Scientific Cumulativity and Conceptual Change: The Case of "Temperature"

Travis Norsen
Marlboro College
Marlboro, VT 05344
norsen@marlboro.edu
(Dated: August 10, 2010)

I examine the historical development of the concept "temperature" from the point of view of questions about the stability of concepts during episodes of theory change. It is argued that the concept retains its identity and meaning through two quite radical developments in surrounding theory, even while these developments uncover novel fundamental characteristics of "temperature" and allow new associated definitions for the concept. I then indicate some of the differing underlying philosophical views which have caused others to view this kind of case very differently, and finally suggest a number of features that I think a theory of concepts would need to possess in order to account for the important aspects of the presented case-study.

I. INTRODUCTION

Starting in the early 1960s, Thomas Kuhn and Paul Feyerabend challenged the traditional cumulativist view of science, and instead emphasized what Kuhn famously dubbed "paradigm shifts" – "those non-cumulative developmental episodes in which an older paradigm is replaced in whole or in part by an *incompatible* new one." [1, p. 92, emphasis added] Feyerabend in particular stressed the changes associated specifically with *concepts* during such episodes. He attacked the idea that concepts retain their meaning when changes in surrounding theory lead them, for example, to be redefined: "the postulate of meaning invariance" as he called this view "is incompatible with actual scientific practice." [2, p. 81]

As pointed out by Dudley Shapere [3] and many others since, Feyerabend's argument rests on a very particular and indeed somewhat peculiar view of the *meaning* of concepts. Those who endorse different theories of concepts have instead seen "meaning invariance" as not only consistent with, but very much ubiquitous in, scientific practice. Hilary Putnam, for example, has argued for the stability of meaning of (among other concepts) "gold" across episodes of theoretical development [4]; John Norton and Jonathan Bain, and also separately Theo Arabatzis, have seen referential stability for the concept "electron" even as physicists' understanding of electrons has grown significantly in complexity [5, 6]; and James Lennox has described the interesting case of the reclassification of "barnacle" (as a crustacean rather than a mollusk). [7]

Here I will present a related case-study of the concept "temperature" – its formation and subsequent development through not one but two major advances in surrounding physical theory. One goal is simply to lay out an additional example of an important scientific concept for which the Kuhn/Feyerabend perspective is *prima facie* implausible, and thus to marginally strengthen the case for the traditional – and, I think, proper – cumulativist view of scientific development.

A deeper goal, though, is to shed light on certain issues pertaining to the theory of concepts: how they are formed, what they mean, and what the relationship is between a concept and its definition. As we will see, there are several unique features of the development of the "temperature" concept which allow some discussion in these directions. It is also hoped that this case-study and the subsequent philosophical discussion might help establish the centrality of these issues – pertaining specifically to *concepts* and their role in science – to the broader questions, such as science's cumulativity, that are of interest to philosophers of science.

Like Lennox's earlier paper on the concept "barnacle", I will be motivated in part by the theory of concepts put forward by philosopher Ayn Rand. [8] Since I won't explicitly use – or presuppose any familiarity with – her account of concepts, there is no need here to discuss her views in any detail. (The interested reader is referred to References [7–9].) I will simply here indicate some general themes of her approach by way of providing an overall orientation to the case-study to follow.

First, for Rand, it is important that concepts have a hierarchical structure: while some (e.g., "table", "bird", "tree") can be formed directly and exclusively on the basis of perception, others are higher-level in the sense that their formation involves – and so requires – previously-formed concepts and/or conceptual knowledge. As we will see, "temperature" is on this account most certainly a higher-level concept, and the non-trivial character of the knowledge required to *form* the concept, will play an important role in the subsequent discussion.

Second, it is important on Rand's view that concepts are open-ended in two senses: (1) they are intended to subsume not just the specific objects on the basis of which they were formed, but an open-ended class of essentially similar objects, and (2) a concept is understood to mean its referents, including all of their characteristics both known and to-be-discovered. As we will see, the development of the concept "temperature" will exhibit both senses of open-endedness.

And third, it is important for Rand's account of concepts that essences are epistemological (not metaphysical) – and, relatedly, that definitions (which ought to state the essence of the concept being defined) do not stipulate or exhaust a concept's meaning, but instead provide a kind of *executive summary* whose propriety is *relative* to some particular context of knowledge. As we will see, this – together with the previous point – provides a very natural framework for understanding how "meaning invariance" can be consistent with the need to revise a concept's definition in light of new scientific discoveries.

Even the explicitly philosophical discussion after the case-study, however, will not focus on Rand's views *per se*. Instead, the purpose will be to understand more deeply why people like Kuhn and Feyerabend see a failure of "meaning invariance" and, more generally, a continual sequence of revolutionary paradigm shifts – where cumulativists (like me) instead see the systematic *expansion* of genuine knowledge. And on the other side, I will point to a certain feature of the "temperature" case-study to bring out what I think is a shortcoming in (especially) Putnam's earlier defense of meaning invariance. More generally, this discussion will lead up to a listing of several features that I think a theory of concepts needs to possess in order to account for the various relevant aspects of the "temperature" case-study.

II. PRELIMINARY REMARKS ON THE CONCEPT "TEMPERATURE"

In Section III, I will recapitulate the historical development of the "temperature" concept, with an emphasis on the original formation of the concept, the major theoretical advances in understanding the meaning of "temperature" and (relatedly) the several ways that the concept is redefined as the surrounding body of theoretical knowledge expands.

Before launching into the history, however, it will be useful to make some preliminary remarks about what, precisely, I take the concept of "temperature" to *be*, and also about what will be systematically *de*-emphasized in my account.

A. Identity of "Temperature"

First, what precisely is the concept of "temperature"? Commentators very frequently begin their explication of this concept by discussing the related concepts of "hot" and "cold" and asserting that "temperature" names the scale or axis along which objects can be ordered by the degree to which they are hot or cold. On page 2 of his treatise on the *Theory of Heat*, for example, James Clerk Maxwell asserts that temperature "is a quantity which indicates how hot or how cold the body is." [10] The author of a classic history of the thermometer similarly declares that temperature "is fundamentally a quantification of the sensations 'hot' and 'cold'..." [11, p. 48] And no doubt on similar grounds, the author of an excellent contemporary thermometry text asserts (on page 1) that: "The concept of temperature must have originated during the earliest stages of human development." [12]

It is certainly true that the concepts "hot" and "cold" must have originated during the earliest stages of human development (just as they are among the first attribute concepts formed by every toddler). For example, they played a famously central role already in the physical theories of the first, pre-Socratic philosophers – not to mention Aristotle.

And while it is not exactly *wrong* to say that "temperature" indicates how hot or how cold a body feels, it is at least rather imprecise: how hot or how cold an object *feels*, though in some sense *primarily* a function of the object's temperature, turns out to depend also on several other factors including the object's texture, size, thermal conductivity, and also on physical properties of the perceiving subject such as the precise degree of warmth of his own body.

As a textbook example of the alleged untrustworthiness of the senses, the phenomenon of a single sample of lukewarm water feeling simultaneously cold and hot (to hands previously immersed in hot and cold water, respectively) is perhaps second only to the straight stick in water that allegedly looks bent. But, really, the example establishes not that the senses deceive us, but simply that the "degree of perceived warmth" of a body is *not the same thing* as its "temperature".[1]

Other examples which support the need for this distinction include the fact that, if you put your hand outside the window of a moving car on a hot day, it will feel cooler (even though, let us say, the temperature of the air inside and out is the same) – and the fact that, say, a block of metal and a block of wood (both left outside on a winter night) won't feel equally cold to the touch in the morning, even though their temperatures are the same.

The upshot is that the concept "temperature" – as dis-

---

[1] For an example of the consequences of failing to make this distinction – i.e., of regarding "temperature" as *simply* a quantification of "degree of perceivable warmth" see Ref. [13], pp. 42-43, where Chang raises the apparently-tricky issue of how it can be that our trust in the reliability of thermoscopes rests on their agreement with sensations of hot and cold, even though "we also allow instruments to augment and even *correct* sensation." (Emphasis added.) Note in particular that, because he doesn't clearly distinguish "degree of perceivable warmth" from "temperature" (as I am suggesting here that one should), Chang sees the need for *correction* of something that was mistaken/invalid. But on my view, it was never that the sensations of hot and cold were mistaken and in need of correction. It is rather that what the thermoscope measures – namely "temperature" – is just one of several factors which jointly contribute to the relevant sensations. The connection to the kinds of issues raised in the Introduction, involving the cumulativity of science, should be obvious.

tinguished from "degree of perceived warmth" – is a relatively sophisticated and technical concept, naming *one* of the several factors on which degrees of perceived warmth depend.

It is, I think, helpful here to analogize "temperature" to the concept which plays a very similar role for visual/color attributes: "hue". Concepts like "red" and "blue" are constructible from direct sense perception just as are "hot" and "cold". And it does not take too much additional sophistication to form a genus concept like "color". But "hue" – though closely related to the concept of "color" and the associated idea of a "color spectrum", is actually quite sophisticated: it denotes one of the several respects in which colors can vary from one another (the others being "saturation" and "brightness"), and so presupposes some method of isolating this one respect.

The quite interesting element of disanalogy here is that, even at the purely perceptual level, colors are experienced as varying along several linearly independent axes. For example, whereas royal blue seems *bluer* than turquoise, robin's egg blue is (say) just as *blue* as royal blue – meaning they have the same *hue* – only it is *whiter* or *softer*. But one has no trouble perceptually *discriminating* turquoise from robin's egg blue: they just don't look the same. By contrast, two bodies which differ from some standard body along two of the distinct dimensions which contribute to "degree of perceived warmth" – say, one is significantly colder but with an equal thermal conductivity, and the other is only marginally colder but with a much greater thermal conductivity – will perhaps *feel* equally cool. Thus, while "temperature" is like "hue" in the sense of naming one very particular respect in which perceptual experiences of the relevant sort can vary (such that both concepts are, so to speak, quite distant from the purely perceptual level), "temperature" is even more sophisticated than "hue" because its formation will require not just a fine distinction, but in addition some kind of experimental apparatus to render the distinction detectable at all.

It should thus be clear that the isolation of this one particular causal factor (from the others which contribute) is going to require the existence and use of a *thermometer* (or, for our purposes here equivalently, a thermoscope – this being the standard term for a thermometer without a series of marks indicating numbered degrees). Such a device will allow one, for example, to discover that two objects can register the same thermal state on the thermometer, even though one feels warmer than the other – it becoming thus apparent that the difference in "degree of perceived warmth" is due to some other difference in the objects. For example, perhaps the two objects are made of very different materials (e.g., one is wood and one is metal), perhaps the two objects are made of the same materials but have very different surface textures or thicknesses, or perhaps (if fluids) the two objects differ in the velocity with which they move past one's hand.

As I will discuss further in the next section, the first thermometers were created in roughly the first decades of the 1600s. During the subsequent century or so, we see people using the newly-invented devices to make precisely these sorts of observations. For example, in 1615, Galileo's friend Sagredo used an early thermometer to demonstrate that water from wells is actually warmer in the summer than in the winter – even "though our senses judge differently" (or so he claims). [11, p. 7] A similar discovery was that "the air of caves and cellars, is not warmer in winter, as our senses would lead us to believe." [11, p. 37] And again, in 1710, we have the first known report of an experiment showing "that blowing air against the bulb of a thermometer with a bellows that had been left with the thermometer for some hours did not change the reading." [11, p. 57]

The claim here is that it is precisely through such exploratory experimentation that the idea of a purely thermal property of objects, distinct from their material nature, size, surface texture, and flow velocity – and also distinct from the "degree of perceivable warmth" to which all of these properties jointly contribute – "proves itself." [14] Thus, the concept of "temperature" was not at all "originated during the earliest stages of human development." Rather, it originated only in the period following the invention of the first crude thermometers in the early 1600s, as these devices (and so the physical attribute they measured) proved themselves useful in practical sciences such as medicine and meteorology.

Of course, this raises a kind of chicken-and-egg problem: what is a "thermometer" if not a device used to measure temperature? And so, wouldn't one have to already possess the concept "temperature" in order to build a thermometer (or recognize some device *as* a thermometer) – in order to then isolate and conceptualize "temperature" in the way that I've sketched here? Looking at the actual history points the way out of this dilemma. The first, crude thermometers were thought of as devices for registering and quantifying (not "temperature", but rather) "degree of perceivable warmth". And this caused no trouble so long as the devices were only used in cases where the *other* relevant causal factors (beyond "temperature") contributing to "warmth" were held fixed: for example, to compare the degrees of warmth of two glasses of water, or of the breaths of two medical patients. The claim is, then, that it is precisely in and by confronting the *puzzles* which arose in the subsequent decades – when these devices began to be used to make more sophisticated comparisons in which their readings seemed to conflict with sensations of warmth – that the concept "temperature" was created – at which point the devices were recognized to register temperature *as opposed to* "degree of perceivable warmth".

It is perhaps also worth noting here that the first known usage of the *word* "temperature" – in the sense relevant here[2] – occurs in 1626 [11, p. 20], and also that another crucial differentiation which is probably required for the genuine formation of the concept "temperature" – namely, the distinction between "temperature" and "heat" – was also being worked out during this same time period, with Francis Bacon in 1620 marking roughly the beginning, and Joseph Black in the middle 1700s marking the conclusion, of this line of development. (For further discussion see, for example, Ref. [15].)

Finally, having begun this discussion by suggesting that other commentators typically *identify* (what I am calling) "temperature" and "degree of perceived warmth," I should clarify that in fact my insistence on distinguishing these is not so unusual. That is, other commentators typically only *begin* by noting the relation of "temperature" to perceptible degrees of warmth, with caveats and clarifications similar to my own following in due course. Maxwell, for example, very quickly notes:

> "We might suppose that a person who has carefully cultivated his senses would be able by simply touching an object to assign its place in a scale of temperatures, but it is found by experiment that the estimate formed of temperature by the touch depends upon a great variety of circumstances, some of these relating to the texture or consistency of the object, and some to the temperature of the hand or the state of health of the person who makes the estimate." [10, p. 2]

He then clarifies that by the use "of the word temperature we fix in our minds the conviction that it is possible, not only to *feel*, but to *measure*, how hot a body is." [10, p. 2, emphasis added]

Given that our goal here is not to present the content of thermodynamics or thermometry – but rather to carefully trace the historical development of the *concept* of "temperature" – it was important to stress not the continuity of "temperature" with more basic concepts such as "hot" and "cold" (though it is clearly in some senses *built* on them), but instead to stress its *distinctness* from these earlier concepts.

[2] An earlier sense of the word was related to the current English word "temperate" and indicated a balance or mixing of opposite elements. In the mid- to late-1500s, one sees this word being used increasingly to denote specifically the balance or mixing of *hot* and *cold*, in the Aristotelian sense of opposite elementary qualities. It was therefore natural to use this word for the novel concept that, I'm arguing, came into existence shortly thereafter. Note also that one sometimes sees the word "temperature" in English translations of Greek texts, such as some of Aristotle's biological writings. This, however, seems always to be a condensation of what was, in the original Greek, an extended phrase (such as translating, as "change in temperature", something literally rendered as "change with respect to hot and cold") or a substitution for the ambiguous Greek meaning "heat" or "the hot". Thanks to Allan Gotthelf, p.c., for help with the translations.

B. Temperature: concept vs. measurement scale

As a second piece of preliminary ground-clearing, it will help also to say what will *not* be emphasized in the historical sketch to follow. I restrict my attention almost exclusively to the *concept* of "temperature" – *not* to the various developing (and competing) experimental *methods* and *scales* for its *measurement*.

That is to say, I intend to (almost) completely gloss over the controversies about such things as: which thermometric fluid (water or mercury or alcohol or air or some other) yields the best-functioning thermometers; whether air- (or some other gas-) based thermometers are generally preferable to those using liquids; which (and how many) fixed points should be used for calibration and how their fixity can be established; how to establish an "absolute" temperature scale that doesn't rest on the arbitrary selection of some particular thermometric substance; and so on.

It's not that these issues are uninteresting or unimportant – just that they are irrelevant to my project. Of course, saying this reveals some things about my philosophic views. For example, I'm not an operationalist, so I don't regard the procedures for *measuring* temperature as constitutive of the meaning of the concept. Indeed, just the opposite: I think it's clear that the quest to systematically improve the procedures for measuring temperatures (by making the measurements more precise, by finding ways to extend the range of temperatures which can be meaningfully measured, by finding less arbitrary temperature scales, etc.) *presupposes* the possession of an already-meaningful, robust concept of "temperature". Those interested in a more operationalist approach in which this quest takes center stage, are referred to Hasok Chang's recent book *Inventing Temperature.* [13] (References [12] and [16] provide excellent, though non-philosophical, overviews of contemporary thermometry from the point of view of the practical physicist or engineer, and Ref. [17] gives a very accessible historical account of these issues through the middle of the 19th century.)

What kind of advance, then, *do* I regard as pertaining to the actual *concept* of "temperature" – and so being relevant to the project here? That will be the story of the next section. But for now I'll just indicate that I recognize two fundamental advances beyond the "thermometric stage" in which the concept first arises. The first is the adoption of the kinetic theory of heat (and the related kinetic theory of gases). This permitted a new definition of "temperature" purely in terms of independently-established intrinsic physical properties. The second is Boltzmann's discovery of the microscopic meaning of "entropy"; this permitted a further re-definition of "temperature" which had the virtue of explaining the previous one(s) and also permitting the application of the concept to new domains (where the kinetic theory notion was simply inapplicable).

III. HISTORICAL DEVELOPMENT OF THE CONCEPT "TEMPERATURE"

Let me then begin my sketch of the historical origins and subsequent development of this concept.

A. The Thermometric Stage

The Greek physician Galen applied the traditional (Empedoclean/Aristotelian) notion of four elements to people, classifying them according to "temperament", i.e., the degrees of such opposing qualities as "hot" and "cold" which allegedly constituted their natures. [18] Galen may have been the first to design a numerical scale along which degrees of hot and cold could be placed – but, not possessing a thermometer, his measurements were restricted to unaided sense perception, and hence (I would say) constituted rudimentary measurements not of "temperature" but only of "warmth."

Devices which do indeed register temperature were also known in Ancient Greece, though (it seems) these were considered mere novelties (being conceived, for example, simply as demonstrations that the sun's heat could be made to move water) and were never actually *used* to measure temperature. The first century (AD) *Pneumatica* by Heron of Alexandria, for example, discusses a number of devices in which water is moved by the application of heat or cold to a bit of trapped air. This work was translated into Latin in the mid-late 16th century, and the construction of the first actual thermometers was evidently motivated by the recognition that the devices in Hero's text could be used for this purpose. [11, 18]

The question of who should be credited with the invention of the thermometer is a question of some controversy. The first published account is in a 1611 work of the Italian physiologist Santorio Santorii:

> "I must inform you of a marvellous method, by which, with the aid of a glass instrument, I am wont to measure the cold or hot temperature of the air, of all districts, all places, and all parts of the body, and so exactly that at any time of the day I can measure with the compass the degrees and ultimate stations of heat and cold; and it is in our house at Padua and we show it freely to all; we promise shortly to bring out a book on medical instruments in which we shall give the figure, construction and uses of this very ancient instrument." [18]

Santorio later explained that the device consisted of an air-filled glass bulb with a long neck, and says that it was adapted from a device "put forward by Heron for another purpose." [18] (It is possible, however, that Santorio invented the device independently, and only credited Heron later when it was pointed out to him that a similar device had been discussed in the *Pneumatica*. [11])

There exists also a letter written in 1638 by a friend of Galileo, which attributes the invention to him:

> "I remember an experiment shown to me, about thirty-five years ago, by our Signor Galileo. He took a glass flask about the size of a small hen's egg, with a neck about two palmi in length and as narrow as a wheat-straw, and having well warmed the said flask in his hands, he then turned the opening upside down into a vessel placed underneath, which contained a little water. When he took away the heat of his hands from the flask, the water at once began to rise in the neck and mounted higher, by more than a palmo, than the level of the water in the vessel. Sig. Galileo made use of this effect to construct an instrument for examining the degrees of heat and cold." [18]

Galileo seems to have claimed the invention of the thermometer for himself in letters from about 1612, and there is some evidence that he may have presented a demonstration thermoscope to students as early as the 1590s, but no account appears in any of his surviving writings of this period.

There are several others from this same period for whom some case for the invention of the thermometer can be made; most probably more than one instance of geniunely independent creation occured. [18] For our purposes, though, these issues are irrelevant; what matters is just that, in the first decades of the 17th century, it became widely recognized that one could use the thermal expansion and contraction of air to measure variations in perceptible hot and cold along the particular axis that came eventually to be conceptualized as "temperature".

In the subsequent decades and centuries, many significant advances in thermometry occured. These include the increasing use of closed devices using the thermal expansion of a liquid (such as water, alcohol, or mercury) instead of air to register the temperature, and also an increasing thoughtfulness and sophistication in regard to the question of numerical scale and its calibration. Such advances are nicely reviewed from a historical point of view in References [11] and [17], from the pratical point of view of the contemporary physicist or engineer in Ref. [12] and [16], and from a philosophical point of view in Ref. [13].

From the point of view of this paper, however, such advances pertain essentially to issues of standardization, reproducibility, and communicability, and do not represent any fundamental change in the way that "temperature" is *conceived*.

One contemporary physics text, after briefly sketching the state of thermal knowledge up to the mid-19th century, notes that: "So far, all we have really done in defining temperature is to make the somewhat circular statement that 'temperature is what one measures with a thermometer'." [19, p. 413] This, of course, is a bit tongue-in-cheek. A proper definition would provide a sensible genus such as "physical property" and would need (in order to avoid the noted circularity) to flesh out what was already known about how a thermometer functions: first, certain materials expand/contract as they are heated/cooled, such that their volumes can be taken to indicate or register their own temperatures; second, heat will flow spontaneously between objects in thermal contact – always from the hotter (higher temperature) body to the cooler (lower temperature) body – until a state of thermal equilibrium (equality of temperatures) is achieved; and so, third, a thermometer (placed in thermal contact with another body for a long time) can be understood to register also the temperature of that body.

The contemporary Oxford English Dictionary provides a more serious "thermometric stage" definition of "temperature" which at least indicates all of these ideas: "The state of a substance or body with regard to sensible warmth or coldness, referred to some standard of comparison; *spec.* that quality or condition of a body which in degree varies directly with the amount of heat contained in the body, and inversely with its heat-capacity; commonly manifested by its imparting heat to, or receiving it from, contiguous bodies, and usually measured by means of a thermometer or similar instrument." [20]

B. The Kinetic Stage

Although the assignment of a specific temperature value to an object requires a physical interaction between it and a thermometer, temperature is – from the very beginning, and quite properly – conceived as a property *of the object* and not merely of the object-device interaction.

That is, the thermometer (under appropriate conditions) is understood as registering the pre-existing temperature of the object. Thus, the question immediately arises: what sort of intrinsic physical property (or properties) of the object does "temperature" actually refer to? That is, what, precisely, is the physical difference between two bodies which differ only in temperature?

One possible answer is as obviously correct as it is useless: temperature denotes the "intensity" or "density" of *heat* contained in the body. Such an answer simply moves the question back without really answering it. We want then to know precisely how a body changes physically when (say) some quantity of heat flows into it. The point is just that the well-known theoretical controversies during especially the 18th and 19th centuries about the physical nature of *heat* can be equivalently thought of as controversies about the underlying, specifically microphysical meaning of temperature.

The widespread view during the first part of this period was the fluid or "caloric" theory of heat, according to which heat was identified with a subtle, elastic fluid (the "caloric", typically thought to consist of particularly fine particles) – temperature then being a measure of what amounts to the density of this fluid (or at least the part of the fluid which was, in a given body at a given time, "sensible" rather than "latent"). It is well-known, of course, that this theory was eventually abandoned, though it is worth stressing that the caloric theory was by no means irrational or ridiculous: it accounted in a very plausible way for many of the empirical principles of heat flow, and could be integrated with some apparent fruitfulness with other ideas such as Dalton's early chemical atomic theory to account for certain qualitative aspects of the behavior of gases.

Despite its plausibilty and reasonableness as a candidate theory, though, the caloric theory was regard with a certain cautious tentativeness by all but its most narrow-minded proponents. An 1820 Dictionary of Chemistry, for example, contains the following entry:

> "CALORIC. The agent to which the phenomena of heat and combustion are ascribed. This is *hypothetically* regarded as a fluid, of inappreciable tenuity, whose particles are endowed with indefinite ido-repulsive powers, and which, by their distribution in various proportions among the particles of ponderable matter, modify cohesive attraction, giving birth to the three general forms of gaseous, liquid, and solid." (emphasis added) [21]

Note that this defines "caloric" as *whatever* lies behind the phenomena of heat, with what is now normally referred to as the "caloric theory of heat" being just one hypothesis about its nature. As Maxwell would later explain, though, the word eventually came to connote so strongly the hypothesis of an imponderable fluid, that it could no longer be used in this neutral, non-committal sense. [10]

The dictionary entry seems to accurately reflect the attitudes of the scientists of the period, who indeed generally considered the controversy between the caloric theory and its (main) rival – the kinetic theory of heat – to be unsettled. For example, Joseph Black explains in his mid-18th century lectures:

> "Heat is plainly something extraneous to matter. It is either something superadded to ordinary matter or some alteration of it from its most spontaneous state. Having arrived at this conclusion, it may perhaps be required of me, in the next place, to express more distinctly this something – to give a full description, or definition, of what I mean by the word *heat* in matter. This, however, is a demand that I cannot satisfy entirely. I shall mention, by and by, the supposition relating to this subject that appears to me the most probable. But our knowledge of heat is not brought to that state of perfection that might enable us to propose with confidence a theory of heat or to assign an immediate cause for it.

> "Some ingenious attempts have been made in this part of our subject, but none of them has been sufficient to explain the whole of it. However, this should not give us much uneasiness. It is not the immediate manner of acting, dependent on the ultimate nature of this peculiar substance, or this particular condition of common matter, that most interests us. We are far removed as yet from that extent of chemical knowledge which makes this a necessary step for further improvement. We have still before us an abundant field of research in the various general facts and laws of action, which constitute the real objects of pure chemical science. And I apprehend that it is only when we have nearly completed this catalogue that we shall have a sufficient number of resembling facts to lead us to a clear knowledge of the behavior peculiar to this substance or modification of matter called heat..." [15, p 150-1]

After reviewing the then-current evidence in favor of both the kinetic and caloric theories of heat, Black concludes: "neither of these suppositions has been fully and accurately considered by their authors, or applied to explain the whole of the facts and phenomena relating to heat. They have not, therefore, supplied us with a proper *theory* or *explication* of the nature of heat." [15, p. 152]

Similarly, Lavoisier and Laplace declare in 1780:

> "We will not decide at all between the two foregoing hypotheses. Several phenomena seem favourable to the [kinetic theory of heat], such as the heat produced by the friction of two solid bodies, for example; but there are others which are explained more simply by the other – perhaps both hold at the same time." (Translated in Ref. [21].)

Sadi Carnot, who in 1824 brilliantly grasped the second law of thermodynamics a quarter century before the first law was formulated, based part of his analysis on the conservation of heat, regarded as a consequence of the caloric theory. But even he recognized the less-than-conclusive character of the evidence for that theory:

> "The fundamental law that we propose to confirm seems to us to require, however, in order to be placed beyond doubt, new verifications. It is based upon the theory of heat as it is understood to-day, and it should be said that this foundation does not appear to be of unquestionable solidity. New experiments alone can decide the question." [22, pg 107] (See also the discussion in Ref. [21].)

The positive evidence supporting the rival kinetic theory of heat was building up during this same period. Already in 1620, Francis Bacon had suggested, on the grounds that new heat comes into existence when an object is rubbed or hammered, that "heat itself, its essence and quiddity, is motion and nothing else." [15, p. 170] The same view was expressed often in the 17th and 18th centuries. Benjamin Thompson (aka Count Rumford) strengthened the evidence significantly with his careful and systematic experiments showing not only that a seemingly unlimited amount of heat can be generated when two objects generate friction by rubbing against one another, but, crucially, that the thermal properties of the objects are not changed (as one might have expected on the basis of the caloric theory). Thompson summarizes, in his paper of 1798:

> "It is hardly necessary to add that anything which any *insulated* body, or system of bodies, can continue to furnish *without limitation*, cannot possibly be *a material substance*; and it appears to me to be extremely difficult, if not quite impossible, to form any distinct idea of anything capable of being excited and communicated in the manner in which heat was excited and communicated in these experiments, except it be MOTION." [15, pg 187-8]

The case for the kinetic theory of heat became conclusive in roughly the middle of the 19th century, after Joule's experiments establishing the value of the "mechanical equivalent of heat" and the associated recognition of what is now called the first law of thermodynamics: energy is conserved when heat (recognized as a microscopic form of energy) is taken into account.

The so-called kinetic theory of gases – repeatedly developed and subsequently ignored or forgotten during this same period – finally became widely accepted at about the same time. As Rudolf Clausius explained it in 1857, according to the kinetic theory of gases the

> "pressure of the gas against a fixed surface is caused by the molecules in great number continually striking against and rebounding from the same. The force which must thence arise is, in the first place, by equal velocity of motion inversely proportional to the volume of the given quantity of gas; and secondly, by equal volume proportional to the *vis viva* [i.e., kinetic energy] of the translatory motion: the other motions do not here immediately come into consideration.
>
> "On the other hand, from Gay-Lussac's law we know that, under constant volume, the pressure of a perfect gas increases in the same ratio as the temperature calculated from $-273°$ C., which we call the absolute temperature. Hence ... it follows that the absolute temperature is proportional to the *vis viva* of the translatory motion of the molecules." [23]

In contemporary notation:

$$\langle\frac{1}{2}mv^2\rangle = \frac{3}{2}k_B T \quad (1)$$

where the left hand side means the translational kinetic energy averaged over the individual molecules, k_B is a universal proportionality constant (eventually to be known as Boltzmann's constant), and T is the (absolute) temperature.

The kinetic model of gases could account not only for the empirical gas laws of Boyle and Charles but also predicted what is now known as Avogadro's hypothesis – for which, at this same time, there was starting to be strong independent empirical support.

By the 1860s, several additional pieces of the puzzle were falling into place. Clausius had recognized the possibility of non-translational kinetic energy playing a role, and used this to account for the differences in specific heats among different chemical species of gas. And Maxwell gave a particularly beautiful proof that, averaging over all possible collisions, when two gases (initially at different temperatures) are allowed to mix, their temperatures (as defined just above) will tend, statistically, to equalize. [24] And Maxwell would subsequently show that the kinetic theory could account also for the various transport phenomena in gases: heat conduction, diffusion, and viscosity.

One thus has at least an in-principle causal explanation for the most important empirical generalizations that were previously noted as supporting the temperature concept – for example, that heat flows spontaneously from hotter to colder objects, and that objects in thermal contact will eventually achieve the same temperature – and thus in particular a causal explanation of the functioning of thermometers: the statistical thermal equilibration process described by Maxwell results in the average translational kinetic energies of the molecules of the object becoming equal to that of the molecules in the thermometric fluid of the thermometer, and the macroscopic volume of that fluid in turn registers the molecular kinetic energies because faster molecules must spread out somewhat in space in order to maintain a constant pressure.

In summary, by roughly the 1860s, the kinetic theory of heat and gases allowed for the identification of a new essential characteristic of the temperature concept: an object's temperature (on a certain now-privileged scale) was a measure of the average translational kinetic energy of its constituent molecules.[3] By "essential" here, I just mean to capture the inevitable sense that the kinetic theory seems to have identified what temperature *really is*. Hasok Chang, for example, notes that "Practical thermometry achieved a good deal of reliabililty and precision before people could say with any confidence what it was that thermometers measured." [13, p. 160] The sense, of course, is that "what it was that thermometers measured" was recognized finally with the kinetic theory identification of temperature with average translational kinetic energy. Moreover, this identification brings with it an explanation of the principles underlying the nature and function of thermometers.

And this newly identified essential characteristic was subsequently used in an updated "Definition of Temperature from the Kinetic Standpoint. According to the kinetic theory the temperature of a system corresponds to a certain value of the mean kinetic energy of translation of the molecules". [27]

[3] In Ref. [13], Chang distinguishes the "ordinal (thermoscope-based) temperature concept" from the cardinal scale that (he suggests) is provided by a genuine thermometer (with a well-defined numerical scale and calibration). (See his pages 41 and 87, for example.) The ordinal/cardinal distinction is indeed important, and he's right that a thermoscope provides only an ordinal scale. Unlike Chang, though, I think that in a way the whole point of all the detailed case histories he re-tells in the book, can be captured by the following simple claim: the more sophisticated temperature measurement techniques and scales *still* provide only an *ordinal* scale for temperature. That is, I think, it is only with the establishment of the kinetic theory of heat and in particular the identification of temperature with average translational molecular kinetic energy, that one has, for the first time, a genuinely *cardinal* scale, with respect to which it is genuinely meaningful to assert, for example, that a given temperature is *twice* another. The meaning of such a statement is, of course, just that the molecules in the one case have, on average, *twice* the translational kinetic energy as they do in the other case. Prior to this kinetic theory identification, any such statement unavoidably involves some element of arbitrary convention. This was pointed out in 1848 by William Thomson (Lord Kelvin): "Although we have thus a strict principle for constructing a *definite* system for the estimation of temperature, yet as reference is essentially made to a specific body as the standard thermometric substance ... we can only regard, in strictness, the scale actually adopted as *an arbitrary series of numbered points...*" [25] Later, the point was made even more forcefully by Ernst Mach. [26]. See Section III C for further discussion of Thomson's allegedly less arbitrary, "absolute" (thermodynamic) temperature scale. Note also that one could imagine perfectly parallel controversies playing themselves out in the history of the concept "hue" which I earlier analogized to "temperature". For example, it would be easy to assign a perfectly definite number to each distinct hue – say, the number characterizing the angle of refraction for monochromatic light of that hue. But then: through a prism made of what material, exactly, and possessing which precise shape? One thus sees clearly that what might *appear* as a cardinal scale for measuring hues (because definite numbers, which can be compared quantitatively, are used) is in fact really still just ordinal, the numbers being after all based on an ultimately arbitrary choice of a prism with some particular shape, material composition, etc. To be clear, I have no idea whether there ever was such a debate – my point is only that there might have been, and this sheds some light on what is and isn't unique – and what is and isn't interesting – about the historical controversies recapitulated in Chang's book.

C. The Statistical Stage

As mentioned already, in 1824 Sadi Carnot identified (what would later be called) the second law of thermodynamics. More precisely, Carnot assumed the *universal* validity of the familiar fact that heat flows spontaneously from higher- to lower-temperature objects – but never vice versa – and established, on the basis of this assumption, an amazing fact about heat engines: the *efficiency* of a reversible heat engine operating between two fixed temperatures (e.g., that of the boiler and that of the condenser) was a universal function of those two temperatures, independent of the working fluid used in the engine, or any other particular details of its construction. (Any element of irreversibility in the engines would only reduce the efficiency.)

For a reversible heat engine between a condensor of temperature T and a boiler of temperature $T + dT$, Carnot's result can be put mathematically as follows:

$$\eta = C(T)\, dT \tag{2}$$

where η is the efficiency (defined as the ratio of the useful mechanical work output W to the quantity of heat absorbed from the boiler Q_{in}) and $C(T)$ is the "Carnot function" – the crucial point being that this function is *universal*, i.e., independent of the material or mechanical details of the particular heat engine in question (provided, of course, that it is reversible).

What particular mathematical function of T $C(T)$ *is*, of course, will depend on which temperature scale one adopts. If, for example, one assumes a reversible heat engine using an ideal gas as the working fluid, and provided one adopts the associated kinetic theory definition of temperature outlined in the previous sub-section, one can show that

$$C(T) = \frac{1}{T}. \tag{3}$$

On the other hand, if one adopts (say) the Celsius scale for temperature, one will have a different (though of course mathematically equivalent) functional dependence.

As William Thomson (aka Lord Kelvin) famously pointed out in 1848, the universality of $C(T)$ allowed it to be used to *define* a new (equally universal) temperature scale, i.e., one in which (unlike all other previously-proposed scales) no arbitrary "reference is ... made to a specific body as the standard thermometric substance." [25] After initially proposing the simplest possibility – namely that $C(T)$ should be a constant, independent of temperature – he recognized that *that* scale would be pointlessly divergent from the already accepted scale based on the thermal expansion and contraction of gases (extrapolated to, and then measured from, "absolute zero").

In 1854, then, and with some help from Joule, Kelvin proposed instead adopting Equation (3) as a *stipulation* for the functional form of $C(T)$, entailing a universal thermometric scale that coincided with the "kinetic stage" definition of "temperature" – and agreed very closely with the scale defined by actual gas-based thermometers. (See pages 173-186 of Ref. [13] for a more detailed account of the historical development here.)

Note that the way I am telling the story here is perhaps a little bit confusing, since the kinetic theory of gases (and the associated kinetic theory definition of temperature) wasn't widely adopted until at least about a decade after Kelvin proposed his universal temperature scale(s). So his motivation for proposing to adopt Equation (3) certainly was not that it re-created the definition I gave in the previous sub-section. It was rather that the adoption of Equation (3) allowed the "universal" temperature scale to coincide with what Kelvin (together with Joule) called "temperature from the zero of the air-thermometer." [13, p. 185]

The reason for this bit of anachronism is that, contrary to most other commentators, I don't see anything particularly profound in Kelvin's idea of a universal temperature scale. At best, it represents an argument for the privileged status of some particular temperature *scale*, but doesn't seem to add anything to one's understanding of what temperature *is*. And, in a way, it doesn't even privilege any particular temperature scale, as Kelvin's own oscillations show: one could stipulate any arbitrary functional form for $C(T)$ and thereby bless, as "universal", any arbitrarily preferred temperature scale. At the end of the day, then, Kelvin's proposal seems to me to be little more than a rather empty nod of support for the already-well-established gas-based temperature scale, whose ultimate justification (as somehow truly privileged) would I think only arise with the kinetic theory as explained in the previous sub-section.

All of this does, nevertheless, play an important role in our project here. For with Kelvin's eventual choice for a universal temperature scale – namely that which turns out to coincide with the kinetic theory definition of temperature – it turns out that the ratio of the boiler and condenser temperatures in a reversible heat engine will be equal to the ratio of the quantities of heat absorbed from the boiler and abandoned into the condenser in one cycle. That is:

$$\frac{T_{\text{boiler}}}{T_{\text{condenser}}} = \frac{dQ_{\text{in}}}{dQ_{\text{out}}} \tag{4}$$

where the quantities of heat flowing in and out are related to the useful mechanical work output W according to the first law of thermodynamics (the conservation of energy):

$$dQ_{\text{in}} = W + dQ_{\text{out}}. \tag{5}$$

Equation (4), however, is equivalent to the following:

$$\frac{dQ_{\text{in}}}{T_{\text{boiler}}} = \frac{dQ_{\text{out}}}{T_{\text{condenser}}}. \tag{6}$$

It is as if there is some new substance, a certain quantity of which – namely the left hand side of Equation (6) – is absorbed by the working fluid of the heat engine during the part of the cycle when it absorbs heat (isothermally) from the boiler, the *same quantity* – the right hand side of Equation (6) – then being transferred into the condenser during the other part of the cycle, when the working fluid passes heat into the condenser.

This hypothetical new substance was named *entropy* by Clausius, who recognized that – for the same reason that the caloric theory of heat had been abandoned – it couldn't be a literal substance at all, because it wasn't conserved. Equation (6), after all, applied only in the case of *reversible* heat engines. For an *irreversible* engine (i.e., for any real heat engine), the entropy transferred from the engine to the condensor would be *greater* than the entropy transferred from the boiler into the engine – i.e., the irreversible processes within the engine itself could create entropy *ex nihilo*.

Thus, like temperature itself, surely entropy was some kind of emergent property of the microscopic/molecular state of the gas. But what kind of property? Clausius had no idea, and simply defined the entropy S through the relation (applicable to a reversible isothermal process)

$$dS = \frac{dQ}{T} \tag{7}$$

– recognizing of course also that entropy could increase "spontaneously" (i.e., without any associated heat flow) if there is some irreversibility involved.

It is then possible to reformulate the second law of thermodynamics as follows: the entropy of a closed system can increase or stay the same, but can never decrease. That is,

$$dS \geq 0. \tag{8}$$

This is a re-formulation of the second law (as opposed to some new kind of statement) because it can be shown – rather miraculously – that this statement about the as-yet mysterious concept of entropy is perfectly equivalent to the two previously-mentioned statements of the second law: that heat flows spontaneously from hotter to colder objects (but never vice versa), and that the efficiency of a reversible heat engine operating between two fixed temperatures is a universal function of those temperatures.

What is important for us is that, eventually, the microscopic meaning of entropy was identified. This is summarized in the famous equation of Ludwig Boltzmann (which, actually was first written in this form not by Boltzmann but by Max Planck):

$$S = k_B \log(g) \tag{9}$$

where S is the entropy of a system with certain given macroscopic properties, k_B is (again) Boltzmann's constant, and g is the *number of distinct (suitably coarse-grained) microstates which share the macrosopic properties of the micro-state to which entropy S is being assigned*. This isn't the place to give a detailed explication of Boltzmann's account of entropy. Suffice it here to say that S measures the "specialness" of states: states which are relatively "unique" (in the sense of there being relatively few other states which are relevantly like them) will have low entropy, while states which are relatively "common" (in the sense of there being many other states which are relevantly like them) will have high entropy. Importantly, Boltzmann's account of entropy (unlike an influential later one due to Gibbs) permits an *explanation* of how the time-asymmetric behavior captured by the second law in the form of Equation (8) can emerge from time-symmetric underlying laws. [28]

We can finally put these pieces together to identify a new, even more fundamental way to define "temperature." With heat and now also entropy being understood in terms of the micro-physical properties of systems (and, importantly, being so understood *independently* of "temperature"), one can mathematically rearrange Equation (7) into the following form:

$$T = \frac{dQ}{dS} \tag{10}$$

or, equivalently for a constant volume process,

$$T = \frac{dE}{dS}. \tag{11}$$

We then regard this as a new, purely micro-physical definition of temperature. This formula or its equivalent is typically asserted as the "fundamental definition of temperature" in modern statistical mechanics textbooks. (See, e.g., pages *xiii* and 41 of Ref. [29].)

All of this was summarized nicely by Max Planck in his contribution to a 60th birthday festschrift for Boltzmann:

> "It seems that Clausius and Maxwell have never tried to give a direct and general definition of entropy in mechanical terms. It was left to Boltzmann to accomplish this step, starting from the kinetic theory of gases and defining entropy in a general and universal fashion as the logarithm of the probability of the mechanical state. .... To [this definition] of entropy there correponds a definition of temperature through $dQ = TdS$." [30, p. 151]

One might wonder: why is this new definition, in terms of entropy, any kind of advance over the "kinetic" definition of temperature, which is also "purely microphysical"? The answer is that it is even more fundamental. The statistical definition *explains* the fact reported in the kinetic definition of temperature: for example, for an ideal monatomic gas of N particles, the number of distinct microstates associated with the total energy being E is proportional to $E^{3N/2}$. (For particles of mass m, the energy constraint reads: $\frac{1}{2}m\sum_i^{3N} v_i^2 = E$. But then, in the $3N$ dimensional velocity space, the energy surface is a sphere of radius $R = \sqrt{2E/m}$ whose "surface area" – proportional to the coarse-grained number

g of distinct microstates consistent with the energy constraint – is $A \sim g \sim R^{3N-1} \sim E^{(3N-1)/2} \approx E^{3N/2}$ since, for macroscopic systems, $N \approx 10^{23} \gg 1$.) Thus, plugging into Equation (9), we have that

$$S(E) = k_B \log(E^{3N/2}) + \text{const} \tag{12}$$

such that the statistical definition of Equation (11) reduces to

$$T = \frac{E}{\frac{3}{2}k_B N} \tag{13}$$

which is, for the assumed kind of system, precisely equivalent to Equation (1).

The statistical definition of temperature also allows a very fundamental explanation of the various empirical principles pertaining to temperature, such as that objects in thermal contact will eventually reach an equilibrium in which their temperatures are the same. Very briefly, two systems which can exchange energy will, with overwhelming probability, end up in a state which maximizes the total entropy of the joint system. But it is easy to show that the partitioning of energy which achieves this is precisely the one for which the "marginal entropies" dS/dE of each sub-system (associated with further energy exchanges) are equal. This then entails, by Equation (11), that the temperatures will be equal.

Furthermore, the statistical definition of temperature also expands the domain of physical systems to which the concept of temperature can be meaningfully applied. Most famously – and most importantly for the development of physics in the 20th century – Boltzmann's account of entropy can be extended to electromagnetic fields ("cavity radiation"), as first undertaken by Max Planck. This allows the temperature of the electromagnetic field in a given region to be defined, according to Equation (11). As is well known, this was precisely the route by which Planck first stumbled upon the correct mathematical formula for the spectrum of cavity radiation, which subsequently triggered the creation of quantum theory in the hands of Einstein and others.

IV. DISCUSSION

Let me begin here with an essentialized summary of the case history sketched in the previous section. The concept of "temperature" was first formed in the 17th century, as it was recognized that there existed a physical property of objects, which could be registered by a certain (then) new kind of device (the thermometer), and which was the predominant – but not exclusive – contributor to the sensations of "hot" and "cold". I have insisted that the concept "temperature" is not *merely* a quantification of the degrees of perceptible hot and cold, but involves also the *differentiation* of temperature from related factors, including especially "heat" and subsidiary properties such as thermal mass and thermal conductivity, as well as other properties of objects such as surface texture and size.

This (rather extensive) body of knowledge which is pre-requisite for the concept "temperature" is summarized (though certainly not exhausted) by what I have called the thermometric-stage definition of temperature: temperature is the physical property, primarily but not exclusively responsible for perceptible degrees of heat, which can be measured using a thermometer.

I then sketched two major developments in thermodynamic theory which vastly increased the depth of understanding of temperature, and allowed for new definitions in terms of deeper fundamentals. The first such development was the discovery of the underlying micro-physical basis of temperature (or, here equivalently, heat): the so-called kinetic theory of heat. The associated kinetic-stage definition of temperature reads: temperature is the physical property (of a macroscopic body) which measures the average translational kinetic energy of its microscopic constituents.

The transition from the thermometric to the kinetic stage was lengthy and controversial, involving an extended period when many or most practicing scientists (at least tentatively) endorsed the now-discredited caloric theory of heat. During this controversy, there was, however, never any question about the meaning of "temperature" or its validity as a concept. Even in the midst of the controversy, for example, proponents of the different theories would have agreed about what the temperatures of various concrete objects were and how that is to be determined. What they disagreed about was only the underlying microscopic nature of thermal phenomena. So this debate could rage and, eventually, be settled without the concept "temperature" in any sense hanging in the balance; indeed, the stable existence of the concept "temperature" is in a way what made the controversy possible (in the sense that what the controversy was *about* was the underlying microscopic basis *of temperature*), and also what made it into a *controversy* (in the sense that, had the proponents of the two theories meant radically different or incommensurable things by "temperature", they wouldn't have actually been disagreeing about anything). Thus, in the transition from the thermometric to the kinetic stage, we have an example of profound (and indeed rather messy) "theory change" in which the temperature concept plays a central role and during which it is redefined, but with the concept itself being a central pillar of stability on which, so to speak, the change occurs.

What makes the concept "temperature" so remarkable is that there is also a second such change in the transition from the kinetic to the statistical stage. This involves the sorting-out of the microscopic basis of "entropy" (much as the establishment of the kinetic stage involved the sorting-out of the microscopic basis of "temperature" itself), concluding with Boltzmann's account of entropy and the associated fundamental statistical-mechanical definition of temperature: temperature as the quantity of additional energy needed (at constant volume) to increase a system's entropy by one unit ($T = dE/dS$).

One of the interesting points here is that the statistical view of temperature allowed, for the first time, the application of the concept to a whole new class of entities (for which the definition in terms of molecular kinetic energies would be meaningless).[4] This is not a change in the extension of the "temperature" concept, but rather only an expansion of the class of objects now known to possess the same physical property that one recognized from the very beginning as[5] measurable with a thermometer.

The main point to be stressed about these developments is that both of them (that is, the development from the thermometric to the kinetic stage and then the development from the kinetic to the statistical stage) are strong candidates for episodes in which scientific knowledge genuinely and substantially *accumulates* but without anything like a Kuhnian paradigm shift. In particular, at the end of each transition, the knowledge integrated by the concept *in the earlier stage* – and in particular, the characteristics cited in the definition associated with that earlier stage – are still endorsed as *true*. Thus, after one adopts the kinetic theory of heat, one still recognizes as true the claim that temperature is a physical property (associated with perceptible degrees of warmth) which can be measured by a thermometer. Indeed, as I have pointed out, not only is the claim that temperature can be measured by a thermometer *consistent* with the claim that temperature is a measure of the average microscopic motion of an object's molecules – the latter actually provides a causal *explanation* of the former.

And similarly, after one adopts the statistical view of temperature, one still endorses as true the idea that the temperature of a material system is proportional to the average translational kinetic energy of its molecular constituents. Indeed, armed with Boltzmann's account of the microscopic meaning of entropy, the statistical definition of temperature *explains why* the temperature is, for the case of a system composed of molecules, a measure of their average translational kinetic energy.

The history of the concept "temperature" thus seems very helpful in responding to the charges of those, like Kuhn and Feyerabend, who claim in effect that every profound theoretical development radically changes the meaning of the involved concepts. Here we seem to have a case (or actually two cases) of just such profound developments in which there is apparently no element of "revolution" at all – and certainly nothing like the alleged conceptual incommensurability.

But this now raises some questions. Feyerabend, for example, also examined the history of the concept "temperature" and came to rather different conclusions from mine here. What gives?

Feyerabend compares, in particular, the notion of temperature as defined through Kelvin's absolute temperature scale in the context of classical (phenomenological) thermodynamics, with the notion of temperature associated with the statistical approach to entropy and the second law of thermodynamics. He then points out that the first version of "temperature" (together with the uniqueness of the Carnot function mentioned above) implies the strict, universal validity of the second law of thermodynamics: heat will always and exclusively flow from a hotter body to a colder one. Whereas:

> "The statistical account ... allows for fluctuations of heat back and forth between two levels of temperature and, therefore, again contradicts one of the laws implicit in the ... thermodynamic temperature. The relation between the thermodynamic concept of temperature and what can be defined in the kinetic theory, therefore, can be seen to conform to the pattern that has been described at the beginning of the present section: we are again dealing with two incommensurable concepts." [2, p. 78]

One curious thing about Feyerabend's argument is that the contradictoriness he emphasizes pertains, it seems, less to the concept "temperature" itself, and more to the validity (strictly universal vs. merely typical for macroscopic systems) of the 2nd law of thermodynamics.

There may be some very isolated and marginal "revolutionary" element when it is recognized that the 2nd law has only a statistical validity. That is: there is a genuine contradiction between the belief that spontaneous entropy decreases for closed systems are strictly and absolutely prohibited, and the belief that they can and will sometimes occur. Whether people actually believed the 2nd law in this absolute sense is perhaps a historical question of some interest. Whether they *should have* so held

---

[4] Actually, one might question "for the first time" on the grounds that, in principle, an ordinary thermometer could register the temperature of a "vacuum" – i.e., a region containing no matter, but only electromagnetic fields – and so render the attribution of a particular temperature to that region perfectly meaningful. Dealing with this in a serious way would involve too much of a tangent; I simply remark that, in practice, it is very difficult to do this, and so, in actual history, one does not find people attributing temperatures to high vacuum regions prior to the 20th century. For example, an 1892 "dictionary of universal knowledge" contains the following discussion: "Guesses have been made from time to time as to the temperature of space, Pouillet, for example, putting it at −238°F. and Fourier at −58°. From our present physical outlook, however, the phrase 'temperature of space' is meaningless. Only where matter is can a true temperature exist. A thermometer placed in space will receive radiations from all sides, and the temperature indicated will depend on the power it has to transform these radiations into the irregular motions which constitute heat in a body. An ideal thermometer, transparent to all radiations, and capable only of receiving heat by contact with other bodies, would remain unaffected if isolated in space." [31] Thus, while it was recognized that heat could be communicated from one body to another *by means of* radiation, the radiation field itself was not yet recognized as a physical system capable of *possessing*, itself, a temperature.

[5] in principle (see the previous footnote)

it is, I think, an epistemological question of considerable interest. But, here, my point is just that even if one grants for the sake of argument that people did – and that they should have – held the 2nd law to have this strict and universal character, Feyerabend's claim that the two temperature concepts are contradictory or incommensurable still makes no sense. Indeed, it is precisely the classical thermodynamic notion of temperature that is *meant* when one says that, according to the statistical mechanical viewpoint, it is *possible* (though exceedingly rare for macroscopic systems) for heat to flow spontaneously from a cooler body to a warmer one. The stability – the "meaning invariance" – of the temperature concept seems to be *presupposed* in the very act of noting the (alleged) contradiction between the two formulations of the 2nd law.

The reason Feyerabend perceives conceptual incommensurability in this kind of case emerges right away:

> "It is certainly possible to redefine the word 'temperature' so that it becomes synonymous with 'mean kinetic energy.' But it is equally certain that on this redefined usage the word has a different meaning from the one associated with it in the classical science of heat..." [2, p. 79]

As has been pointed out before (see, e.g, p. 55 of Ref. [3]) Feyerabend here simply takes for granted that the meaning of a concept is its definition. It then follows as a matter of course that theory change, perhaps resulting in a new definition, yields a concept with a new meaning – and one which is therefore fundamentally different from the original concept.

The identification of the concept's meaning with its definition thus leads to a conclusion which, in light of the above case-study, seems absurd: it's just obvious, for example, that Joseph Black (in the mid-18th century) and Max Planck (in the early 20th century) meant the same thing by "temperature" (even if the latter knew considerably more about the relevant phenomena than the former). What the case-study then suggests is the adoption of a different account of the meaning of a concept. In particular, what seems to be needed is a theory in which the concept means *all* of the characteristics of the referents – or more precisely, in which a concept means the referents themselves. Such a view was suggested already by Shapere, and developed more thoroughly by Hilary Putnam. [4] It is on the basis of such a view of meaning (explicitly acknowledged or not) that many commentators have denied the charges of conceptual incommensurability across episodes of theory change. [4–7]

From one point of view, this is all that's needed. A concept doesn't simply consist in and doesn't simply mean its definition, or some other condensed description (or statement of necessary and sufficient condiditions for class inclusion). The concept functions, rather, as a storehouse for collecting and organizing knowledge about a certain class of existents, and its meaning *is* those existents, including not only *all* of the already-established knowledge about them that is integrated by the concept, but also (by deliberate intention) facts about the referents which are not yet known. Then, in at least many cases in which some profound new discovery is made about the nature of the referents of a concept, and even in cases in which the concept's *definition* is updated in light of such new discoveries, it becomes possible to say that, nevertheless, the *meaning* of the concept – and indeed the concept itself – has not changed.

From another point of view, though, this answer alone is insufficient and not even very helpful. Yes, a concept means (at least in part) its referents. But the question is: how, precisely, does one in fact pick out a class of existents such that the class will (normally) remain stable – i.e., such that the associated concept will (normally) retain its meaning – even as knowledge continues to expand in the future? Putnam's account of reference (an extension, to concepts, of Kripke's account of the reference of proper names) doesn't help much here. Indeed, Putnam essentially dismisses – at least from the purview of philosophy – the very question: "Since, in many cases, extension is determined socially and not individually, owing to the division of linguistic labor, I believe that this problem is properly a problem for sociolinguistics." [4] At the end of the day, the process by which, according to Putnam, one forms a concept amounts to pointing at some concrete object (ostention) and declaring: "by 'X' I mean this and everything like it."

But: like it in what respect? Here Putnam retreats to what amounts to traditional moderate realism or essentialism. To use his own famous example, the meaning of "water" is ultimately: anything sharing *the same microstructural essence* as whichever bits of water were pointed to at the initial "dubbing". Thus, if it is later discovered that there is a substance sharing all of the "surface properties" of water (e.g., it is clear, potable, sometimes falls from clouds, etc.) it is nevertheless *not water* if its molecular structure turns out to be XYZ instead of H_2O. The point here is that, for Putnam, the knowledge one possesses *initially* about the referents of a concept – that is, the knowledge on the basis of which one first *forms* the concept – is actually *irrelevant* to the concept's meaning, i.e., to the determination of the precise class of existents to which the concept refers.[6]

---

[6] Some further asides about moderate realism and the issues of this paper are appropriate here. First, one of the interesting and important aspects of the concept "temperature" is that it is a concept for an attribute or property, not for a kind of entity. So it is prima facie implausible that the metaphysics of "natural kinds" – which is often deployed by or at least suggested by the defenders of the stability of conceptual reference across episodes of theory change, could apply in this case. But the kinds of issues (pertaining to the growth and development of science) that have *motivated* the deployment of this metaphysics for this kind of problem, are equally present for the case of "temperature" as in the more-commonly-discussed cases of concepts for

There is, then, a kind of underlying agreement between Putnam and Feyerabend. Arabatzis has nicely summarized Feyerabend (and Kuhn) as promoting "a contextual view of concepts, according to which concepts obtain their meaning from the theoretical framework in which they are embedded. It follows that when a theoretical framework changes the concepts embedded in it change too." [6] Putnam, as if conceding the premise that any role played by a background knowledge context renders the concept subjective[7] , provides an account of reference in which context plays no role at all – indeed, in which reference is determined by characteristics of which one was wholly unaware when the concept was initially formed.

The problem is, both views seem totally implausible as accounts of the formation and development of the concept "temperature" (which, I want to suggest, is typical). Previously-established background knowledge (e.g., that some things are hot and some are cold, that there is a kind of spectrum of degrees of warmth, that generally speaking cooler things warm up if placed near a hotter thing and vice versa, that for the most part things placed in contact with each other or with the same third thing will after some time attain roughly the same degree of warmth, but that there are puzzling exceptions to this, that air expands slightly when heated and contracts slightly when cooled, etc.) played an ineliminable role in the formation of the concept "temperature." The actual concept-formation process is *very different* from the sort of ostensive dubbing event that Putnam seems to have in mind. (Indeed, picking off a physical property (like temperature) *ostensively* seems inconceivable in principle: you can't point to "a temperature" or a mass or an electric charge or ...)

But contra Feyerabend, the background context of knowledge in which the "temperature" concept is born does not *get in the way* of the concept's having a stable, rigid meaning. For one thing, none of the relevant background beliefs (sketched in parentheses in the previous paragraph) are overturned by later developments (even if one allows the temporary wide acceptance of the caloric theory of heat as a "development"). This is the meaning of the claim that the later stages of development of the concept "temperature" illustrate cumulativity, not revolution. And anyway, it just seems completely clear that even, for example, in the 20th century – hundreds of years after the concept was originally formed – physicists who talk (say) about the temperature of the cosmic microwave background radiation, or attribute negative absolute temperatures to certain unusual kinds of systems, *mean exactly the same thing*, by "temperature", that people meant in the 17th century.

To summarize this discussion about the meaning and reference of concepts, it seems that what's needed is an account of concepts in which the stability of their reference obtains, not *in spite of* their arising in a context, but rather *by means of* their arising in a context[8] – in particular, by means of the specific knowledge in that context on the basis of which a certain class of existents is integrated in contrast to others from which its members differ.

By way of closing, let me simply list some features that I think a viable account of concepts would need to have in order to comply with and illuminate the features of the development of "temperature" which I have highlighted and which I consider to be typical of concepts in general. The theory must

- reject the discovery/justification dichotomy and locate the justificatory basis of concepts in the *process* of concept-formation (which will have to be emphasized and explained);
- respect and explain the hierarchical structure of knowledge in general and concepts in particular and thereby clarify how concepts can arise in a context without that implying Feyerabendian incommensurability or a failure of "meaning invariance";
- regard concepts as storehouses for collecting and organizing knowledge (about a certain class of existents) which storehouses are constructed, from the beginning, to be addition-friendly [8, p. 66-7] on the premise that knowledge is an ever-expanding sum;

(typically theoretical) entities. This of course suggests that the retreat to moderate realism is the wrong strategy even for the cases of concepts for (alleged) "natural kinds." On the other hand, endorsing such an approach and attempting to apply it to the case of "temperature" would leave one in a quandry, for there are not one but *two* theoretical developments which could quite plausibly be cited as the discovery of the true, micro-structural "essence" which legitimizes (and constitutes the real referent of) the concept. That is, which is going to be the analogy, for "temperature", of the discovery that water is essentially H_2O? That temperature is a measure of average molecular translational kinetic energy? Or that temperature is a measure of the amount of additional energy required (at constant volume) to increase the entropy by one unit? Both answers seem equally compelling, yet the proposed account would seem to require that one be privileged. This of course is a second way in which the concept "temperature" suggests against the (general) applicability of moderate realism. More positively, the case-study strongly suggests that the felt qualitative difference between knowing (say) that water is cool, clear, and potable – and knowing that water is H_2O – is illusory. In moving from the earlier stage to the later stage, one is not achieving, for the first time, a grasp of the single, true, final *metaphysical* "essence" of water. One is grasping a causally deeper distinguishing characteristic, to be sure. But this is a movement along a continuous and open-ended path, not a discrete step from the absence of a certain kind of knowledge to its presence.

[7] See here pp. 42-3 of Ref. [8] where Rand writes: "Concepts are not and cannot be formed in a vacuum; they are formed in a context.... [But t]his does not mean that conceptualization is a subjective process..."

[8] Cf. pp. 78-82 of Ref. [8].

- treat concepts as meaning their referents, with definitions treated not as exhaustive stipulations of meaning, but rather as condensations [8, p. 48] of what is known about the referents in a given context of knowledge.

Such a theory will inevitably support the perspective on the historical development of "temperature" that has been sketched above. The further conjecture, then, is that such a theory will also help us to see the degree to which the "temperature" concept is *normal*, i.e., indicative of the pattern of development of scientific concepts as such, and (thus) indicative also of the pattern of development of knowledge generally.

---

[1] Thomas S. Kuhn, *The Structure of Scientific Revolutions*, 3rd Ed., University of Chicago Press, 1996 (1962)

[2] P.K. Feyerabend, "Explanation, Reduction, and Empiricism" pp. 28-97 in *Scientific Explanation, Space, and Time*, Herbert Feigl and Grover Maxwell, eds., University of Minnesota Press, 1962

[3] Dudley Shapere, "Meaning and Scientific Change," in *Mind and Cosmos*, R.G. Colodny, ed., Univ. of Pittsburgh Press, 1966; reprinted in *Scientific Revolutions*, Ian Hacking, ed., Oxford, 2004

[4] Hilary Putnam, "The Meaning of Meaning", in *Language, Mind, and Knowledge*, K. Gunderson, ed., University of Minnesota, 1975

[5] Jonathan Bain and John Norton, "What Should Philosophers of Science Learn from the History of the Electron?" in *Histories of the Electron: The Birth of Microphysics*, Jed Buchwald and Andrew Warwick, eds., MIT, 2001

[6] T. Arabatzis, "Conceptual Change and Scientific Realism: Facing Kuhn's Challenge" in S. Vosniadou, A. Baltas and X. Vamvakoussi, eds., *Reframing the Conceptual Change Approach in Learning and Instruction*, Elsevier, 2007

[7] James G. Lennox, "Ayn Rand on concepts, context, and the advance of science", forthcoming in *Concepts and Their Role in Knowledge*, Allan Gotthelf, ed., Ayn Rand Society Philosophical Studies No. 2, Pittsburgh

[8] Ayn Rand, *Introduction to Objectivist Epistemology*, 2nd Edition, Plume, 1990

[9] Allan Gotthelf, "Ayn Rand on Concepts: Another Approach to Abstraction, Essences, and Kinds" forthcoming in *Concepts and Their Role in Knowledge*, *op cit.*

[10] J.C. Maxwell, *Theory of Heat*, Peter Pesic, ed., Dover, New York, 2001 (1871)

[11] W.E. Knowles Middleton, *A History of the Thermometer and its Uses in Meteorology*, Johns Hopkins Press, Baltimore, Maryland, 1966

[12] Thomas D. McGee, *Principles and methods of Temperature Measurement*, John Wiley and Sons, New York, 1988

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

[14] F. Steinle, "Concept Formation and the Limits of Justification: 'Discovering' the Two Electricities" in J. Schickore and F. Steinle, eds., *Revisiting Discovery and Justification*, pp. 183-195, Springer, 2006

[15] Duane Roller, "The Early Development of the Concepts of Temperature and Heat: The Rise and Decline of the Caloric Theory" in *Harvard Case Histories in Experimental Science*, Volume 1, James Bryant Conant and Leonard K. Nash, eds., Harvard University Press, 1957

[16] T.J. Quinn, *Temperature*, Academic Press Inc., London, 1983

[17] Martin Barnett, "The Development of Thermometry and the Temperature Concept" *Osiris*, Vol. 12 (1956), pp. 269-341

[18] F.S. Taylor, "The origin of the thermometer", Annals of Science, 5:2, 129-156 (1942)

[19] Hugh Young, *Mechanics and Heat*, McGraw-Hill, 1964

[20] Entry on "temperature", Oxford English Dictionary, http://dictionary.oed.com

[21] S. Psillos, "A Philosophical Study of the Transition from the Caloric Theory of Heat to Thermodynamics: Resisting the Pessimistic Meta-Induction", *Stud. Hist. Phil. Sci.* 25(2), pp. 159-190, 1994

[22] S. Carnot, *Reflections on the Motive Power of Heat*, R. Thurston, trans., 2nd edition, John Wiley and Sons, New York, 1897

[23] Rudolf Clausius, "Ueber die Art der Bewegung welche wir Wärme nennen," *Annalen der Physik*, Vol. 100, pp.353-80 (1857), english translation reprinted in Stephen Brush, *The Kinetic Theory of Gases*, Nancy Hall, ed., Imperial College Press, London, 2003

[24] J. C. Maxwell, "Illustrations of the Dynamical Theory of Gases," *Phil. Mag.*, Vol. 19, pp. 19-32; Vol. 20, pp. 21-37 (1860), reprinted in S.G. Brush, *op cit.*

[25] Willian Thomson, "On an Absolute Thermometric Scale", *Philosophical Magazine*, Vol. 23, pp. 313-7 (1848)

[26] Ernst Mach, *Die Principien der Wärmlehre*, Leipzig, 1896, translated excerpt in Brian Ellis, *Basic Concepts of Measurement*, Cambridge University Press, pp. 183-196 (1968)

[27] William C. McC. Lewis, *A System of Physical Chemistry*, Volume I, Kinetic Theory, Second Edition, Longmans, Green and Co., London, 1921

[28] S. Goldstein, "Boltzmann's Approach to Statistical Mechanics", in *Chance in Physics: Foundations and Perspectives*, edited by Jean Bricmont, Detlef Dürr, Maria C. Galavotti, Giancarlo Ghirardi, Francesco Petruccione, and Nino Zanghi, Lecture Notes in Physics 574 (Springer-Verlag, 2001)

[29] Charles Kittel and Herbert Kroemer, *Thermal Physics*, Second Edition, Freeman and Company, New York, 1980

[30] Max Planck, "Über die mechanishe Bedeutung der Temperatur und der Entropie" in *Festschrift Ludwig Boltzmann*, Barth, Leipzig, 1904, excerpts translated in Carlo Cercignani, *Ludwig Boltzmann: The Man Who Trusted Atoms*, Oxford, 1998

[31] Chamber's Encyclopedia: A Dictionary of Universal Knowledge, Vol. 10, 1892, full text available through books.google.com