Croatian Journal of Philosophy
Vol. XVIII, No. 54, 2018

On the Measurability of Measurement Standards

PHIL MAGUIRE and REBECCA MAGUIRE
National University of Ireland, Maynooth, Ireland

Pollock (2004) argues in favour of Wittgenstein's (1953) claim that the standard metre bar in Paris has no metric length: Because the standard retains a special status in the system of measurement, it cannot be applied to itself. However, we argue that Pollock is mistaken regarding the feature of the standard metre which supports its special status. While the unit markings were arbitrarily designated, the constitution, preservation and application of the bar have been scientifically developed to optimize stability, and hence predictive accuracy. We argue that it is the 'hard to improve' quality of stability that supports the standard's value in measurement, not any of its arbitrary features. And because the special status of the prototype is tied to its ability to meet this external criterion, the possibility always exists of identifying an alternative, more stable, standard, thereby allowing the original standard to be measured.

Keywords: Measurement standard, stability, accuracy, prediction.

1. *Introduction*

Wittgenstein (1953: 29, §50) makes the following claim:

> "There is *one* thing of which one can say neither that it is 1 metre long, nor that it is not 1 metre long, and that is the standard metre in Paris. – But this is, of course, not to ascribe any extraordinary property to it, but only to mark its peculiar role in the language-game of measuring with a metre-rule."

Pollock (2004) argues that this claim is correct. Because the prototype metre has a special status in the metric system, it cannot be measured within that system. His view is that there is no fundamental unit of length beyond the prototype metre. It is not the case that bar was selected because it happens to match an *a priori* concept of the metre. Rather, the bar is the essence of the metre and measuring metric length does not make sense without it.

According to Pollock (2004: 153), "measurement consists in nothing more than the comparison of the object of measurement with some (arbitrarily chosen) standard". In other words, the value of measurement comes not from some intrinsically meaningful process of evaluation, but from a process of comparison enabled by an arbitrary standard. When we ask "how long is that object?" we are not seeking information about its true length, whatever that might mean. After all, we can see exactly how long the object is. What we want to know is how its length compares with that of other objects, a comparison process which requires some arbitrarily selected standard for quantifying length, be it metres, feet, hands or fingers. Pollock explains that "measurement simply consists in determining the ratio of one object's length to the length of some standard". Wittgenstein (1953: 103, §279) makes a related observation, highlighting the meaninglessness of measurement without comparison:

> "Imagine someone saying: 'But I know how tall I am!' and laying his hand on top of his head to prove it."

Taking another example, somebody might step outside on a warm day and exclaim "I wonder how hot it is?" Clearly they can feel how hot the air is, it's touching their skin. But measurement is not about providing independent, theory-free descriptions of phenomena, it's about relating things together. What this person wants to know is how the air temperature today compares with that of previous days. In sum, the utility of measurement comes about, not from the result of the measurement itself, but from the comparisons it enables.

Pollock (2004: 152) argues that this feature of measurement, a system for comparison rather than description, has been lost on the majority of philosophers: "...philosophers simply [do] not understand the concept of measurement". Salmon (1986: 210) is proposed as epitomising this confusion:

> "...if the reference-fixer does not know how long S is, he cannot know, and cannot even discover how long anything is. Measuring an object's length using S only tells him the ratio of that object's length to the length of S."

Also (208):

> "If one knows only that the length of the first is n times that of the second without knowing how long the second object is, one knows only the proportion between the lengths of the two objects without knowing how long either object is."

Pollock (2004: 149) states that Salmon here demonstrates a failure "to understand the very concept of measurement, as well as what it means to know the length of something." There is nothing to the act of measurement beyond expressing relationships between objects. There is no absolute scale of measurement, no apodictic system for quantifying unit length. And with no natural unit, there is no concept of measurement beyond an arbitrarily selected standard being used to express length ratios. For Pollock, that's all there is to measurement.

The question "how long is that object?" presupposes a system of measurement involving a standard of comparison. Because the prototype metre is a necessary condition for the existence of the metric system, the question "how long is the standard metre?" is not a proper question. The standard is a criterion for measuring in the metric system, and it makes no sense to apply a criterion to itself (Pollock 2004: 155). The description of the prototype metre as "one metre long" only functions as a name or label, not as the description of a measurable length of the object. Pollock (2004) therefore concludes that Wittgenstein is correct: We cannot say that the standard metre is a metre in length, or that it is not a metre in length.

In summary, Pollock's (2004) argument hinges on first, the idea that the standard metre is identified arbitrarily, and second, that it retains a special status in the metric system and so cannot be applied to itself. This article investigates whether these two assumptions are valid. In brief, we will argue the following: Pollock is correct in assuming that there is nothing to the act of measurement beyond expressing relationships between objects. However, he is wrong in assuming that this implies that measurement standards are selected arbitrarily (and hence immune to being measured themselves). Not all systems are equally capable of expressing relationships in a useful way. Specifically, standards that feature the property of *stability* are better, because they enable superior predictions. Because measurement standards are obligated to meet the external property of stability, they are susceptible to being improved upon, and thus open to being measured themselves. Although temporarily enjoying dominant status, working standards do not have immunity to being overthrown in the game of measurement. Accordingly, Wittgenstein's statement about the metre bar having no measurable length value must be wrong.

2. *A Brief History of Length*

Before examining Pollock's (2004) argument, we provide a brief overview of the history of the metre and the standard metre bar in Paris.

Measurement standards in medieval Europe varied widely between different jurisdictions, which were often little more than single market towns. The French revolution in 1789 provided the motivation to abolish the multitude of length measures associated with the *ancient régime* and replace them with a new decimal system based on a universal and easily replicable standard (Crease 2011).

The new movement towards standardization provoked much debate as to which environmental property could provide a globally recognizable standard. One proposal was to use the length of a "seconds pendulum", that is, a pendulum which swings through a half-period in exactly one second. However, it was soon discovered that the length of such a pendulum actually varies from place to place. For example, the French astronomer Jean Richer demonstrated a 0.3% difference in this length when calibrated in Cayenne (in French Guiana) versus Paris (Crease 2011).

In light of this, the commission for measurement reform eventually came to the decision that the new unit of length should be equal to one ten-millionth of the distance from the North Pole to the Equator, when measured along the meridian passing through Paris. This was a concept expressed in a single sentence that everybody on earth could agree on. During the surveying process, the commission ordered the production of a series of platinum bars based on preliminary calculations. Following the survey's completion, the bar with length closest to the meridional definition was identified. This bar, which subsequently became known as the *"mètre des archives"* was placed in the National Archives on June 22nd 1799 (Wikipedia "History of the metric system", 2018).

The simple meridional definition had been intended to ensure international reproducibility. In practice, however, nobody was in a rush to replicate a survey of the distance between the Equator and North Pole. The definition was so impractical to verify that it became irrelevant, being replaced instead by artefact standards. When it was later established that the circumference quadrant was actually 10,019km, as opposed to 10,000km, this had no bearing on the use of the metre. The use of artefacts was already providing a de facto standard, unconnected and arbitrary relative to any other worldly definition.

Countries adopting the metre as a legal measure during the 19th century purchased standard metre bars with which to calibrate their own national standards. These, however, were prone to wearing down with use. Because different standard bars in different countries were being worn down at different rates, there was no mechanism for verifying whether everybody was adhering to the same standard. In light of these difficulties, an international treaty, known as the Metre Convention, was signed in Paris on 20th May 1875. An organisation known as the Bureau International des Poids et Mesures (BIPM) was established in Sèvres, just outside Paris. This organization was entrusted with the responsibility of conserving prototypes and carrying out regular comparisons between different national standards, so as to ensure international consensus.

The BIPM set about creating a new state of the art international prototype metre, accompanied by a set of copies earmarked for international distribution. These bars were made of a special alloy, consisting of 90% platinum and 10% iridium, making them significantly harder than pure platinum. They were also fashioned in the shape of an X, thus minimizing the effects of torsional strain during length comparisons.

One of these bars was "sanctioned" to be identical in length to the *mètre des archives* on September 28th 1889, during the first meeting of the Conférence Générale des Poids et Mesures (CGPM). Following this moment of consecration, the new bar became the international prototype metre, and the old 1799 bar began to fluctuate in length.

In 1960, at the 11th CGPM, a new definition of the metre was agreed, based on wavelengths of radiation from the krypton-86 atom. In 1983, at the 17th CGPM meeting, the metre was redefined again in terms of the distance travelled by light in a vacuum per second.

For the purpose of analysing the validity of Wittgenstein's original claim and Pollock's (2004) defence of it, we will initially consider the role of the 1889 metre bar as an active standard, as it was in 1953 when Wittgenstein's comments were first published.

3. *Length versus Unit*

Pollock (2004) repeatedly emphasises that the selection of the standard is arbitrary, meaning that it is completely self-sufficient and has no connection with any external phenomena: The "standard is arbitrarily chosen and agreed upon by the community. Only practical considerations bar us from using anything at all as a standard" (154). Also: "This arbitrary nature of standards of measurement seems to be lost on many philosophers" (154).

Intuitively, the selection of prototypes by the BIPM does not seem arbitrary. For example, the prototype kilogram is deliberately forged of platinum-iridium alloy, an inert metal with very high density (to negate a buoyancy effect), extreme resistance to oxidation, low magnetic susceptibility and high resistance to contamination and wear. In addition, the artefact is carefully isolated under multiple nested bell jars and subject to periodic cleaning with ether and ethanol followed by steaming with bi-distilled water (Wikipedia "Kilogram", 2018). When Pollock describes the prototype metre as arbitrary, he is referring, not to its material, preservation and application, but to the markings on that bar which designate one metre. It is the designation of a *unit* that is arbitrary.

But what is the size of that unit? We propose that the size of the unit does not exist independently of the medium of the platinum-iridium bar onto which it is inscribed. The bar does the work of preserving the size of the unit, rendering the concepts of 'unit' and 'standard' inextricable. Asserting that the unit is arbitrary is therefore meaningless: there *is* no unit that can be addressed independently of its embodiment by the metre bar itself.

All measurement units depend on an underlying standard which embodies their size. For example, the Imperial and metric systems were originally associated with different processes for realizing their respective units. However, by 1964, the definition of the inch was *tied* to that of the metre, meaning that both units serve as different labels for describing measurements in the same fundamental system. To turn a measurement from centimetres to inches, one simply divides by 2.54. Although this ratio is one that arose by historical chance, it has no measurement value in itself, serving merely as a cosmetic treatment of an underlying measurement result. The value of both the Imperial and metric systems lies in the embodiment of unit length by a standard.

Pollock (2004) misses the idea that the size of a unit must be realized by some sophisticated practice, believing instead that the concept of objective length is universally appreciated following its 'discovery': "Although we discovered the concepts of length (and mass) we invented the concept of a metre for our own convenience; as a means of making judgments about length, which we could record and/or communicate to others" (154). Thus, for Pollock, the aspect of the prototype metre that gives it its special status in the metric system is not its role in enabling reliable judgments about length, but merely its role in designating a unit of measurement. In the following section we argue that Pollock's attitude overlooks a crucial property of measurement standards, namely that of *stability*.

4. *What Makes the Standard Special?*

Let's imagine what would happen if the only role of the metre bar was to designate a unit of measurement, as Pollock (2004) assumes, without any regard to realizing the size of that unit.

Under this scenario units of length could be perfectly replicated and maintained by any measured object. For example, I could take a wooden stick and mark on it exactly the same unit lengths as exist on the prototype metre bar. In Pollock's world the stick functions just as well as the original standard. Indeed, every act of measurement is equivalent to forging a new standard. Once my desk is identified as having some particular length value, it too becomes part of the standard, and, following the assumptions inherent to Pollock's view, ceases to have measurable length because of its new special role in the system. The original standard is not special anymore. We can no longer cling to Wittgenstein's statement that only *one* thing has no measurable length value: Every object which is measured becomes just as good at realizing length as the original standard, hence losing its property of measurable length. Pollock doesn't care what object the markings are made on. After all, the choice is arbitrary.

In practice, measurement does not work like this. The standard encompasses, not just the physical bar, but a whole set of procedures for handling, comparing and making copies of the bar, as well as the background knowledge and assumptions involved in those procedures. For example, in 1927 the defined *mise en pratique* of the prototype metre was altered, without affecting the prototype artefact, or its unit markings. At the 7th CGPM it was clarified that any measurement of the bar should now be "subject to standard atmospheric pressure, with the prototype 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 1928: 49). The preservation of a given measurement standard resides in the understanding of its *mise en pratique* by active practitioners; used improperly the metre bar might prove no more useful in measurement than a metre stick.

If we accept that some artefacts and procedures enable superior judgements about length (e.g. a platinum-iridium bar, when used in appropriate manner, makes a better standard than a stick) then we are admitting the existence of some external criterion that standards are intended to meet.

Consider, for example, a fanatical dictator who issues a diktat defining the length of his beard as the new standard for measurement. Relying on this unstable beard length might cause bridges to fall down, buildings to collapse, and ships to sink.

Is this a problem? If we maintain that a standard of length can be selected arbitrarily, then it has no obligations to achieve anything. It is only relative to the external goal-directed expectation that measurement standards should keep bridges up and ships afloat that we can describe the beard-length standard as wanting. In sum, an arbitrary standard, without external connection to any practical function, does not support the property that we intuitively understand as length.

Danjon (1929) highlights the difficulty of interpreting the ephemeris time standard (using the position of the sun, moon, planets and stars) as a fiat with no external obligations:

> …Although Newton's law has been saved, it is experiencing a quite extraordinary adventure: henceforth called upon to gauge the passage of time, it becomes in part unverifiable and ceases to be what could strictly be termed a law…Since we would ask these laws to provide a measure for the passage of time, we could no longer subject them to experimental control without entering into a vicious circle. (Danjon 1929)

Consistent with Danjon's critique, Chang (2001, 2004, 2007), van Frassen (2008, 2009) and Tal (2011, 2012, 2013, 2016) all reject the traditional view of arbitrary, apodictic definitions at the heart of measurement. Acknowledging the real-world application of measurement, they recognise the role of a 'hard-to-isolate' external criterion, supporting goal-directed activities, as being at the heart of the practice.

5. *What is Measurement For?*

To properly understand the role of the prototype metre in the metric system we need to consider what measurement is for in the first place.

Intuitively, people make measurements because measurements are valuable. But what is it about measurements that makes them useful? Tal (2012, 2013, 2016) proposes that the goal of prediction lies at the heart of measurement, insofar as measurement accuracy, and hence the calibration of scientific instruments, is defined in terms of predictive accuracy. When I measure the length of my desk I am effectively making a prediction about what will happen when it interacts with other measured objects (e.g. will my desk fit through that door?) Even if we imagine cases where measurement is carried out for its own sake, without any expectations for prediction, the concept of reliable relationships still applies. For example, somebody who measures how fast they run around a race track expects those timings to enable comparisons involving other runners, suggesting who would win a hypothetical race between them. In order to be of value, a measurement system must provide reliable information about the relationships between measured phenomena, information which enables accurate predictions.

Tal's (2012) goal-based view stands in contrast to the widespread supposition that measurement and prediction are distinct epistemic activities. He argues that traditional accounts of measurement have overlooked its practical role in prediction, ignoring the key associated concepts of uncertainty, reliability and inference. For example, theorists such as Campbell, Stevens and Suppes "took 'measurement' to be synonymous with either 'number assignment' or 'scale construction', and neglected the 'applied' aspects of measurement such as accuracy, precision, error, uncertainty, and calibration" Tal (2013: 1164). In practice, measurement outcomes are obtained from instrumental readings by a chain of inferences, and the inferences drawn depend on the particular theoretical and statistical assumptions associated with the measurement apparatus. According to Tal (2013: 1165), "this way of viewing measurement raises a host of representational questions that have been either neglected or only partially addressed by traditional accounts".

The idea that measurement might be goal-directed raises the issue of how a theoretical quantification could be coordinated with empirical measurement. The issue here is that the empirical adequacy of a given theory and the reliability of a related measurement process appear to depend on each other in a circular fashion (Tal 2013: 1160). For example, in order to establish a theory of weight, it is necessary to test the predictions of that theory, a task which itself requires a reliable method of measuring weights. Conversely, testing the reliability of such measurements presupposes existing theoretical knowledge about weight against which it can be calibrated (Tal 2013: 1160).

The traditional philosophical approach to this problem, which Pollock (2004) espouses, has been to assume that coordination is achieved, and circularity avoided, by establishing apodictic definitions for quantification, which are arbitrary, self-supporting and internally complete. These definitions are assumed to be "analytic statements that require no empirical testing" (Tal 2013: 1160), thus severing the link between measurement and any external goal-directed outcome, such as prediction. For example, Ernst Mach noted that different types of fluid expand at different nonlinearly related rates when heated and concluded that there can be no fact of the matter as to which fluid expands most uniformly, since the very notion of equality among temperature intervals has itself no determinate application prior to a conventional choice of standard thermometric fluid with which to establish it (Tal 2013: 1161). The eventual choice of standard, for Mach, was a conventional one. Poincaré similarly argued that the processes scientists use to mark equal time durations (e.g. pendulum swings) are chosen for the sake of convenience (Tal 2013: 1161).

Pollock (2004: 155) echoes a similar conventionalist sentiment when he insists that "we simply chose a length that we found convenient and *called* it a metre. That is all there is to choosing a standard of measurement." However, though the examples noted by Mach and Poincaré seem, at first blush, to indicate arbitrariness at the heart of measurement, this arbitrariness results from pushing measurement beyond the existing limits of science and technology, thereby exhausting justification. The arbitrary decision here is to choose between several highly sophisticated systems, each of which does so well at measuring that their various merits are hard to distinguish.

For instance, while Mach and Poincaré recognized that choices of coordinative principles are often constrained by considerations of simplicity and convenience, they were not suggesting that these choices are completely arbitrary, but rather that working standards are selected because they are "good enough" to provide useful practical reference (see Galison 2003), an attitude subsequently adopted by the BIPM.

6. *Stability*

Metrology is the science of measurement and standardization, carried out by metrologists, who are experts in highly reliable measurement. Despite the fact that it is an independent discipline with its own journals and controversies, the methods and tools of metrology have received little attention from philosophers (Tal 2011: 1083). A central philosophical question in metrology is how the process of standardization works. What exactly is it that metrologists are doing to develop and maintain accurate standards of measurement? How are these methods justified from an epistemic perspective and how do they resolve the apparent circularity of theoretical quantification and empirical measurement?

Chang (2001, 2004, 2007) and van Fraassen (2008, 2009) argue that the apparent circularity is not vicious. According to their view, constructing a quantity concept and standardizing its measurement are co-dependent, iterative tasks. With each iteration the quantity concept is re-coordinated to a more stable set of standards, which allows theoretical predictions to be tested more precisely, facilitating the subsequent development of standards, and so forth (Tal 2013: 1162). This corresponds with the BIPM's view of their own standards, which are not intended as absolute but rather based on a *'mise en pratique'*, that is, a set of instructions allowing the unit to be realized in practice with the highest level of accuracy. The difference between this view and the traditional philosophical approach is that it does not seek to resolve circularity through absolutism. Rather, it treats the standard as a working realization of an external criterion known in metrology as ‘stability’.

Stability refers to the tendency of an apparatus to produce the ‘same’ measurement outcome over repeated runs, as well as replicating the outcomes of similar instruments around the globe. What this means in practice is that discerning any predictable fluctuation in a standard should be as hard as possible; the standard should be as uncorrelated as possible with any changes in the environment. This is the external criterion that measurement standards are designed to meet. Under the guidance of the BIPM, a worldwide network of metrological institutions is responsible for comparing, adjusting, maintaining, disseminating and refining stable standards (Tal 2016: 297).

One of the notable successes of these institutions is the standard measure of time used in almost every scientific context, known as Coordinated Universal Time (UTC) (Tal 2016: 297). UTC is regarded as overwhelmingly stable insofar as a variety of standardization labs around the world manage to closely reproduce it on an ongoing basis. Standardization can be regarded as a process for ensuring independent agreement: Despite being displaced in space and time, and having no causal interaction with each other, the labs can produce results which agree with each other. In other words, they are able to make highly accurate predictions about the measurements that other labs will report each day. Metrologists labour relentlessly to identify standards that support greater predictive accuracy. If standards were chosen arbitrarily, as Pollock (2004) maintains, the world would have no need for metrologists.

As regards the prototype metre bar, its value comes, not from those features which have been selected arbitrarily, but from those which have been carefully calibrated to maximise stability. The standard reflects the realisation of centuries of accumulated theoretical and technological efforts, involving the identification of materials that best support predictive accuracy under varying conditions. Contrary to Pollock’s understanding, the metre bar’s utility is not related to its ability to *designate* a unit of length. After all, anyone could just as well hold their two fingers in the air, refer to the distance between them, and say “this is the length of a metre”. Designating an arbitrary distance is easy, but to be rendered useful, the size of that unit must be preserved by some stable standard. The utility that the prototype metre bar provides lies in its capacity to maintain and replicate that designated distance.

7. *Measuring the Standard*

In sum, measurement and stabilization are one and the same concept. To measure a property such as length is to stabilize it relative to a standard which can reliably preserve that property, thus enabling accurate predictions to be made. Stability is the backbone of measurement utility, and working standards merely approach that ideal without ever realizing it completely.

We return now to the original question of whether the prototype metre bar has measurable length. We have argued that what makes measurement standards valuable is their capacity to enable reliable judgements about length, and hence support accurate predictions. If a standard has been designed to meet an external goal-based criterion, this opens up the possibility of improving the standard and replacing it with a more stable version, thus allowing the original system to be measured. Because stability relates to external events and relationships, no standard can ever represent the final word on stability. As soon as we identify measurement value as being related to stability, we recognize that working standards provide a useful, yet incomplete representation of the concept of measurement.

A continuing trend in metrology is to eliminate as many as possible of the artefact standards, and instead define practical units of measurement in terms of fundamental physical constants. As of writing, the only remaining artefact standard is the International Prototype Kilo (IPK), shortly to be replaced, like that of the other BIPM base units, by a definition entirely based on fundamental constants.

For example, the metre bar prototype was officially superseded in 1960, at the 11^{th} CGPM, when a new definition was agreed based purely on a universally replicable *mise en pratique*. Specifically, the metre was redefined as equal to 1,650,763.73 wavelengths in a vacuum of the radiation corresponding to the transition between the levels $2p^{10}$ and $5d^{5}$ of the krypton-86 atom. This new definition democratised and diversified the materialization of standards, by allowing anyone with the appropriate lab equipment to realize the metre for themselves. Increasing levels of scientific sophistication and greater levels of shared practical knowledge between metrologists have obviated the need for a remaining link to a localised artefact.

It should be noted that the shift from artefact to decentralized standards has not changed the practicalities of metrology any more than the de jure abandonment of the gold standard in 1976 changed the nature of international economics. In practice, the Parisian artefact standards were rarely consulted. The *only* comparison of national standards with the international prototype was carried out over a 15 year period between 1921 and 1936, revealing a variability of around 0.2 μm (Nelson 1981). Like the gold standard, the role of artefact standards was chiefly to shore up confidence in the system as a whole. As scientific knowledge became more widespread, sophisticated and interconnected, this role was no longer necessary.

The definitions of decentralized standards, just like artefact standards, involve an arbitrary component which is needed to establish a convenient unit quantity. For example, in the krypton-86 standard, the value “1,650,763.73” was selected so as to ensure historical continuity with the preceding definition. The number is arbitrary, insofar as any other number would work just as well. However, as previously argued, the number by itself does not provide utility. Instead, it's the stability of krypton radiation wavelengths that supports reliable judgements about length.

The issue of standards being vulnerable to measurement applies just as equally to decentralized standards as it does to localized artefact standards. In order to maintain their status, measurement standards are obligated to deliver reliable judgements, and to support accurate predictions. When competitors can gain an advantage using an alternative system, a working standard immediately loses its status.

The new BIPM base units, which tie base units to fundamental constants, state ideal conditions that cannot be realized by a material object or process, only by an abstract entity, these conditions can be approached more and more closely in practice, yet never perfectly realized (e.g. achieving a perfect vacuum to measure the speed of light). Accordingly, the realization of standards is left entirely open and prone to change when metrologists discover new physical principles that make it possible to materialize the unit with greater stability than before.

Just as with artefact standards, the incomplete understanding of stability leaves decentralized standards perennially vulnerable to refinement.

8. *Conclusion*

Pollock's (2004) argument begins promisingly, with the observation that the utility of measurement stems from its capacity to support comparisons, and not from providing absolute, theory-free descriptions. However, he makes a critical error by falling back into the trap of absolutism, assuming that the concept of length is 'objectively' known, as opposed to something whose practical realization we must work relentlessly towards.

The prototype metre in Paris was selected by metrologists as a useful working standard because it did a good job. It was never intended as the absolute, inviolable definition of the metre. Pollock's (2004) arguments regarding the irreproachable role of the metre standard are directly undermined by the BIPM's 1960 declaration from the 11th CGPM, according to which "the international prototype does not define the metre with an accuracy adequate for the present needs of metrology" (Tal 2011: 1082–1083). If the metre bar was really the foundation of measurement, how could its accuracy ever be found lacking?

Pollock (2004) overlooks the crucial idea that measurement is a goal-directed activity based on clear external objectives, and thus open to continuing refinement. When a system asserts its own supremacy, it severs any ties to delivering in practice, and the system ceases to have utility. For instance, any measurement standard which is beyond reproach, such as the dictator's beard, cannot measure at all, because it is freed of any responsibility to provide practical results in the real world. To be useful, a measurement standard must hold the potential to be found lacking—to be measurable—by some alternative system which delivers superior results in practice. Thus, while current measurement standards do a great job, they do not completely define what we demand of measurement.

For example, in 1988, the International Prototype Kilogram (IPK), which continues to serve as the standard for mass, was removed from its vault in Paris. It was found that the mass of the prototype had drifted downwards relative to the set of national copies distributed globally in 1884, at a rate of change of about 0.5 parts per billion per year (Crease, 2011). By definition, the prototype has no measured value, and hence no measured error. From this frame of reference, the copies around the world are gaining mass. However, because that is a clearly counterproductive interpretation, the BIPM 'inferred' that the prototype must be unstable and somehow losing mass, thus making an implicit comparison of the mass of the IPK to some more stable reference frame.

In conclusion, Wittgenstein's original claim regarding the measurability of the prototype metre must be mistaken. The prototype holds its status as a standard, not because it has been arbitrarily singled out as having a special role in some language game, but because it delivers results in practice which are hard to beat. Measurement standards should thus be interpreted as well-established recommendations for how to achieve the best possible measurement results given the current state of technology. As soon as we succumb to the assumption that standards somehow encapsulate the foundations of measurement itself, and are thus immune to reproach, we cease to be engaged in measurement.

References

BIPM 1928. Comptes Rendus de la 7e CGPM (1927).

Chang, H. 2001. "Spirit, air, and quicksilver: The search for the 'real' scale of temperature." *Historical Studies in the Physical and Biological Sciences* 31 (2): 249–284.

Chang, H. 2004. *Inventing Temperature: Measurement and Scientific Progress*. Oxford: Oxford University Press.

Chang, H. 2007. "Scientific progress: Beyond foundationalism and coherentism." *Royal Institute of Philosophy Supplement* 61: 1–20.

Crease, R. P. 2011. *World in the balance: the historic quest for an absolute system of measurement*. New York: W. W. Norton & Company.

Danjon, A. 1929. "Le temps, sa definition pratique, sa mesure." *L'astronomie* XLIII: 13–22.

Galison, P. 2003. *Einstein's Clocks, Poincaré's Maps: Empires of Tim*e. New York: W. W. Norton & Company.

Nelson, R. A. 1981. "Foundations of the interational system of units (SI)." *The Physics Teacher* 596–613.

Pollock, W. J. 2004. "Wittgenstein on the standard metre." *Philosophical Investigations* 27 (2): 148–157.

Salmon, N. U. 1986. *Frege's puzzle.* Atascadero: Ridgeview Publishing Company

Tal, E. 2011. "How accurate is the standard second?" *Philosophy of Science* 78 (5): 1082–1096.

Tal, E. 2012. *The Epistemology of Measurement: A Model-Based Account*. PhD Thesis, University of Toronto.

Tal, E. 2013. "Old and new problems in philosophy of measurement." *Philosophy Compass* 8 (12): 1159–1173.

Tal, E. 2016. "Making time: A study in the epistemology of measurement." *British Journal for the Philosophy of Science* 67 (1): 297–335

van Fraassen, B. 2008. *Scientific Representation: Paradoxes of Perspective*. Oxford: Oxford University Press.

van Fraassen, B. 2009. "The perils of Perrin, in the hands of philosophers." *Philosophical Studies* 143: 5–24.

Wikipedia "History of the metric system" (2018, February 9). In *Wikipedia, The Free Encyclopedia*. Retrieved 17:49, February 9, 2018, from https://en.wikipedia.org/w/index.php?title=History_of_the_metric_system&oldid=823583734

Wikipedia "Kilogram". (2018, February 9). In *Wikipedia, The Free Encyclopedia*. Retrieved 17:49, February 9, 2018, from https://en.wikipedia.org/w/index.php?title=Kilogram&oldid=824750032

Wittgenstein, L. 1953. *Philosophical Investigations*. G. E. M. Anscombe and R. Rhees (eds.). G. E. M. Anscombe (trans.). Oxford: Blackwell.