Written for *Kuhn's The Structure of Scientific Revolutions at 60*,
ed. by K. Brad Wray. Cambridge University Press

Kuhn's '5th Law of Thermodynamics': Measurement, Data, and Anomalies

Alisa Bokulich & Federica Bocchi
Department of Philosophy
Boston University
abokulic@bu.edu   fbocchi@bu.edu

1. Introduction

A central component of Thomas Kuhn's philosophy of measurement is what he calls the *fifth law of thermodynamics*. According to this "law," there will always be discrepancies between experimental results and scientists' prior expectations, whether those expectations arise from theory or from other experimental data. These discrepancies often take the form of what Kuhn calls *quantitative anomalies*, and they play a central role in both normal and revolutionary science. Whether the effort to resolve these anomalies is taken to be a part of normal or revolutionary science depends in part on the ever-evolving and context-dependent standards of what Kuhn calls *reasonable agreement*. In *The Structure of Scientific Revolutions*, Kuhn identifies as one of the most important types of experiments those aimed at determining the values of the fundamental physical constants. Why would he emphasize this seemingly obscure class of experiments? The answer, we argue, requires paying closer attention to, first, the historical context of a prominent research program in the Physics Department at Berkeley when Kuhn arrived in the 1950s, and second, Kuhn's broader philosophy of measurement and data. As we show, the *fifth law of thermodynamics* and the failure of *reasonable agreement* played a fundamental role in both.

In Section 2, we reconstruct Kuhn's philosophy of measurement and philosophy of data, as laid out primarily in his 1961 paper "The Function of Measurement in Modern Physical Science," where he introduces this fifth law. We discuss the important role of quantitative anomalies in Kuhn's philosophy, noting his emphasis on the iterative process of improving reasonable agreement. Section 3 turns to the historical context at Berkeley and the research program initiated by the long-time physics chair, Raymond T. Birge, who first called attention to the widespread discrepancies and inconsistencies in the experimental data on fundamental constants. We illustrate the quantitative anomalies uncovered in this research on constants, using the example of the speed of light (c), for which there were many different (and inconsistent!) experimentally determined values measured during the Birge-Kuhn era.

We follow this important research program forward in time in Section 4, highlighting Kuhnian elements taken up by the metrology institution subsequently charged with periodically adjusting the values of the fundamental constants, known as Committee on Data for Science and Technology (CODATA). In particular, we identify three striking points of similarity: First, like Kuhn, these metrologists emphasize the iterative and ever-changing standards of reasonable agreement, prioritizing the identification of quantitative anomalies. Second, the metrology community also expresses a fundamental skepticism about scientists' ability to ever know the "true value" of a fundamental constant. Third, in the absence of any access to the true values of the constants, these metrologists emphasize the values of consistency and coherence as the only arbiters in deciding what numerical value to adopt. We connect these points to the ongoing effort to determine the value of the gravitational constant (*G*), which is the fundamental constant that Kuhn emphasizes as being particularly problematic in the *Structure*.

In Section 5, we discuss Kuhn's later reflections on the formative role that his earlier work on the philosophy of measurement and data had for the development of his views in the *Structure*. By paying closer attention to Kuhn's work on the philosophy of measurement we are also able to recover a key notion of scientific progress in Kuhn's thinking that goes beyond the increase in puzzle-solving ability later identified in the Postscript to the *Structure*. We conclude by reflecting on the continuing relevance of Kuhn's views for the philosophy of metrology and philosophy of data today.

2. Kuhn's Philosophy of Measurement & Data

In *The Structure of Scientific Revolutions*, Kuhn highlights as one of the "most important of all" classes of experiments in normal science the "determination of physical constants" ([1962] 1996, p. 27). He notes, for example, the "improved values of the gravitational constant [G] have been the object of repeated efforts ever since [the 1790s] by a number of outstanding experimentalists" (pp. 27-28). Kuhn goes on to mention the ongoing work to improve the values of several other fundamental constants, such as "the astronomical unit [AU], Avogadro's number [N_A], Joule's coefficient [μ_{JT}], the electronic charge [e], and so on" (p. 28). This emphasis on the iterative determination of the values of fundamental constants might prima facie be surprising given the paucity of attention this topic has received from philosophers of science. However, when Kuhn arrived at Berkeley in the mid-1950s he was surrounded by discussions of the adjustment of fundamental physical constants, which was one of the towering achievements of his colleague Raymond T. Birge, who had served as chair of the Berkeley physics department for over 20 years.[1] While the "true values" of the fundamental physical constants are assumed to be constant, their values can only be known empirically through experimentation and measurement. And as Kuhn and his Berkeley colleagues were acutely aware, the empirically determined values of these constants over time are anything but constant.

In order to better understand Kuhn's views on this work, and its possible influences, we must look to Kuhn's writings on the philosophy of measurement and philosophy of data, which are worked out in detail in a paper he published just one year before the *Structure* titled "The Function of Measurement in Modern Physical Science."[2] It is in this paper that Kuhn introduces the eponym for the title of our paper: *the fifth law of thermodynamics*. In physics, of course, there are only the three traditional laws of thermodynamics, related to the conservation of energy, the increase of entropy, and the limit of absolute zero.[3] Both the fourth and fifth laws of thermodynamics are humorous aphorisms of the experimental sciences. The more familiar fourth law (which Kuhn relegates to a passing footnote) states that "no piece of experimental apparatus works the first time it is set up." However, Kuhn's primary interest is in articulating what he calls the *fifth law of thermodynamics*, which states that "no experiment gives quite the expected numerical result" (Kuhn [1961] 1977, p. 184).

---

[1] We will return to discuss Birge's program of the periodic adjustment of physical constants and the legacy of its influence on the philosophy of metrology in the next section.

[2] This paper, originally published in 1961 in the journal *Isis*, is reprinted in Kuhn's (1977) collection of essays *The Essential Tension*, which is the source for our page numbers cited here.

[3] To complicate the numbering, there is also a "zeroth law of thermodynamics", which establishes a kind of transitivity relation for systems in thermal equilibrium.

These prior "expected numerical results" can come from two kinds of sources: On the one hand, they can be based on data from other experiments that have measured the same quantity, either by the same experimental method (an attempted replication), or via a different experimental approach. On the other hand, expected numerical results can also come from predictions that were calculated on the basis of theory. For example, in a collision experiment involving rigid bodies in a low friction environment, one would theoretically expect momentum to be conserved. As anyone who has been in a high school physics lab knows, however, it can be extremely difficult to get the experimental data to come out exactly the way the theory predicts. Indeed, Kuhn rightly argues that the numbers will never agree exactly, even in the best-case scenario, due to the resolution limits of the measuring instruments employed, which will always yield fewer decimal places than one could in principle calculate from theory. More importantly, Kuhn argues that the theoretical expectations will involve various idealizations, and the experimental setup various "approximations," such that the concrete experiment will never be a perfect realization of the theoretical schema of the experiment. There will be any number of interfering influences, some perhaps known or guessed at, but others unknown. For all these reasons, experimenters are constrained by the fifth law of thermodynamics and can at best hope for what Kuhn calls "reasonable agreement." This notion of reasonable agreement is a crucial one for Kuhn, which we will discuss in more detail shortly, but before doing so it may be helpful to provide some background and terminology in the philosophy of measurement and data to help understand Kuhn's ideas.

In trying to better understand the discrepancy that Kuhn identifies in his fifth law of thermodynamics, it is helpful to turn to Pierre Duhem's chapter on "Experiment in Physics" from his book *The Aim and Structure of Physical Theory*.[4] In this chapter, Duhem introduces the following distinction: "When a physicist does an experiment, two very distinct representations of the instrument . . . fill his mind: one is the image of the concrete instrument that he manipulates in reality; the other is a schematic model of the same instrument, constructed with the aid of symbols supplied by theories" (Duhem ([1914] 1954), pp. 155-156). Duhem goes on to explain that "[t]he schematic instrument is not and cannot be the exact equivalent of the real instrument" (p. 157). He continues, "the physicist, after reasoning on a schematic instrument that is too simple . . . will seek to substitute for it a more complicated scheme that resembles reality more. This passage from a certain schematic instrument to another which better symbolizes the concrete instrument is essentially the operation that the word *correction* designates in physics" ([1914] 1954), p. 157, emphasis added). Kuhn describes this very process using the example of Newton's *theoretical* description of a pendulum and a *concrete* pendulum in the lab:

> "The suspensions of laboratory pendula are neither weightless nor perfectly elastic; air resistance damps the motion of the bob; besides, the bob itself is of finite size. . . . If these three aspects of the experimental situation are neglected, only the roughest sort of quantitative agreement between theory and observation can be expected. But determining how to reduce them . . . and what allowance to make for the residue are themselves problems of the utmost difficulty. Since Newton's day much brilliant research has been devoted to their challenge." (Kuhn [1961] 1977, p. 191)

As Kuhn emphasizes, this gap between theory and data is not one that is initially directed at the theory, or yet attempting any sort of confirmation or falsification. Instead, the theory is initially

[4] Duhem's book was originally published in 1906, and his work influenced many of the authors that Kuhn cites. Here we quote from the second edition of Duhem's book, published in 1914.

held “fixed,” while scientists attempt to bridge the gap through a complexification of our understanding of the experimental apparatus and an extended process of wrestling with various ‘corrections’ to the data.

The process that Kuhn describes here is similar to what metrologists today would refer to as creating a model of the measurement process and an associated uncertainty budget. The international organization responsible for coordinating the vocabulary and standards of measurements in metrology (i.e., the science of measurement) is the International Bureau of Weights and Measures, known by the acronym of its French title BIPM, which was established by the Metre Convention in 1875. In discussing measurements in physics, it is helpful to briefly review some of the basic terminology in metrology to bring philosophical clarity to the discussion. The quantity one is trying to measure in the world is known as the *measurand*, which is assumed to have some definite, but unknown (and arguably unknowable) *true value*. When a measurement is performed using some apparatus, the scientist will obtain what is called a measurement *indication*—a reading on the dial, for example. A measurement result requires turning this measurement indication into a measurement *outcome* which is the scientist’s considered *estimate* of the measurand’s “true value”. As Eran Tal explains,

> “‘indication’ . . . does not presuppose reliability or success . . . but only an intention to use such outputs for reliable indication of some property of the sample being measured. . . . A measurement outcome, by contrast, is an estimate of the quantity value associated with the object being measured . . . inferred from one or more indications. . . . and include[s] either implicitly or explicitly, an estimate of uncertainty” (2012, pp. 143-144).

Turning measurement indications into measurement outcomes (or “results”) can involve all sorts of theoretical calculations, such as equations that allow one to convert the measured quantity (e.g., travel time of a light signal measured in seconds) into the quantity of interest (e.g., distance measured in kilometers), a process that is known as *data conversion* (see, e.g., Bokulich 2020a). More broadly, it depends on having a detailed understanding, or model, of the measurement process, as well as an estimate of the various sources of error that may be affecting the measurement indications given by the measurement apparatus. These sources of error, or uncertainty, can be detailed in an uncertainty budget, which tries to estimate how these various sources of error likely influenced the measurement. Kuhn is calling attention to these many possible sources of error in his discussion of examples like the “vacuum is not perfect” and “the ‘linearity’ of vacuum tube characteristics” (Kuhn [1961] 1977, p. 184)). The data that is obtained from a measurement indication, thus often needs to be corrected or processed (i.e., undergo *data correction*) in order to turn it into a measurement outcome.

This metrological distinction between indications and outcomes helps us understand Kuhn’s philosophy of data: as he notes in the *Structure*, “the measurements that a scientist undertakes in the laboratory are not ‘the given’ of experience but rather ‘the collected with difficulty’ ” ([1962] 1996, p. 126). This key insight is also elaborated in his earlier article on the function of measurement:

> “the scientist often seems rather to be struggling with the facts, trying to force them into conformity with a theory he does not doubt. Quantitative facts cease to seem simply ‘the given.’ *They must be fought for and with*.” (Kuhn [1961] 1977, p. 193; emphasis added)

These uses of the term “given” are of course references to the Latin roots of the word datum, or data (from the past-tense of the verb, *dare*, to give). Not only must data be *fought for*, with the design and execution of careful experiments, but as Kuhn says they must also be *fought with,* through various forms of data correction and data processing. How does the scientist judge when the data corrections have been adequately completed? Kuhn explains, "the tests for reliability of existing instruments and manipulative techniques must inevitably be their ability to give results that compare favorably with existing theory" (Kuhn [1961] 1977, p. 194). In other words, when the data are in "reasonable agreement" with theory.

What counts as a "reasonable agreement," and when is a discrepancy a problematic anomaly? As Kuhn explains, there is no single, universal criterion; it is instead highly context dependent. "'Reasonable agreement' varies from one part of science to another, and within any part of science it varies with time" (Kuhn [1961] 1977, p. 185). First, reasonable agreement is usually field-dependent: the standard of accuracy for an experimental result to be in agreement with theory varies widely across research fields. For example, Kuhn notices that in "spectroscopy 'reasonable agreement' means agreement in the first six or eight left-hand digits in the numbers of a table of wave lengths" (Kuhn [1961] 1977, p. 185). In contrast, "there are parts of astronomy in which any search for even so limited an agreement must seem utopian. In the theoretical study of stellar magnitudes agreement to a multiplicative factor of ten is often taken to be 'reasonable'" (p. 185). Evolving norms for reasonable agreement—as established by a given subfield, for a particular quantity, and at a moment in time—set the standards and expectations that practitioners in that field must abide by in their research.

As scientists come to understand the measurement process and the various sources of error better over time—resulting in improved experimental design and protocols, a more detailed model of the measurement process, and improved uncertainty budgeting and data correction—the gap between theory and measurement data should gradually close. George Smith (2014) describes this as a process of "closing the loop," noting its critical role in the Newtonian research program, and spectacularly illustrated by the discovery of Neptune from the no-longer-reasonable agreement between Newton's theory and the observed orbit of Uranus.[5]

Similarly, Kuhn notes that "in overwhelming proportion, these discrepancies disappear upon closer scrutiny. They may prove to be instrumental effects, or they may result from previously unnoticed approximations in the theory, or they may, simply and mysteriously, cease to occur when the experiment is repeated under slightly different conditions (Kuhn [1961] 1977, p. 202). Even if the mismatch between the measurements and theoretical prediction tends to shrink as measurements get more and more refined, the gap is never completely eliminated. In the most troubling manifestation of the *fifth law of thermodynamics*, however, an anomaly can prove recalcitrant:

> "a quantitative anomaly [can resist] all the usual efforts at reconciliation. Once the relevant measurements have been stabilized and the theoretical approximations fully investigated, a quantitative discrepancy proves persistently obtrusive" (Kuhn [1961] 1977, p. 209).

It is in the course of this "mop-up" work (as Kuhn, in an overly disparagingly way, describes it) of stabilizing the measurements and theoretical approximations that the most productive

---

[5] The Neptune case illustrates the *theory-data discrepancy*, and is discussed many places, including in Smith (2014). For a fuller discussion of a case involving *data-data discrepancy*, or more specifically how the iterative comparison between two different experimental approaches to measuring the same quantity can help identify and resolve anomalies in each experimental method, gradually closing the loop, see Bokulich's (2020b) "Calibration, Coherence, and Consilience in Radiometric Measures of Geologic Time."

quantitative anomalies are revealed. At this point, the disagreement between theory and data is no longer reasonable, and the quantitative anomaly could precipitate a crisis.

3. Kuhn, Birge, & the Adjustment of Physical Constants at Berkeley

With this background on Kuhn's philosophy of measurement, data, and quantitative anomalies in place, we have one of two key components needed to understand Kuhn's puzzling remarks in the *Structure* about one of the most important classes of experiments in normal science being the iterative improvement of the values of the fundamental constants. However, the complete story also requires paying attention to the historical context in which Kuhn found himself at U.C. Berkeley when he was working out these ideas and writing the *Structure*.

When Kuhn arrived at Berkeley in 1956, one of the most prominent figures in the Physics Department was Raymond T. Birge—a figure who has largely been forgotten today. Birge served as the Chair of the Physics Department from 1932 to 1955, and is credited with being the architect behind the department's rise to international prominence, helping to recruit the likes of E. O. Lawrence and J. R. Oppenheimer (Helmholtz 1980). Birge's impact on Berkeley was so great, that when the new physics building was completed in 1964, it was named Birge Hall. More importantly for our purposes here, Birge was also the architect behind the project to coordinate a set of best current values for all the major fundamental physical constants to be prescribed for the entire physics community—a Herculean feat that made him few friends. When Birge started this project in the late 1920s, the physics community was in disarray, using many different—and often inconsistent—values for the fundamental constants. Birge reports a "surprising lack of consistency, both in regard to the actually adopted values and to the origin of such values" (Birge 1929, p. 2), not least because of 'some peculiar national flavor' (1957, p.40) in the choice of values. This lack of standardization undermined any attempt to compare measurement results from different researchers across different labs and in different countries.

Birge's task was to comb through the entire physics literature for the various experimental determinations of the value of some physical constant—which were often arrived at via quite different experimental methods—and then decide which values should be combined to produce an averaged best estimate for the constant, and which values should be discarded as outliers. This project often required that Birge go back to the raw data (measurement indications) from these experiments and reprocess the data himself using improved data correction methods to achieve a better value (or measurement outcome) for the constant. This project of estimating a best value for a given fundamental constant is complicated by the fact that many fundamental constants are part of an interdependent web, and hence cannot have their values fixed in isolation from the values of other fundamental constants. For example, the Rydberg constant, R_∞, is defined by Bohr's formula in terms of electron mass (m_e), electric charge (e), Planck's constant (h), and the speed of light (c), and hence the adopted values for all these fundamental constants must "form a self-consistent system, as judged by the Bohr formula for R_∞" (Birge 1929, p. 71). In order to see how this project of determining the values of fundamental constants relates to Kuhn's fifth law of thermodynamic*s*, it is helpful to examine some concrete examples.

The first fundamental constant that Birge discusses in his seminal 1929 article is the speed of light, *c*. Birge notes that there are three different cutting-edge experimental methods for determining the speed of light: First, there is Albert Michelson's (1927) method using a rotating mirror; second, there is Jean Mercier's (1924) method measuring the velocity of stationary electric waves along a wire; and a third method is Edward Rosa and Noah Dorsey's (1907) indirect determination from the measured ratio of electrostatic to electromagnetic units for an electric charge. The values for *c* obtained from these three different measurement approaches are listed in Figure 1.

Experimental Method	**Value Speed of Light (*c*)** km/sec	**Uncertainty/Error** km/sec
Michelson (1927)	299,796	± 4
Mercier (1924)	299,700	± 30
Rosa & Dorsey (1907)	299,710 (reported) 299,790 (Birge corrected)	± 30 ± 10

Figure 1: Values of speed of light based on different experimental determinations discussed in Birge (1929).

The first thing to note is that the experimental values for the speed of light in the center column are not all identical. While at one point in the history of physics they may have been considered to be in "reasonable agreement," by Birge's time, the demands of high-precision measurement physics required a consistent value to higher-resolution. At first glance, Rosa and Dorsey's *reported* value appears to be in closer agreement with Mercier's value, suggesting that Michelson's result is the outlier. However, as Birge notes, in calculating their value for the speed of light Rosa and Dorsey used the *international ohm*, and this needs to be converted to *absolute ohms* in order to obtain a proper value for the speed of light; hence, their result needs to undergo *data correction*. As Birge further emphasizes, it is essential that the "probable error" or uncertainty of a measurement value also be taken into consideration when determining whether or not results are in "reasonable agreement".[6] Birge goes on to reassesses the uncertainties associated with the Rosa-Dorsey measurement values by changing their *maximum* uncertainty value of ± 30 km down to a *probable* error value of ±10 km. When these data corrections are taken into account, it turns out their measured value for the speed of light is in closer agreement with Michelson's, and it is instead the Mercier value that is the outlier. Birge concludes that Rosa and Dorsey are in fact in "beautiful agreement with Michelson's recent value" (Birge 1929, p. 10) and given that the Rosa-Dorsey probable error is over twice that of Michelson's, Birge recommends adopting Michelson's result as the recommended value for *c*. Unfortunately, this "beautiful agreement" would not last, and the estimate for the speed of light would drop precipitously in the next decade, before rising again, as seen in Figure 3 (from Henrion & Fischhoff 1986, p. 793) which we discuss below.

Notice that in the case of the speed of light, Birge opted to use the other values of *c* only as a kind of coherence test to help pick out one "best" value.[7] Alternatively, he could have combined the various experimental values for the fundamental constant, such as by weighting and averaging the values in some way, or even just by expanding the associated uncertainty estimate.[8] In his excellent article on the history and philosophy of the adjustment of fundamental physical constants, Fabien Grégis (2019a) describes these two approaches as *arbitrage* and

[6] There are subtle differences between "uncertainty" and "probable error", which we do not engage here.
[7] For a discussion of coherence testing of multiple measurement methods, see for example Bokulich 2020b on radiometric dating.
[8] As we discuss below, in his later reassessment of the speed of light in 1941, Birge does opt for the "compromise" over "arbitrage" approach.

compromise respectively, noting that Birge adopts different methods for different constants at different times, depending on the nature of the discordant data. We'll come back to discuss the compromise approach in a moment, but first let us take a closer look at Birge's philosophical views underlying the process of adjusting fundamental physical constants.

In this same 1929 article, Birge lays out a remarkably Peircean[9] view on the process of determining the values of fundamental constants. In particular, Birge sees this as an ongoing, iterative project: "The need is continuous since the most probable value of to-day is not that of to-morrow, because of the never-ending progress of scientific research" (Birge 1929, p. 2). In addition, Birge highlights the social dimensions of this iterative project of determining the value of physical constants: "there is required the unbiased cooperation of many persons situated in scientific laboratories throughout the world" (p. 2). Although the true value of the constant may not be determinable at any given moment of history, it should be a probable value, and more importantly the same value must be coordinated and adopted across the entire scientific community. Though grounded in the best available scientific evidence, the particular value chosen at any given time will, as Birge acknowledges, involve an irreducible element of subjective judgement. For example, the adjuster must make a judgement about what data is "good" and hence will be incorporated into the adjustment, and what data is "bad," and hence may be discarded.

Another noteworthy aspect of Birge's philosophy of data is the careful attention he paid not only to the value of the fundamental constant, but also to its associated uncertainty or probable error.[10] In metrology, measurement uncertainty provides the boundaries to the likely range of values within which the "true" measurement value is supposed to lie. He writes, "Some estimate of the probable error is . . . as important as the constant itself" (Birge 1929, p. 4). Such a strong emphasis on uncertainty marks an epochal change in precision measurement physics (see Cohen and DuMond 1957; Gregis 2019a). Birge reflects on this important turning point in the physics community in 1943:

> "Previous to the time that I began publishing critical values of the general physical constants [in 1929], it was rather exceptional to attach a probable error, or other measure of reliability, to suggested "most probable" values. It seemed to me, however, desirable to attempt such an estimate, even in cases where the estimate was admittedly almost a pure guess." (Birge 1943, p. 213)

Acknowledging the necessity of uncertainty estimates, even if highly speculative, is connected to the recognition, encoded in the fifth law of thermodynamics, that measurements are never a perfect revelation of the true value of any quantity.

Indeed, Birge saw this as one of the strongest arguments for using the now widely adopted least squares method[11] for combining the different measured values of a particular

---

[9] The influence of Charles S. Peirce's pragmatic philosophy on Birge is explored by one of us in another work (Bokulich, Book Manuscript In Progress).

[10] As we noted before, although "probable error" was the quantity and term used in Birge's time, today the metrology community prefers to use "uncertainty", rather than "error". See Grégis (2019b) for a more detailed discussion of the shift in the meaning of uncertainty.

[11] Birge in his 1932 expository article defines the least squares method as a statistical method for "(1) the calculation of the "most probable" values of certain quantities, from a given set of experimental data, (2) the calculation of the "probable error" of each of the quantities just evaluated, (3) the calculation of the reliability, or probable error, of the probable errors so evaluated."(Birge 1932, p. 207). The method goes back to the 18th century work Adrien Marie Legendre and Carl Friedrich Gauss's work in geodesy (see Struik 1954 for a history).

constant: “Now in order to evaluate the probable error it is necessary to use the method of least squares. One great objection, it appears to me, to certain methods which have been proposed as substitutes for least squares, is that they give no objective criterion for the error” (1929, p. 4). In a later work, Birge recalls his seminal role in spreading the use of the least squares method from astronomy and geodesy to the rest of physics: “When I started my work on the calculation of the general constants, the method of least squares was in common use by surveyors and by astronomers. But it was seldom used by anyone else” (Birge 1957, p. 42). Birge himself made many contributions to the refinement of the least squares method, which had an impact far beyond determining the values of constants. Birge also notes, however, the limits of the least squares method. He writes,

> “one must use some judgement in applying the method of least squares. Otherwise the results may well be absurd. Such a solution applies *only* to observations which are affected merely by accidental errors of observation. If a particular observation deviates too widely from a smooth curve, it should be rejected before attempting to treat the data by least squares” (Birge 1929, p. 6).

He recognizes here that although the least squares method is well-suited for handling *random errors*, it does not address the possibility of *systematic errors* that will skew the result in a particular direction. Another problem with least squares, noted by subsequent researchers (Taylor et al. 1969, p. 379) and discussed by Grégis (2019a, p. 49), is that by discarding outliers, this method leads to an underestimation of uncertainty and an overconfidence in the convergence of results.

Later, in his 1957 article, Birge grapples with a further problem, now often referred to as the *bandwagon effect* (or intellectual phase locking), which he defines as follows: the “tendency of a series of experimental results, at a certain epoch, to group themselves around a certain value”— even when that value turns out not to be correct. Convergence, or an agreement of measurement values, is often seen as a hallmark of their accuracy or truth. But, of course, there can be other explanations for such a convergence, apart from the measurements correctly hitting upon the “true value.” Birge recounts a conversation he had with his Berkeley colleague, the Nobel-prize winning experimental physicist Ernest Lawrence, who suggested an alternative explanation based on his own experience as an experimentalist:

> “In any highly precise experimental arrangement there are initially many instrumental difficulties that lead to numerical results far from the accepted value of the quantity being measured. . . . [T]he investigator searches for the source or sources of such errors, and continues to search until he gets a result close to the accepted value. *Then he stops!* . . . In this way one can account for the close agreement of several different results and also for the possibility that all of them are in error by an unexpectedly large amount.” (Birge 1957, p. 51)

Interestingly, Kuhn draws attention to this same phenomenon in his “Function of Measurement” paper when he notes that measurements can be ‘self-fulfilling prophecies’ in the sense that they are adjusted to conform with an expected standard (Kuhn 1961 [1977], p. 196). Birge offers as a partial solution to this problem the use of multiple *different* experimental approaches to measuring a quantity, which insofar as they involve very different experimental paths are less likely to suffer from the same systematic errors. Given the common target quantity and much of the common (potentially erroneous) background knowledge, it is unlikely that all systematic errors can be avoided in this way. As we will see next, Birge's own work on the adjustment of fundamental physical constants would also turn out to be plagued by the bandwagon effect.

The experimental work on the speed of light in the decades following Birge's 1929 recommended value turned out to be an illustration of the bandwagon effect in action, as Max Henrion and Barauch Fischhoff (1986) have argued. In 1941 Birge published a reassessment of the value of the speed of light, prompted by new experimental determinations of *c* that some in the physics community took as evidence that the speed of light was not in fact a constant, but rather was either steadily decreasing or varying sinusoidally.[12] Birge rejected this as nonsense—to echo Kuhn, the theory of the constancy of the speed of light was, for Birge, never in doubt, rather the anomaly was to be located in the data, which had be forced into conformity with the theory. Birge then turned back to Michelson's (1927) value to identify previously unrecognized systematic errors, such as Michelson's mistaken use of the *wave* index of refraction instead of the correct *group* index of refraction in correcting from the measurement in air to the needed speed in a vacuum, applying the needed corrections to Michelson's original data, and even exploring possible tectonic changes that might have affected the measured distances of the base line used in the experiment (Birge 1941, p. 93). Birge similarly revises the Rosa and Dorsey (1907) value, removing a rounding error and substituting in an updated value for one of the quantities in the calculation, again illustrating the ongoing, iterative process of *data correction* (e.g., Bokulich 2020a; Bokulich and Parker 2021). To these two determinations he adds six more recent experimental measures, concluding that these eight values for the speed of light are all that need to be taken into account in generating the most probable value for *c*.

Unlike in his 1929 determination of the speed of light where he opted for *arbitrage*, selecting one best value, in this 1941 reassessment of the constants he adopts the *compromise* approach, weighting the different values according to the reciprocal of the square of the probable error to obtain the value 299,776 ±4 km/sec. This value is considerably smaller than Birge's 1929 adopted value of 299,796 ±4 km/sec—and notably well outside the uncertainty bounds of the previous estimate. To address the claim that the value of the speed of light might actually be changing, Birge calculates the weighted average of five measurements of the speed light carried out before the Rosa and Dorsey (1907) experiments. Birge concludes that these

> "five older results [yielding by a weighted average 299,873 km/sec] . . . are entirely consistent *among themselves*, but their average is nearly 100 km/sec greater than that given by the eight more recent results. The cause of the sudden change in the experimentally determined values of ***c***, at the opening of the 20th century might be an interesting subject for investigation, but I would hesitate to believe this 100 km/sec change is real." (Birge 1941, p. 100)

The fact that the five older values for *c* agreed well with each other, and that the eight new values also agreed well with each other—but not with the five older values—suggests two possible hypotheses: First, this robust result across multiple experiments from these two different time periods could be interpreted as evidence that the speed of light was in fact physically changing over time—a possibility that, as we saw earlier, Birge rejected. Second, this differential clustering of experimental values for the speed of light at two different time periods could alternatively be interpreted as a *social* phenomenon—the bandwagon effect.

---

[12] For a discussion and references see Birge 1941, p. 92; Birge 1957, p. 50; and Henrion & Fischhoff 1986, pp. 793-794.

Table 1. Velocity of light

Author	Method	Epoch	Corrected result	Adopted probable error r	Adopted weight $100/r^2$	Original published result
Cornu-Helmert	TW	1874·8	299990	200	0·0025	299990
Michelson	RM	1879·5	299910	50	0·0400	299910
Newcomb	RM	1882·7	299860	30	0·1111	299860
Michelson	RM	1882·8	299853	60	0·0278	299853
Perrotin	TW	1902·4	299901	84	0·0142	299901
Rosa-Dorsey	EU	1906·0	299784	10	1·000	299710
Mercier	WW	1923·0	299782	30	0·111	299700
Michelson	RM	1926·5	299798	15	0·444	299796
Mittelstaedt	KC	1928·0	299786	10	1·000	299778
Michelson, Pease and Pearson	RM	1932·5	299774	4	6·250	299774
Anderson	KC	1936·8	299771	10	1·000	299764
Hüttel	KC	1937·0	299771	10	1·000	299768
Anderson	KC	1940·0	299776	6	2·778	299776

TW=toothed wheel ; RM=rotating mirror ; EU=electric units ;
WW=waves on wires ; KC=Kerr cell.

Figure 2: Birge's 1941 table showing the clustering of the 5 earlier experimentally measured values for the speed of light around a significantly higher value than the 8 later measurements that cluster around a much lower value for *c*—a difference he interpreted as a sort of Bandwagon effect, rather than evidence that the constant was really changing. (Reprinted with permission of IOP Publishing.)

Focusing on the eight more recent determinations of the speed of light (listed in Fig. 2 above), Birge writes, "[T]hese eight results, obtained by *six* different investigators, using *four* completely different experimental methods, agreed. . . . one would scarcely anticipate that the several final systematic errors should all be in the *same* direction and of roughly the *same* magnitude" (Birge 1957, p. 50). But as Birge had realized by 1957, this is in fact just what had happened! Although Birge had hoped in 1941 that "after a long and, at times, hectic history, the value of *c* [had] at last settled down" (1941, p. 101), the consensus value for *c* was about to jump again, up to around 299,792.4 km/s. Henrion and Fischhoff (1986) provide the following helpful graph showing how estimates of the speed of light changed dramatically in this period of 1929 – 1973 (Fig. 3)

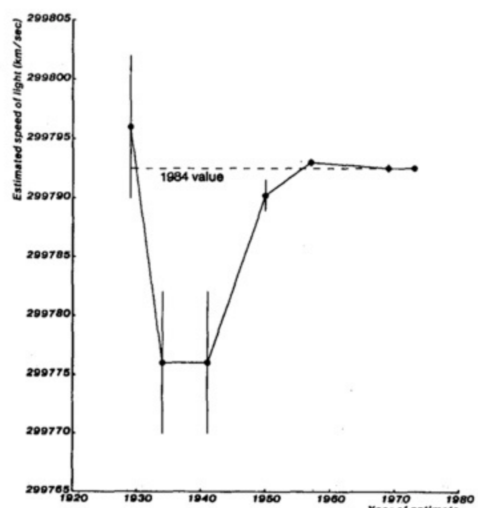

Figure 3: Recommended values for speed of light from 1929-1973, from Henrion and Fischhoff 1986, p. 793).

These remarkably large swings in the "consensus" recommended values of *c* within a space of just 30 years sent ripples throughout the physics community. In his 1957 paper describing this surprising reversal, Birge remarks,

> "This may well be the last paper I will ever write on the subject of the general physical constants. For that reason I should like to take the opportunity to consider briefly some aspects of the human side of the subject. . . . For if I have, to any degree, succeeded in calling attention to the numerous pitfalls that menace every research worker in science, and that lead so often to false results and conclusions, I consider that to be a far more valuable accomplishment than any specific scientific advance." (Birge 1957, p. 39)

It was in this growing mood of crisis and turn to philosophy by one of the most influential members of the physics department, that marked Kuhn's first year at Berkeley. It is no wonder that Kuhn decided to formulate the fifth law of thermodynamics as the centerpiece of his philosophy of measurement and data, just a few years later.

1957 also marked a symbolic passing of the baton from Birge to another Berkeley physics colleague, Kenneth Crowe, and two other California physicists, E. Richard Cohen and Jesse DuMond, who took up Birge's adjustment project with the publication of their book *The Fundamental Constants of Physics*, which they dedicated to Birge that same year. It should be emphasized that the speed of light was not the only fundamental constant whose experimentally determined values were misbehaving.[13] Similar problems plagued the value for electron charge, *e*, in the wake of R. A. Millikan's oil-drop experiments, which, as Cohen, Crowe, and DuMond discuss in this book, had a "systematic error [that] remained completely unsuspected for a period of about 15 years. . . . Because of the great importance of *e* and its close relationship to many other atomic constants this error had quite far-reaching effects" (p. 116). There were similarly problems with Newton's universal constant of gravitation, *G*—which was one of the fundamental constants whose determination Kuhn singles out in the *Structure* as being the object of repeated effort by experimentalists ever since the 1790s ([1962] 1996, p. 27).[14] It was becoming increasingly clear that the project of updating and coordinating a consistent set of values for all the fundamental physical constants, with an ever-growing influx of new experimental data, was neither a project for just one individual, nor a project that was ever truly finished. By the late 1960s, the Committee on Data for Science and Technology, known as CODATA, was formed to oversee the project for the entire physics community, and the readjustment of fundamental physical constants would eventually come to be regularized to every four years.

---

[13] In 1983 it was decided that *c* would no longer be empirically determined, but instead would become conventionally defined through the redefinition of the meter, tying both to the standard second, which is given in terms of a cesium atom transition (see Tal 2011for discussion of standard second and see Quinn 2011 for conventional stipulation of *c* through redefinition of meter). That is why when you look up the value for *c* today it says 299,792,458 m/s "exact"—it is because the value became one that was conventionally stipulated, not because of any change in experimental methodology that eliminated all uncertainty—something that would violate Kuhn's *fifth law of thermodynamics*! As W. Rowley further clarifies, "in making the speed of light a fixed constant, we are not attempting to dictate the laws of nature, but merely changing the viewpoint. We are not stating that the speed of light can never change; rather that, if it does, then the size of the metric length unit will change in sympathy so that the numerical value is preserved" (Rowley 1984, p. 284).

[14] We will return to briefly discuss ongoing efforts to determine the gravitational constant, G, in the following section.

4. Data, Anomalies, & the CODATA Philosophy of Metrology

The subsequent history of the process of adjusting the fundamental physical constants is an important one that is only now beginning to receive attention from philosophers of science (e.g., Smith 2010; Grégis 2019a). Relevant for our project here are the number of striking points of similarity between the emerging philosophy of metrology espoused by the CODATA metrologists and Kuhn's own philosophy of measurement.[15] Cohen and DuMond articulate their philosophical approach to the adjustment of physical constants more clearly in a subsequent review paper, where they open with a discussion of the moving goal post of what Kuhn describes as "reasonable agreement," and how the discovery of anomalies propels this process forward:

> "[T]he very process of improvement in accuracy and reliability (which the specialists in reviewing the constants themselves stimulate by calling attention to the discrepancies and troubles) whets the appetite for increasing precision, so that the discrepancies, which would have been of negligible magnitude a few years before, become of increasing importance." (Cohen and DuMond 1965, p. 538)

What counts as reasonable agreement is not fixed once and for all, but rather evolves with the increasing standards of precision as the program to measure a fundamental constant unfolds. To put it in Kuhnian terms, it is only by knowing in detail what to expect the measurement values to be, that scientists can recognize a *quantitative anomaly* in the numbers not turning out as expected. This leads to corrections in either the measurement process, data, or background theory, which in turn yields more precise expectations. As we saw with the case of the value of the speed of light, this is often not a linear process of convergence, but it is one where, gradually over time, the expectation of number of decimal places to which results should agree increases.

Within this CODATA community, the search for anomalies, or what they call discrepancies, is one of the most important parts of the readjustment process—perhaps even more important than the new value of the constant itself. Cohen and DuMond write, "it should be clear that the prime object of these re-evaluations of the constants must always be to *look for discrepancies* and to resolve them by finding errors in either theory or experiment which account for them" (1965, p. 540). One implication of this approach, is the recommendation *not* to expand the uncertainty estimates attached to values of the fundamental constants in the hope that the "true value" will be contained within that expanded uncertainty, as the "safety" approach would recommend. Grégis (2019a) has described this as the dilemma of *safety* versus *precision* in the philosophy of measurement, noting that for the CODATA scientists, *precision* was to be favored over safety, because of its ability to more readily reveal *anomalies*. Although the values of the constants are less likely to be revised outside of the previous uncertainty bounds on the "safety" approach, doing so makes the measurements a less sensitive instrument for detecting anomalies. Since it is the disagreement and discrepancies that drive science and lead to new discoveries, the narrower uncertainty estimations of the "precision" approach are to be preferred.

The dismissal of "safety" by the CODATA group working on the revisions of fundamental physical constants also relates to another point of overlap with Kuhn's philosophy,

---

[15] Due to the lack of citations between Kuhn on the one hand and his colleagues Birge, Crowe, Cohen, and DuMond on the other hand, we make no strong pronouncements about arrows of causation, noting only points of similarity and possible synergy. Our primary interest is in the continuing value of these two threads of ideas for understanding philosophy of measurement and data today.

namely a fundamental skepticism about scientists' ability to ever come to know the true value of a fundamental constant. Cohen and DuMond express this antirealism when they write,

> "No one can guarantee that an evaluation of the fundamental constants at a given epoch yields the 'true' values. Absolute truth, if these words have any meaning, is beyond the realm of physics. All we can do at each time of re-evaluation is to try to determine a set of values which, in the sense of least squares, and in the light of accepted theory at that time, does least violence to a chosen budget of observational data then believed to be the 'best'." (Cohen and DuMond 1965, p. 540)

Like Kuhn, these physicists take absolute truth to be unattainable, and so outside of scientific practice, instead adopting a more pragmatic or instrumentalist view. The antirealism of the CODATA metrologists seems to be shaped by the same general considerations that influenced Kuhn: first, a kind of skeptical induction arising from the surprising twists and turns of the history of readjustments, and second, a kind of two-world metaphysics, where true values are something forever beyond our reach.[16]

A third point of similarity between Kuhn and these metrologists involved in the readjustment of fundamental physical constants is an emphasis on the values of consistency and coherence as key arbiters in a conflict between different experimental results, coupled with a commitment to scientific holism. Cohen and DuMond write,

> "[F]ew physicists or chemists fully realize in what a complicated, intricate way the fundamental constants, together with the measurements from which they are derived, are interconnected and interrelated. Everything depends upon everything else . . . and one flaw in the picture propagates its defect, to a greater or lesser extent, throughout all the numerical values of the fundamental constants and conversion factors we seek." (1965, p. 538)

This is reminiscent of Kuhn's idea of the knowledge within a paradigm forming a strongly interconnected web that if modified has to be all together "shifted and laid down again on nature whole" (Kuhn [1962] 1996, p. 149). The fact that the fundamental constants of physics are interconnected, exhibiting this holism, also means that they must be evaluated together, and since there is no external arbiter, they must be assessed by the values of consistency and coherence:

> "[T]he greatest merit in a re-evaluation of the constants resides not in the numerical output values . . . but in the fact that the reevaluation constitutes a new test of the validity of all our theoretical preconceptions and their experimental verification over the widest possible domain. The only test of such validity we have is the consistency of the data, and this is indeed all we ask for." (Cohen and DuMond 1965, p. 540)

This emphasis on holism and the search for quantitative anomalies has come to be a central part of CODATA's philosophy of metrology, as we also see in the writings of Barry Taylor, a metrologist who joined the U.S. National Bureau of Standards (now NIST) in 1970 and who is still involved in the most recent readjustment (Tiesinga et al. 2021). To emphasize the fallible and ever-iterative project of determining the values for fundamental constants, Taylor recommends that all reported values be accompanied by a warning label:

---

[16] For an introduction to these arguments for antirealism and some realist responses, see for example McMullin (1984).

"[I]n order to bring home to the average worker that a set of recommended constants is not inviolate and handed down on stone tablets, every table of constants, whether original or reprinted, should probably start off with some type of warning label (preferably in large bright red letters) such as:

Warning!

Because of the intimate relationships which exist among least-squares adjusted values of the fundamental constants, a significant shift in the numerical value of one will generally cause significant shifts in others." (Taylor 1971, p. 497)

Like for Kuhn, there is a delicate balance to be struck by the scientific community between, on the one hand, typically suppressing anomalies by excluding outliers from the least-squares adjustment, and on the other hand, periodically allowing these anomalies to trigger an adjustment that, because of this tightly interconnected web, can end up having far-reaching consequences, even for other accepted values that had been thought secure.

The periodic adjustment of fundamental physical constants is still an ongoing project today, and since 1998 has been institutionalized by CODATA to be undertaken every four years.[17] Figure 4 shows the results of the most recent (2018) adjustment for twenty-seven fundamental physical constants in relation to the previous 2014 values.

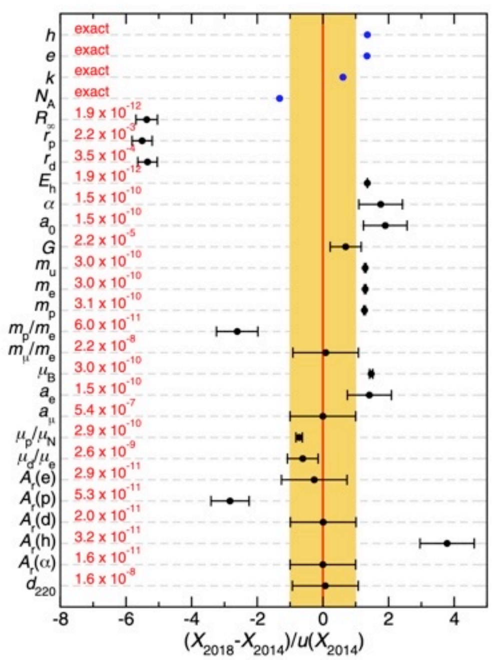

Figure 4: Comparison of the new 2018 values for 27 fundamental physical constants (listed on y-axis) with the previously recommended 2014 values for those constants. The vertical solid red line and yellow band represent the 2014 values and their standard uncertainties. The black circles with

[17] The next adjustment slated for 2022 has not been released yet, and typically lags by a year or so from the official date. More information about the CODATA Task Group on Fundamental Physical Constants and the periodic readjustments can be found here: https://codata.org/initiatives/data-science-and-stewardship/fundamental-physical-constants/

error bars show the difference between the 2018 and 2014 values divided by standard uncertainty of 2014 value (From Tiesinga et al. 2021, p. 55).

Particularly noteworthy is how many constants have updated values that fall outside of the uncertainty bounds of their previously recommended values. The constants without error bars that are listed as "exact" are ones that are no longer empirically determined values, but rather are conventionally defined values due to revisions in the International System of units (SI).

Let us return to the fundamental physical constant that was the primary focus Kuhn's discussion of constants in the *Structure*—the gravitational constant, *G*—and assess the status of current efforts to determine its value. As Kuhn notes, *G* does not appear in Newton's *Principia*, and it was only introduced later when his universal law of gravitation was formulated as the following equation:

$$F = G\frac{m_1 m_2}{r^2}$$

The gravitational constant, or "big G" as it is sometimes called, has been the subject of the longest program of experimental investigation of any of the fundamental constants, but at the same time has also had the most problematic history in terms of resisting a reduction in the uncertainty of its value. Kuhn cites a classic review article by J.H. Poynting that "reviews some two dozen measurements of the gravitational constant between 1741 and 1901" (Kuhn [1962] 1996, p. 28). A more recent review of experimental determinations of G that also begins with Poynting's review, but brings it up to the present, is that of Christian Rothleitner and Stephan Schlamminger (2017). They provide the following figure (Fig. 5) summarizing recent measures of G from 1982 through 2014:

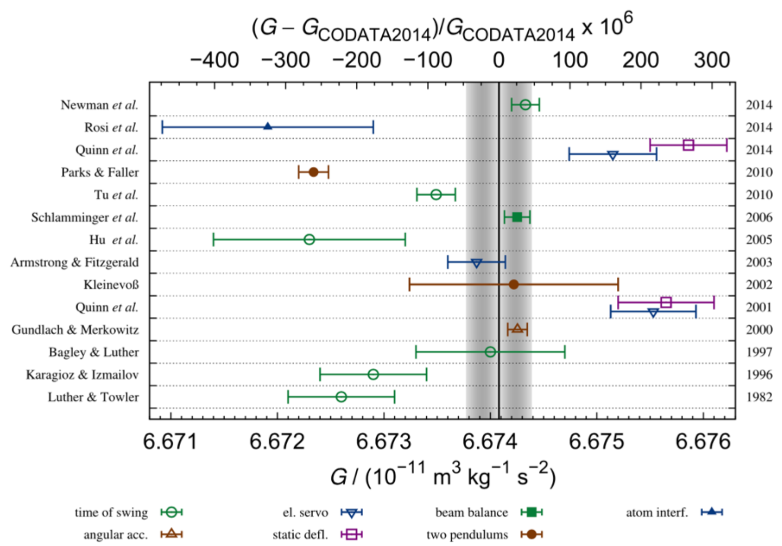

Figure 5: Measurement of the gravitational constant, *G*, from 1982 through 2014, involving seven different experimental methods, with the 2014 CODATA recommended value given by the vertical black line. (From Rothleitner and Schlamminger 2017, p. 22; with permission from AIP Publishing).

One of the continuing problems with the gravitational constant (*G*) is that there is a large spread in the data arising from different experimental methods, far more than the uncertainties identified in any individual measurement. As Rothleitner and Schlamminger note, there are three possible explanations for the inconsistent data:

> "1. Some or all of the experiments suffer from an unknown bias . . . [or] systematic effect that shifts the measured result from the true value by a predictable amount. . . .
> 2. Some or all of the experiments underestimate the relative uncertainty of the measurement. Hypothetically, all of the reported values of the measurements may be correct, but the uncertainties reported may be too small. If the true uncertainty were five times larger, the data set would be perfectly consistent. . . .
> 3. The most exciting, yet least probable explanation is that new unknown forms of physics can explain the variation in the data." (2017, p. 22)

The experimental measurements of *G* are thus a very dramatic illustration of Kuhn's fifth law of thermodynamics that no experiment yields quite the expected value. Even more troubling is that the different experimental determinations of *G* are not in "reasonable agreement"—their scatter is far outside of the uncertainty bounds attached to each experiment. As in Kuhn's day, the value of the gravitational constant remains a puzzle, and it is not yet known whether this quantitative anomaly will someday be resolved within the current physical paradigm, or will turn out to someday require fundamentally new physics, hence precipitating a scientific revolution.

5. Conclusion: Kuhn's Account of Progress in Measurement

The history of experiments to determine the values of the fundamental constants offers a striking illustration of Kuhn's *fifth law of thermodynamics*, with no experiment giving quite the expected result. Instead, determining the values of the constants involves a long process of data wrangling and remeasurement in an effort to iteratively improve their *reasonable agreement* with both theory and other experimental data. Kuhn knew well that there was a long history of troubling anomalies for even the most central of constants, such as the speed of light (*c*) and the gravitational constant (*G*). While Kuhn was working out his views on the "fifth law" and the philosophy of measurement at Berkeley in the years leading up to the *Structure*, the experimentalists in the Physics Department were simultaneously struggling with the realization that social phenomena, such as the bandwagon effect or "intellectual phase locking", could influence experimental data. As we saw, this was coupled with an emerging philosophy of metrology that viewed any talk of "true values" as beyond the realm of physics. Instead, the adequacy of an interdependent web of fundamental constants was to be determined by the values of consistency and coherence. There was thus a remarkable synchronicity—if not synergy—between these two intellectual developments concerning the history and philosophy of physics at Berkeley in the late 1950s and early 1960s.

In his later reflections, Kuhn remarked on what a central role this paper on the philosophy of measurement had for his thinking in the *Structure*: "Earlier at Berkeley I was asked to do a command performance . . . on 'the role of measurement in xyz.' . . . The paper that ultimately emerges is *The Function of Measurement in Physical Science*, and that really was extremely important . . . that's where the notion of normal science enters my thinking" (Kuhn 2000, p. 295). Despite the central role of this work in Kuhn's own thinking, his views about the philosophy of measurement and philosophy of data have been largely eclipsed by the theory-centric focus of the *Structure*. This is unfortunate because it obscures a key notion of scientific progress that Kuhn identifies in this 1961 work—one that goes beyond the instrumentalist increase in puzzle-solving ability identified in the Postscript to the *Structure* (p. 206). In this pre-*Structure* paper he writes, "I know of no case in the development of science which exhibits a loss of quantitative accuracy as a consequence of the transition from an earlier to a later theory. . . . Probably for the same reasons that make them particularly effective in creating scientific crises, the comparison of numerical predictions, where they have been available, has proved particularly successful in bringing scientific controversies to a close." (Kuhn [1961] 1997, p. 213; italics removed)

This emphasis on quantitative accuracy, or perhaps more accurately *quantitative resolution* (understood as an ever-increasing number of decimals places required for reasonable agreement) is a remarkably robust and paradigm-neutral form of progress for Kuhn to have identified. And it is moreover a type of progress that has traditionally been neglected, along with a broader neglect of measurement and data in the philosophy of science until relatively recently. More generally, we hope that by situating Kuhn's views in the *Structure* within both the broader philosophical context of his contemporaneous views on measurement and data, and the broader historical context of work being done by Kuhn's colleagues at Berkeley to determine the values of fundamental physical constants, we can gain a deeper appreciation of the extent to which Kuhn's philosophy remains relevant for the philosophy of metrology and philosophy of data today.

References

Birge, R. T. (1929). "Probable Values of the General Physical Constants." *Reviews of Modern Physics* 1 (1): 1–73.

__________(1932). "The Calculation of Errors by the Method of Least Squares" *Physical Review* 40: 207 – 227.

__________(1941). "The General Physical Constants: As of August 1941 with Details on the Velocity of Light Only." *Reports on Progress in Physics* 8 (1): 90–134.

__________(1943). "Comments on" The Probable Accuracy of the General Physical Constants"." *Physical Review* 63.5-6 (1943): 213.

_________(1957). "A Survey of the Systematic Evaluation of the Universal Physical Constants." *Il Nuovo Cimento* 6 (S1): 39–67.

Bokulich, A. (2020a). "Towards a taxonomy of the model-ladenness of data." *Philosophy of Science* 87, no. 5 (2020): 793-806.

__________(2020b). "Calibration, Coherence, and Consilience in Radiometric Measures of Geologic Time." *Philosophy of Science*, *87*(3), 425-456.

Cohen, E. R., & DuMond, J. W. (1957). The fundamental constants of atomic physics. In *Atoms I/Atome I* (pp. 1-87). Springer, Berlin, Heidelberg.

__________(1965). Our knowledge of the fundamental constants of physics and chemistry in 1965. *Reviews of Modern Physics*, *37*(4), 537-594.

Cohen, E. R., Crowe, K. M., & Dumond, J. W. (1957). *The fundamental constants of physics* (Vol. 1). Interscience Publishers.

Duhem, P. ([1914] 1954). *The Aim and Structure of Physical Theory*, 2nd edition. Translated from French by Marcel Rivière Princeton: Princeton University Press.

Grégis, Fabien. (2019a). "Assessing Accuracy in Measurement: The Dilemma of Safety versus Precision in the Adjustment of the Fundamental Physical Constants." *Studies in History and Philosophy of Science Part A* 74: 42–55.

___________ (2019b). "On the Meaning of Measurement Uncertainty." *Measurement*, *133*, 41-46.

Helmholz, A. C. (1980). "Raymond Thayer Birge." *Physics Today* 33 (8): 68–70.

Henrion, M. and Fischhoff, B. (1986). "Assessing Uncertainty in Physical Constants." *American Journal of physics*, *54*(9), pp.791-798.

Kuhn, Thomas S. ([1961] 1977). "The Function of Measurement in Modern Physical Science." *Isis* 52 (2): 161–93. Reprinted in The Essential Tension: Tradition and Innovation in Scientific Research." University of Chicago Press, 2011. Pp. 178-224.

_____________ ([1962] 1996). *The Structure of Scientific Revolutions*. 3rd ed. Chicago, IL: University of Chicago Press.

_____________ (2000). *The Road since Structure: Philosophical Essays, 1970-1993, with an Autobiographical Interview.* Chicago, IL: University of Chicago Press.

McMullin, E. (1984). "A Case for Scientific Realism." *Scientific Realism*, ed. J. Leplin, 8–40. Berkeley: University of California Press.

Mercier, J. (1924). De la Synchronisation Harmonique et Multiple." J*ournal de Physique et le Radium*, 5, 168–179.

Michelson, A. A. (1927). "Measurement of the Velocity of Light Between Mount Wilson and Mount San Antonio." *The Astrophysical Journal*, 65(1), 1–14.

Poynting, J. H. (1910-1). "Gravitation Constant and Mean Density of the Earth." *Encyclopaedia Britannica* (11th ed.; Cambridge, 1910–11), XII, 385–89.

Quinn, T. (2011). "Time, the SI and the Metre Convention." *Metrologia*, *48*(4), S121.

Rosa, E. B., & Dorsey, N. E. (1907). "A New Determination of the Ratio of the Electromagnetic to the Electrostatic Unit of Electricity." Bulletin of the Bureau of Standards, 3(3), 433–540.

Rothleitner, C. & Schlamminger, S. (2017) "Invited review article: Measurements of the Newtonian constant of gravitation, G." *Review of Scientific Instruments* 88, no. 11: 111101.

Rowley, W. R. C. (1984). "The Definition of the Metre: from Polar Quadrant to the Speed of Light. *Physics Bulletin*, *35*(7), 282.

Smith, G. E. (2010). "Revisiting Accepted Science: The Indispensability of the History of Science." Edited by Sherwood J. B. Sugden. *Monist* 93 (4): 545–79.

___________(2014)"Closing the Loop: Testing Newtonian Gravity, Then and Now." *Newton and Empiricism*, ed. Z. Biener and E. Schliesser, 262–351. Oxford: Oxford University Press.

Struik, D. (1957). *A Concise History of Mathematics*. London: Bell and Sons, Ltd.

Tal, E. (2011). "How accurate is the standard second?." *Philosophy of Science* 78, no. 5: 1082-1096.

Taylor, B. N. (1971). "Comments on Least-Squares adjustments of the Constants." In D. N. Langenberg, & B. N. Taylor (Eds.). *Precision measurement and fundamental constants* (pp. 495–498). National Bureau of Standards.

Taylor, B. N., W. H. Parker, and D. N. Langenberg. 1969. "Determination of e h , Using Macroscopic Quantum Phase Coherence in Superconductors: Implications for Quantum Electrodynamics and the Fundamental Physical Constants." *Reviews of Modern Physics* 41 (3): 375–496.

Tiesinga, E., Mohr, P.J., Newell, D.B. and Taylor. B. (2021). "CODATA Recommended Values of the Fundamental Physical Constants: 2018." *Reviews of Modern Physics* 93 (2): 025010.