stability empirically. Hence these empirical comparisons provide a kind of independent standard by which to judge the mass of the IPK. In his article, “The Kilogram: The Present State of Our Knowledge”, BIPM metrologist Terry Quinn summarized the results of the last periodic verification as follows:

It appears that the copies are increasing mass as a function of time with respect to the international prototype to the extent of about $0.5\mu\text{g}$ per year (5 parts in 10^{10} per year). An alternative, and from the evidence of Fig. [4] perhaps more probable, interpretation is that the mass of the international prototype is falling with respect to that of its copies. (Quinn 1991, p. 83)

Although there may be a problem of underdetermination whether it is in reality the some dozen copies all increasing and the IPK holding steady, or the IPK losing mass relative to those dozen

copies, the latter interpretation is, given our background knowledge, more probable. Here we see quite clearly that background knowledge about contamination and cleaning, and additional empirical knowledge from intercomparisons, do in fact provide an independent (albeit fallible) standard by which the mass of the IPK can be judged.

Indeed, it was in the wake of this last intercomparison that the metrology community judged that we were in a situation analogous to Witteveen's scenario *b*, and hence initiated the long and difficult process to stipulatively redefine kilogram, a process which was not complete until 28 years later in May 2019. I want to argue that it was precisely the metrologists' discovery that we were in a scenario like *b* that prompted them to make the stipulative redefinition, moving from scenario *a* (the IPK artefact definition of the kilogram) to the current scenario *c* (the new Planck-constant-tied definition of the kilogram⁸). Without the possibility of being in a scenario like *b*, it is not clear how one can make sense of the decision to undertake the difficult redefinition. The decision was neither arbitrary nor based on a whim, which if it was purely a matter of stipulation it would seem to be; rather it was a decision made reluctantly under an increasingly heavy burden of evidence. Moreover, as I hope it is becoming increasingly clear, it was crucially important that metrologists keep this independent "common sense" standard in mind. If in fact the IPK had continued to lose mass, but had stipulatively been maintained as exactly 1kg, then, for example, the doses of pharmaceutical drugs would have had to be steadily adjusted, potentially reaching harmful thresholds for the human body. The world is under no obligation to respect our stipulative pronouncements on the accuracy of scientific types, and insofar as scientists must use those scientific types to guide their interactions with the world, they cannot afford to ignore that failure of coherence.

As we have seen in this case study, there is a complicated interplay between stipulation and empirical discovery in the specification of scientific types. Although scientific types function as a stipulative standard, and hence seem to be infallible, there is an independent standard by which such scientific types can be judged, and that is in terms of a broader coherence with a background network of scientific practice, evidence, theory, and auxiliary knowledge—what in anticipation of the Duhemian analysis we might collectively call a “non-symbolic” or “common sense” standard (without the latter term in any way suggesting that it need be obvious). With this understanding of the myth of the absolute accuracy of measurement standards in hand, we are now in a position to see how it can be extended to other scientific types.

4.2: Revising Holotypes

Passing comparisons of holotypes to measurement prototypes can be found in the literature, though such comparisons have never been explored in detail. For example, in noting the ICZN's understanding of a holotype as an objective international standard of reference, Witteveen remarks, "In this sense a type specimen is not much different from other archived reference standards, like the International Prototype Kilogram" (2015, p. 583). I want to argue that they are indeed alike, but in a sense that has not been widely appreciated for either. As Witteveen, Haber, and others have noted, the ICZN's code does have several articles pertaining to the revision of type specimens. However, it is worth taking a closer look at the precise wording of these articles in order to better understand what motivates them.

⁸ Of course in the case of the kilogram the new definition is no longer in terms of an artefact scientific type.

There are four articles in the *International Code of Zoological Nomenclature* that pertain to the revision of type specimens. The first is Article 70.3 titled *Misidentified type species*:
If an author discovers that a type species was misidentified . . . the author may select, and thereby fix as type species, the species that will, in his or her judgement, *best serve stability and universality*, either 70.3.1. the nominal species previously cited as type species . . . or 70.3.2 the taxonomic species actually involved in the misidentification. (ICZN Code, Article 70.3; emphasis added)

Although this article concerns type species, rather than type specimens, it is nonetheless relevant to assessing the method of typification. The type species is the name-bearing type of the nominal genus; i.e., it is the species that first fixes the reference of, or "defines," a new genus taxonomic category. It thus functions in the same way as a type specimen serving as the objective standard of reference for the name it bears. There are two things to note about the above article: First, the notion of misidentification is conceptually possible within the rules of the system of typification and, moreover, it is treated as temporally prior to the revision of the type. Such a misidentification can, thus, be cashed out in terms of the analog of Witteveen's scenario *b* in Fig. 2. The second key feature to note in the above article is the importance placed on a type species being able to "serve stability and universality." When identification goes awry, the goal that the scientist is required to keep in mind is that of finding a specimen that enables a coherence of practice across time (i.e., stability) and across space (i.e., universality).

The second article in the ICZN's code relevant to the revision of types is Article 74 and concerns when the name-bearing type has been based on a series of specimens (syntype) rather than a single specimen. Article 74 allows the name-bearing function to be concentrated from a collection of specimens to a single individual that then becomes the type specimen (lectotype) and deprives the other specimens of their former name-bearing function. This article can be invoked, for example, when the specimens in the syntype series are found to belong to different species. As Recommendation 74G indicates, "the designation of lectotypes should be done as part of a revisionary or other taxonomic work to enhance the stability of nomenclature, and not for mere curatorial convenience" (ICZN Code, Article 74.G). As we see here, recognition of the need to revise taxonomy should precede the revision of types and when types are not working to enhance stability, they should be revised accordingly.

The third article relevant to the revision of types is Article 75, which involves the conditions for designating a new type specimen (neotype) when the previous type specimen is recognized to be inadequate. This can happen for a number of reasons as the ICZN code states (with emphases added):

1. When there is "the express purpose of clarifying the taxonomic status or the type locality⁹ of a nominal taxon" (75.3.1).
2. When a holotype is believed "to be lost or destroyed" and a physical type specimen is needed (i.e., no published descriptions or pictures are adequate) (75.3.4).
3. "A neotype may be based on a different sex or life stage if necessary or desirable to *secure stability* of nomenclature" (75.3.5).

⁹ 'Type locality' is being used here in a broader sense (than that of stratotype discussed in section 2.2) to mean the geographical or stratigraphical location where the type specimen was collected.

4. "When an author considers that the taxonomic identity of a nominal species-group taxon cannot be determined from its existing name-bearing type . . . and *stability or universality* are threatened thereby" (75.5).
5. "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" (75.6)

This last condition of the article (75.6) was the one invoked in the case involving *T.s. infernalis* cited by Haber (2012) and discussed in section 3. A recurring theme in these conditions for the revision of type specimens is that they are failing to secure a stability and coherence ("universality") of nomenclatural practice. And as we see in this last condition especially, the revision can be prompted when the type specimen gets too far out of "taxonomic accord" with the prevailing usage. As I will argue more below, this suggests that "taxonomic accord with prevailing usage" is functioning here as an independent standard by which to judge the accuracy of a type specimen in picking out the taxon to which it belongs. In other words, a type specimen might carry the name 'S' that designates taxon S, but by the criterion of "taxonomic accord with prevailing usage" the type specimen belongs to a different taxon, T, thus landing us in the scenario *b* that Witteveen claims is impossible in the ICZN's system of typification. As Article 75 makes clear, such a situation is possible according to ICZN code rules, and it is the nomenclature that should, in such cases, bend to the prevailing usage.

The final article in the code relevant to our discussion is Article 81 regarding the purpose and extent of plenary power of the International Commission on Zoological Nomenclature:

The Commission has the plenary power . . . to modify the application of provisions of the Code to a particular case, if such application would in its judgement disturb *stability or universality or cause confusion*. For the purpose of preventing such disturbance and of *promoting a stable and universally accepted nomenclature*, it may by use of its plenary power, conserve, totally, partially, or conditionally suppress, or give a specified precedence to, or make available any name, type fixation or other nomenclatural act, or any publication, and establish replacements. (81.1)

As this "catch-all" article emphasizes, there are two criteria or standards by which to judge the adequacy of a scientific type: 1. By the rules of the principle of typification, and 2. By the extent to which that scientific type promotes a stable and universal system of nomenclature, minimizing confusion. When all goes according to plan, these two criteria align: indeed, the system of typification is adopted precisely because it is typically the best method for securing a stable and universal—or "coherent"—practice. But when these two criteria do come apart, Article 81.1 makes it clear that it is the second "coherence" criterion that should ultimately take precedence. In other words, coherence with prevailing community practice is functioning as an independent "common sense" standard by which the adequacy of the scientific type can be judged.

4.3 Revising Stratotypes and the Definition-Correlation Debate

As we saw in section 2.2, a stratotype (or more precisely a global boundary stratotype section and point (GSSP)), is a specific place marked in a sequence of rock strata constituting the standard for the definition and recognition of the boundary marking the beginning of a named chronostratigraphic (and geochronologic) unit. Like holotypes and prototypes, stratotypes are intended to have a permanence, but nonetheless there are provisos that allow for their revision.

As I will show, also like holotypes and stratotypes, those provisos can best be understood in terms of the ultimate purpose of stratotypes as a tool for securing a stability and coherence of scientific nomenclature and practice across a diverse community of scientists.

There are a number of interesting conceptual issues and ongoing methodological debates surrounding the practice of stratotypes. One that is particularly relevant to our present concern with the myth of the absolute accuracy of scientific types is what might be called the definition-correlation debate. Stratigraphers have long debated whether correlation should precede definition, or whether definition should precede correlation. As discussed before, it is only by time-correlating distant stratigraphic sections, and integrating them into a single chronostratigraphic scale (and hence Geologic Time Scale), that key events in Earth's history, such as global climate changes and mass extinctions, can be identified, and cause and effect can be investigated. The question concerns the best way to secure such a consistently defined and globally correlated chronostratigraphic scale.

One philosophical approach goes back to Hollis Hedberg (1903-1988), who was a key architect in developing the international system of nomenclature and practice for stratigraphy, and the editor of the first edition of the *International Stratigraphic Guide* published in 1976. In this *Guide*, Hedberg argues that definition must precede correlation: "Only after the type limits [boundary stratotypes] of a chronostratigraphic unit have been established [defined] can the limits be extended geographically beyond the type section [globally correlated]" (Hedberg 1976, p. 86). Hence, according to Hedberg and subsequent thinkers such as Marie-Pierre Aubry (Aubry et al. 2000), a stratotype boundary must be conventionally defined first before the substantive scientific project of correlation, by means of biostratigraphic (e.g., first appearance of an index fossil) or other events (e.g., stable isotope excursions) can be undertaken; and by being prior, those definitions should be immune to changing scientific views about correlation.

The alternative philosophical view in this debate is that correlation must precede definition. This is the view legislated by the revised guidelines of the International Commission Stratigraphy (ICS) adopted in 1996 and currently in effect. As then chairman of the ICS, Jürgen Remane, and colleagues write:

To define a boundary first and then evaluate its potential for long-range correlation (as has been proposed in some cases) will mostly lead to boundary definitions of limited practical value. (Remane et al. 1996, p. 78)

On this view, boundary stratotypes should be placed not arbitrarily, but in those localities with a maximum (or at least sufficient) potential for global correlation.

One of the GSSPs that was placed prior to the revised guidelines (on the former "definition precedes correlation" approach) was the stratotype that defined the base of the Silurian System. This GSSP was plagued with problems of correlatability, specifically "biostratigraphic deficiencies and even [a] lack [of] the key index fossils" (Rong et al. 2008, p. 315). A formal restudy by the Subcommittee on Silurian Stratigraphy subsequently led to both the first revision of a GSSP and to the formulation of the general guidelines for how to revise a stratotype (Rong et al. 2008). In a forerunner to this revision paper, they respond to the objection (by Holland et al. 2003) that, once defined, a GSSP should never be revised. They write,

The central important point raised by Holland et al. (2003) is the issue of stability in Silurian stratigraphy. We agree that stability is of paramount importance now that the Silurian GSSPs have been defined for almost two decades and it is in the interests of both stability and precision in international correlation that these restudies are being

undertaken. It is our view that reliance on a poorly defined GSSP does not lead to stability. The examples of the two GSSPs selected for restudy illustrate this point well. (Melchin et al. 2004, p. 124)

As these quotations make clear, both sides recognize that the ultimate purpose of a stratotype is to be a tool for securing a stability and coherence of practice in the face of ever-evolving scientific knowledge—a focal function that they share with holotypes and prototypes. And, as we also saw in the case of holotypes, stratotypes are at the intersection of nomenclature and (stratigraphic) taxonomy, having one foot in each; hence they are not just about having a fixed system of definitions for their own sake, but rather are for the broader purpose of securing a coherent (universal and stable) system of scientific practice. Thus the failure to secure such a coherent system of practice becomes an independent "common sense" standard by which to judge the inadequacy of the stratotype.

5. A Duhemian Approach to Scientific Types

My aim in this paper has been to call attention to an important component of taxonomic and nomenclatural scientific practice that has been largely overlooked by philosophers of science, namely typification and corresponding notion of a scientific type. Through a close examination of the use of holotypes in biology, stratotypes in geology, and measurement prototypes in metrology, I have abstracted the following general definition: *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*. I then showed the fertility of this new conception by using this general framework to address three debates: the necessity-contingency debate about holotypes, the absolute accuracy debate over measurement standards, and the definition-correlation debate about stratotypes. In particular, we can recognize that although scientific types are defined conventionally through an act of stipulation, the fundamental purpose of scientific types is to secure a stability and coherence of scientific practice as new discoveries are made. As we saw, this purpose can function as an independent "common sense" standard by which scientific types can be evaluated and various debates concerning them can be resolved.

With this detailed examination in hand, we are now in a position to further develop our general philosophical account of scientific types. The account I want to defend draws its inspiration from Pierre Duhem's philosophy of science, and more specifically the framework he uses to critique conventionalism. Although Duhem never (to my knowledge) discusses the practice of typification, his approach contains a number of insights that together can provide the foundation for a philosophy of scientific types. These are found in his book *The Aim and Structure of Physical Theory* (hereafter *The Aim*), which despite its named focus on physical theory is actually motivated by reflections on biological classification.¹⁰ In particular, there are

¹⁰ *The Aim* both opens and closes with a discussion of biological classification, where Duhem describes how zoologists aim to reconstruct the true tree of life, but only have access to these "relations of real family affiliation" indirectly through comparative anatomy. Thus he sees biological taxonomy proceeding iteratively, aiming towards a natural classification where the abstract (or what he elsewhere calls "symbolic") entities and relations of the taxonomy reflect the true familial relations of real organisms. Although interesting and underexplored, the details of Duhem's philosophy of biology are not directly relevant to our project here.

three elements I want to draw on: first, his understanding of theories as symbolic systems; second, his critique of conventionalism; and third, his notion of common sense. Since these three elements of Duhem's philosophy are often not widely appreciated, it is worth briefly reviewing them in turn.

The first key element is Duhem's understanding of scientific theories as symbolic systems. No matter how close a taxonomy comes to what Duhem calls a "natural classification" (Duhem [1914] 1954, p. 297), it will always be an abstraction or symbolic representation of a concrete world. Traditionally this notion of symbolic system has been interpreted too narrowly to mean a theory mathematized by a system of equations, but on my view, Duhem intends the point to apply much more broadly to any theoretical system, including, for example, biological theories. Duhem highlights this duality between the symbolic system and the concrete world at many different levels of scientific theorizing. So, for example, he distinguishes between the concrete experimental apparatus and "a schematic model of the same instrument constructed with the aid of symbols supplied by theories", noting that "it is on this ideal and symbolic instrument that [the scientist] does his reasoning" (p. 156). He moreover notes that when the latter is too simplistic a representation of the former, it can be corrected. Similarly Duhem distinguishes between "concrete circumstances" or "practical facts" on the one hand, and "theoretical facts" on the other, noting how the former are turned into the latter through "the intermediary of measurements" (pp. 133-4). Thus a central component of Duhem's philosophy of science is this distinction between a symbolic system and the world it is supposed to represent.

The second, and for our purposes most important, element from Duhem's philosophy of science is his critique of conventionalism (the conventionalism found, for example, in the work of Henri Poincaré and Edouard Le Roy). The conventionalist thesis that Duhem wants to reject is 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" (p. 209). He considers the example, cited by the conventionalists, that "free fall" of a body means its "acceleration is constant." The conventionalist argues that such a law can never be contradicted by experiment because "it constitutes the very definition of what is meant by 'falling freely,'" and that if a fall was not uniformly accelerated, then we would be forced to conclude that it was not free (p. 209). Duhem's response to this claim of the conventionalists is to distinguish between the "symbolic" meaning of the phrase as defined within a system of mechanics and the "common sense" meaning:

The words 'free fall of a heavy body' now have two distinct meanings . . . they have their *real* meaning, and they mean what common sense means in pronouncing them; for the physicist they have a *symbolic* meaning, and mean 'uniformly accelerated motion.' (p. 209-10)

Duhem then asks us to suppose that we observe that a certain fall regarded by common sense as a free fall has a slightly variable acceleration. The proposition which in our theory gives its symbolic meaning to the words 'free fall' does not represent with sufficient accuracy the properties of the real and concrete fall we have observed. (p. 210)

In this case Duhem argues, contra the conventionalists, that there are in fact two options available to the scientist: First,

we may declare that we were wrong in establishing a connection between the concrete fall we have observed and the symbolic free fall defined by our theory. . . . We may seek to eliminate by means of suitable 'corrections' the 'causes of error,' such as air resistance, which influenced our experiment. (pp. 210-211)

This strategy is to hold the symbolic system fixed and attribute the error to our measurements or the world. According to the conventionalists, this is the *only* option, since 'uniformly accelerated motion' is the very *definition* of 'free fall'. Although this will be the most commonly used strategy, Duhem argues the conventionalists are wrong in thinking it is the only strategy. A second option, according to Duhem, is that

we can declare that we were right in regarding the fall studied as a free fall and in requiring that the theoretical definition of these words agree with our observations. . . . Since our theoretical definition does not satisfy this requirement, it must be rejected; we must construct another mechanics on new hypotheses, 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, p. 210)

Although Duhem grants that the first strategy is typically to be preferred, it is up to the wisdom ("good sense" or *bon sens*) of the scientist to know when the second option is called for.

The point I especially want to emphasize here is Duhem's recognition of *two* definitions of a term: a "symbolic one" (taken to be true by definition) and a "common sense" one. For Duhem, the success of science rests on keeping both these two separate definitions in play and comparing one against the other. While the hope is that they will normally coincide, when they do come apart there is a substantive scientific decision regarding how to proceed. Although deference to our symbolic systems is usually called for, sometimes the wise choice is to have our symbolic systems bend to common sense.

This notion of "common sense" is the third element I want to draw on from Duhem's philosophy. This too is often misunderstood, so a few brief comments are in order. The first point is that in calling it 'common sense' he does not in any way mean to imply that it is 'obvious'; rather, he means that it is outside the framework of a particular symbolic system. Second, Duhem actually has a more complex classification of different notions of "common sense," which are borrowed in part from Blaise Pascal and come out more explicitly in his work *La Science Allemande* (see Martin 1991 for a discussion). It includes not just *bon sens* (good sense) but also *connaissance commune* (communal knowledge or common knowledge) and *finesse d'esprit* (astuteness). For Duhem, success in science requires a balance between what he calls the "geometrical" mind working within a symbolic system and the mind of "common sense" understood in these three ways.

With these three elements in place, we are now in a position to bring them to bear on further developing a philosophical account of scientific types. I want to argue that scientific systems of typification, such as the examples from biology, geology, and metrology discussed here, are symbolic systems in Duhem's sense. Moreover, they are symbolic systems that are at the intersection of (a conventionally stipulated) nomenclature and (efforts towards a "natural classification") taxonomy. Scientific types, as we have seen, are designed to fix the reference of the names they bear: '*T.s. infernalis*' refers to whatever species the *T.s. infernalis* type specimen belongs to, and 'one kilogram' refers to whatever mass the IPK possesses. As such, these objects

serve as the "symbolic" standard of reference for their kind, and hence are taken *by convention* to be infallible in belonging to the kind they name or define.

However, as we have seen in the three examples of scientific types discussed here, there are independent standards for judging which taxons these scientific types belong to. Following Duhem we can refer to these broadly as the "common sense" or "communal knowledge" standards. In the case of prototypes, the independent (non-symbolic) standard by which to judge whether the mass of the IPK was exactly 1.000 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 the BIPM and around the world. Although perhaps underdetermined as to where the error lies, a *finesse d'esprit* guided the metrologists in the decision to locate the error in the scientific type and initiate the process of redefinition. In the case of holotypes, we saw the independent (non-symbolic) standard in the ICZN's deference to "taxonomic accord with prevailing usage"—a kind of "communal knowledge" independent standard—which can judge a type specimen not to belong the species it names. Similarly, a stratotype definition of the base of a chronostratigraphic unit, such as the Silurian, can be judged by the independent standard of enabling a stability and coherence of practice in stratigraphic correlation, which is its aim. In all three cases, 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. Scientific types are also judged by independent, common sense or communal knowledge standards.

In thinking about the scientific practice of typification, it is helpful to keep in mind not just the structure of this symbolic system, but also its aim. The aim of scientific typification, recall, is to secure a stability and coherence of nomenclature and practice in the face of ever-evolving scientific knowledge. Scientists usually cede a priority of accuracy in nomenclature and definitions to the international organizations governing these systems of typification, and choose to locate any errors, not in these type standards, but in their particular taxonomic or measurement practices. They do so because this is typically the best way to secure a stability and coherence of scientific nomenclature and practice. However, 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; indeed all systems of typification have clauses allowing for such an option in exceptional circumstances, as we have seen. In the end it is nomenclature that serves taxonomy, not the reverse.

Although types can perform a variety of discipline-specific functions, they share a common focal function, which I have here used as a basis for constructing the first general definition of a scientific type. In addition to providing this definition, I have outlined a novel Duhemian philosophy of scientific types and demonstrated the fertility of this framework for a number of debates: the necessity-contingency debate in biology, the absolute accuracy of standards debate in metrology, and the definition-correlation debate in geology. There are other important philosophical debates about types to explore within each of these subfields, such as the ethical debate in biology about whether type specimens should be collected for vulnerable species (Havstad 2019). By offering a general definition and theory of types, I hope to provide a foundation for further philosophical work in this area. Just as there has been a productive synergy between philosophical accounts of natural kinds and scientific taxonomic practices, so

too there is much to be gained from developing a deeper understanding of the practices and philosophy of scientific types.

Acknowledgements

This paper was first presented at the *XX Jornadas Rolando Chuaqui Kettlun: Filosofía y Ciencias* conference in Santiago, Chile, in August 2019, and I am grateful to my generous hosts José Tomás Alvarado and Pablo Acuña, and the audience members there, for giving me the opportunity to develop and refine these ideas in such a stimulating and congenial environment. I am also grateful to Doug Erwin and the graduate students in my Philosophy of Geosciences Research Group, especially Aja Watkins, Leticia Castillo Brache, and Federica Bocchi, for their feedback on an earlier draft.

References

- Aubry, M.-P., J. Van Couvering, W. Berggren, and F. Steininger (2000), "Should the Golden Spike Glitter?" *Episodes* 23 (3): 203-210.
- BIPM (2006) *The International System of Units (SI)*. Bureau International des Poids et Mesures, 8th edition. https://www.bipm.org/utis/common/pdf/si_brochure_8.pdf
- BIPM (2019a) *The International System of Units (SI)*. Bureau International des Poids et Mesures, 9th edition. ISBN 978-92-822-2272-0. <https://www.bipm.org/utis/common/pdf/si-brochure/SI-Brochure-9-EN.pdf>
- BIPM (2019b) *A Concise Summary of the International System of Units (SI)*. Bureau International des Poids et Mesures. <https://www.bipm.org/utis/common/pdf/si-brochure/SI-Brochure-9-concise-EN.pdf>
- Bokulich, A. (2014), "Pluto and the 'Planet Problem': Folk Concepts and Natural Kinds in Astronomy" *Perspectives on Science* 22 (4): 464-490.
- Bokulich, A. (2020), "Calibration, Coherence, and Consilience in Radiometric Measures of Geologic Time" *Philosophy of Science*.
- Cowie, J. W., W. Ziegler, A.J. Boucot, M.G. Bassett, and J. Remane (1986) *Guidelines and Statutes of the International Commission on Stratigraphy (ICS)*. Volume 83. Frankfurt a. M.: Courier Forschungsinstitut, pp. 1-14.
- Duhem, P. ([1914] 1954) *The Aim and Structure of Physical Theory*, 2nd edition. Translated from French by Marcel Rivière Princeton: Princeton University Press.
- Girard, G. (1990) "The Washing and Cleaning of Kilogram Prototypes at the BIPM" Sèvres: Bureau International des Poids et Mesures. <https://www.bipm.org/utis/common/pdf/monographies-misc/Monographie1990-1-EN.pdf>
- Girard, G. (1994) "The Third Periodic Verification of National Prototypes of the Kilogram (1988-1992)" *Metrologia* 31: 317-336.
- Haber, M. (2012) "How to Misidentify a Type Specimen" *Biology & Philosophy* 27: 767-784.
- Havstad, J. (2019) "Let Me Tell You 'Bout the Birds and the Bee-Mimicking Flies and Bambiraptor" *Biology & Philosophy* 34: 25 <https://doi.org/10.1007/s10539-019-9681-3>.
- Hedberg, H.D. (1976) *International Stratigraphic Guide*, First Edition. John Wiley & Sons, Inc.
- Holland, C., M. Bassett, and B. Richards (2003) "Stability in Stratigraphy" *Lethaia* 36: 69-70.
- Hull, D. (1982) "Exemplars and Scientific Change," in P.D. Asquith and T. Nickles (eds), *PSA 1982, Vol. II*, Philosophy of Science Association, East Lansing, pp. 479–503.
- ICZN (accessed 2019) *International Code of Zoological Nomenclature*. Online: <https://www.iczn.org/the-code/the-international-code-of-zoological-nomenclature/the-code-online/>
- Kendig, C. (2016), *Natural Kinds and Classification in Scientific Practice*. New York: Routledge.
- LaPorte, J. (2003) "Does a Type Specimen Necessarily or Contingently Belong to its Species?" *Biology & Philosophy* 18 (4): 583-588.
- LaPorte, J. (2004) *Natural Kinds and Conceptual Change*. Cambridge: Cambridge University Press.

- LaPorte, J. (2018) "Rigid Designators", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), URL = [<https://plato.stanford.edu/archives/spr2018/entries/rigid-designators/>](https://plato.stanford.edu/archives/spr2018/entries/rigid-designators/).
- Levine, A. ((2001), "Individualism, Type Specimens, and the Scrutability of Species Membership" *Biology & Philosophy* 16: 325-338.
- Martin, R. N. (1991) *Pierre Duhem: Philosophy and History in the Work of a Believing Physicist*. La Salle: Open Court Publishing Co.
- McGowran, B (2005) *Biostratigraphy: Microfossils and Geologic Time* Cambridge: Cambridge University Press.
- Melchin, M., J. Rong, F. Gradstein, T. Koren, and S. Finney (2004) "Stability in Stratigraphy" *Lethaia* 37: 124-125.
- Pyle, R. and E. Michel (2008) ZooBank: Developing a Nomenclatural Tool for Unifying 250 Years of Biological Information. *Zootaxa* 1950: 39-50.
- Quinn, T. (1991) "The Kilogram: The Present State of Our Knowledge" *IEEE Transactions on Instrumentation and Measurement* 40 (2): 81-85.
- Remane, J. M. Bassett, J. Cowie, K. Gohrbandt, H.R. Lane, O. Michelsen, W. Naiwen (1996) "Revised Guidelines for the Establishment of Global Chronostratigraphic Standards by the International Commission on Stratigraphy (ICS)" *Episodes* 19(3): 77-81.
- Rong, J., M. Melchin, S. Williams, T. Koren, J. Verniers (2008) "Report of the Restudy of the Defined Global Stratotype of the Base of the Silurian System" *Episodes* 31 (3): 315-318.
- Smith, A., T. Barry, P. Brown, J. Cope, A. Gale, P. Gibbard, J. Gregory, M. Hounslow, D. Kemp, R. Knox, J. Marshall, M. Oates, P. Rawson, J. Powell, and C. Waters (2015) "GSSPs, Global Stratigraphy and Correlation" in D. Smith, G. Bailey, J. Burgess, and A. Fraser (eds.) *Strata and Time: Probing the Gaps in Our Understanding*. Geological Society, London, Special Publications, 404: 37 - 67. <http://dx.doi.org/10.1144/SP404.8>
- Tal, E. (2011), "How Accurate Is the Standard Second?" *Philosophy of Science* 78: 1082-1096.
- Walsh, S., F. Gradstein, J. Ogg (2004), "History, Philosophy, and Application of the Global Stratotype Section and Point (GSSP)" *Lethaia* 37: 201-2018.
- Witteveen, J. (2015) "Naming and Contingency: The Type Method of Biological Taxonomy" *Biology & Philosophy* 30: 569-586.