Logical Positivism's Project: Formalization of Science

Nicholaos Jones

Logical Positivism is a philosophical movement from the early 20th century. It takes its name from two other movements. The first consists of efforts by mathematicians and logicians to translate natural language into formal language as a way to regiment our reasoning. Alfred North Whitehead and Bertrand Russell's *Principia Mathematica* is the paradigm work from this movement: written entirely in formal notation, following an ancestor of the logical system we now teach in our *Introduction to Logic* courses, it is an attempt to derive all of arithmetic from purely logical principles. The attempt fails. But the motivating idea is tempting, namely, that we can make intellectual progress on abstract issues by substituting the precision of formal language for the ambiguity of natural language. The Logical Positivists endorse this same motivation, albeit for philosophy and science more generally (not only for mathematics). Whence the "Logical" in Logical Positivism.

The second movement from which Logical Positivism takes its name is Positivism, a philosophy developed by the French philosopher Auguste Comte in the mid-19th century. Comte is not as well known as the British philosopher John Stuart Mill, who credits Comte's work with liberating him from the influence of Jeremy Bentham. But, from among his many contributions to developing foundations for the then-new discipline of sociology, Comte is best known for his so-called *law of the three stages*. The law is a claim about historical trends in human development, and it states that we pass through three successive stages:

1) first, a *theological* stage, wherein we search for ultimate causes and purposes, attribute apparent anomalies to supernatural agency (such as gods), and secure social cooperation and agreement through (threats of) violence and force;

2) second, a transitory *metaphysical* stage, wherein we continue our search for cause and purpose, replace supernatural agents with abstract entities (such as Platonic forms), and secure cooperation and agreement through lawyers and jurists; and

3) third, a *positive* stage, wherein we abandon the search for ultimate causes of phenomena in favor of discerning universal laws of nature that govern phenomena, wherein we abandon the search for ultimate purposes in favor of relative purpose, and wherein we secure cooperation and agreement through industrial means.

The Logical Positivists identify with the third of Comte's stages: they reject the pretensions of metaphysics, denying that it has any capacity to provide us with

knowledge about the world; and they endorse the efforts of scientists (sometimes called "natural philosophers") only insofar as those efforts are properly grounded in publicly accessible evidence—that is, concrete observations and experimental results, as opposed to abstract posits. Whence the "Positivism" in Logical Positivism.

Rudolph Carnap, one of the original Logical Positivists, epitomizes these inspirations in "The Elimination of Metaphysics through the Logical Analysis of Language." Here is his second paragraph:

> The development of *modern logic* has made it possible to give a new and sharper answer to the question of the validity and justification of metaphysics. The researches of applies logic or the theory of knowledge, which aim at clarifying the cognitive content of scientific statements and thereby the meanings of the terms that occur in the statements, by means of logical analysis, lead to a positive and a negative result. The positive result is worked out in the domain of empirical science: the categories are clarified; their formal-logical and epistemological connections are made explicit. In the domain of *metaphysics*, including all philosophy of value and normative theory [ethics], logical analysis yields the negative result *that the alleged statements in this domain are entirely meaningless*. Therewith a radical elimination of metaphysics is attained….

Carnap gives examples. Here is what he has to say about the word "God":

> In its *mythological* use, the word has a clear meaning. It, or parallel words in other languages, is sometimes used to denote physical beings [with special moral qualities]. Sometimes the word also refers to spiritual beings which…manifest themselves nevertheless somehow in the things or processes of the physical world and are therefore empirically verifiable. In its *metaphysical* use, on the other hand, the word "God" refers to something beyond experience. The word is deliberately divested of its reference to a physical being or to a spiritual being that is immanent in the physical. And as it is not given a new meaning, it becomes meaningless. To be sure, it often looks as though the word "God" had a meaning even in metaphysics. But the definitions which are set up prove on closer inspection to be pseudo-definitions. They lead either to logically illegitimate combinations of words (of which we shall treat later) or to other metaphysical words (e.g., "primordial basis," "the absolute," "the unconditioned," "the autonomous," "the self-dependent," and so forth), but in no cases to the truth-conditions of its elementary sentences.

For Carnap, and for other Logical Positivists, metaphysical statements "do not serve for the *description of states of affairs* [but instead for] the *expression of the general attitude of a person toward life*." Metaphysical statements are neither true nor false: they are more akin to poetry or music, albeit expressed with less artstry (as Carnap goes on to say). (So Logical Positivists are not atheists: they deny that *God exists* is capable of being true or false, and therefore neither affirm nor deny the existence of God.)

For statements that describe states of affairs (and therefore sometimes provide us with knowledge of the world), Logical Positivists hold that we must turn to

modern science rather than metaphysics (or theology).[1] Their reasons are complex, and the subfield of professional philosophy known as Philosophy of Science originates from the Logical Positivists' project of presenting and defending their reasons to others. But, complex as their reasons are, it helps to recall *why* they bother with the project: not from some hostility toward metaphysics, but from living through some of the most groundbreaking, exciting, and revolutionary breakthroughs in the history of science—namely, the onset of Relativity Theory and Quantum Mechanics—and wanting to understand how physical science managed such achievements…and why metaphysicians continue languishing over the same problems as the ancient Greeks.[2]

The project of Logical Positivism has a positive component and a negative one. The negative component involves classifying statements as *metaphysical* or *empirical* and explaining why the metaphysical ones are *meaningless* (in the sense of being neither true nor false). The positive component involves explaining why the empirical statements are meaningful (capable of being true or false), and further explaining why modern science is so successful in providing us with *true* empirical statements. We are going to focus on this positive project.

Explaining why modern science (especially physics) is the paradigm of successful inquiry about the world involves giving an account of its "epistemological underpinnings"—that is, an account of how science works, and why its working that way gives it such unrivalled success.

For Logical Positivists, giving an *account* of modern science's epistemological underpinnings means *reconstructing the reasoning scientists use to achieve their results*. "Reconstructing" does not mean *describing*; it means *making up reasoning that, in principle, could achieve the same results that the scientists achieve in practice*. Keep this in mind: philosophy of science is not sociology or psychology of science! For example, when Carnap says, "The work in theoretical physics consists mainly in constructing calculi and carrying out deductions within them," he is not *describing* what theoretical physicists do, and he is not claiming that they ever regiment their thinking with formal logic. Instead, he is characterizing an *ideally pristine technique* capable of simulating the results of theoretical physics, shorn of distractions due to human limitations.

For the sake of comparison and illustration, here's a Kantian-inspired account:

---

[1] Compare how one is supposed to know *that God loves everyone* with how knows *that the average force of gravity near the Earth's surface is 9.80665 meters per second per second*.

[2] General Relativity and the Standard Model of particle physics often are touted as the most rigorously and extensively confirmed scientific hypotheses of all time. The general attitude in the scientific community is that the evidence available (at least as of 2009) confirms these hypotheses to a high degree and that no available evidence disconfirms either hypothesis.

> Scientists make observations about correlations; those observations are passively "filtered" through "forms of intuition" in our minds; these "forms of intuition" automatically project causal relations onto the correlations; because the "forms of intuition" also determine how we experience the world, the projected causal relations turn out to be true whenever the observations about correlations are true.

This is, of course, *extremely* oversimplified. But you should be able to glean the basic account: modern science is successful in providing us with knowledge of the world, because it is especially good at *observing correlations*, and because our minds determine how we experience the world.

As you might expect, Logical Positivists have no patience for this kind of account: it's *metaphysical*, positing "forms of intuition" to which we have no empirical access; and it's *vague*, providing no detail about *how* projections of causal relation work. Logical Positivists want to do better: they want an account that is clear and precise and avoids metaphysical baggage. They think the key their ambition involves using the tools of modern formal logic. This is why Carnap discusses physical theories in terms of physical calculi and deductions.

Questions for Discussion

1- What are the components of a *physical calculus* (primitive signs, semantical rules)? In what sense do these components provide a *calculus*, and in what sense is the calculus *physical* as opposed to *mathematical*?

2- What is a *deduction* or *derivation* in a physical calculus? How does it work, and what components of the physical calculus does it involve?

3- What is the theory of thermic expansion? How does Carnap *formalize* it?

4- What is the general *procedure* for formalizing a physical theory? What's the *point* of formalizing a theory?

5- What kind of understanding of the world does modern science provide?

Supplement: A Guide to Carnap's *The Interpretation of Physics*

Carnap is one of the more demanding writers in the Logical Positivist tradition, especially if you don't have a background in mathematics, logic, *and* physics. So, just to get us going, I'll provide some guidance to his essay on the interpretation of physics. This should help you to follow what he is saying. It also might help you for other readings for the course, if you decide to imitate my technique.

Let's start with identifying some technical terms (words we don't usually encounter in our ordinary life) and trying to define them. We'll try for *pedagogical* definitions: a phrase saying what the term means, as well as examples to illustrate the meaning. I won't cover all the technical terms from Carnap's essay; but a few should get you going in the right direction, and these suffice for understanding the central theses of Carnap's essay.

abstract calculus: a collection of primitive signs (for names, predicates, and functions) as well as axioms formulated with those signs. Example: the signs 'a' and 'c' for name, the sign 'Px' for a predicate, the sign "=" for a function; the axioms 'a=c' and 'Pc'

interpretation: a set of rules assigning semantic meaning to the elements of an abstract calculus, specifying which names designate which objects, which predicates designate which properties, and which functions designate which relations. Example: Let 'a' designate the person named Angelface, 'c' the person named Carnap, 'Px' the predicate *is a philosopher*, and = the identity relation; then a=c means *Angelface is identical to Carnap*, and Pc means *Carnap is a philosopher*

physical theory: a universal statement, or collection thereof, describing certain features of worldly events or relations among such events

"When we use the word 'fact,' we will mean it in the singular sense [a particular observation about a particular individual] in order to distinguish it clearly from universal statements. Such universal statements will be called 'laws' even when they are as elementary as the law of thermal expansion…" (Carnap, An Introduction to the Philosophy of Science (Basic Books: 1966), p.5).

Now let's paraphrase the central theses from Carnap's essay, and identify the location where the ideas first occur in his text. I think there are two:

1- Any physical theory can be reconstructed as an abstract calculus and an associated customary interpretation of that calculus (p.311).

2- Physical theories need not be visualizable, nor even fully interpreted, to be understood (p.318).

Now let's give some examples of what Carnap means by his theses. A good way to do this is to give an *instance* for each thesis. I pull the first example from Carnap's text, because he gives it to us; but I add some detail I got from reading more about his example, because I'm not familiar with the physics he mentions. I also had to do some digging to spell out an example for the second, because Carnap only gives hints.

1- Any physical theory can be reconstructed as an abstract calculus and an associated customary interpretation of that calculus.

> Example: the theory of thermic expansion.

The law of thermic expansion governs how a rod expands lengthwise upon being heated to a certain temperature (see p.310). It states:

> For any solid body x with length l1 and temperature T1 at time t1, and with length l2 and temperature T2 at time t2, and with coefficient of thermal expansion β, $l2 = l1(1 + \beta(T2 - T1))$.

This law helps to explain (among other things) why roads buckle during period of high heat: when the temperature of the road rises, the side-to-side length of the road increases; when the road expands, it hits up against compacted dirt, which is also expanding; if the dirt is stronger than the road, the road will buckle to make room for its expansion.

Here's how Carnap "reconstructs" this law as an abstract calculus with a customary interpretation. First, he stipulates "primitive signs" for names, predicates, and functions along with some "semantical rules":

> x ranges over solid bodies; t1 and t2 range over times, L(_,_) denotes length of _ at _ (in meters), and T(_,_) denotes temperature of _ at _ (in Centigrade); C (_) denotes coefficient of thermal expansion of; = denotes mathematical identity, + denotes addition, – denotes subtraction, * denotes multiplication; other symbols denote specific numerical values

Then he writes the following statement with the primitive signs as well as some other standard signs from formal logic (like the upside down A, which means "for all," the ampersand which means "and" and the arrow which means "only if"):

$$\forall x \forall t2 \forall t1 [(L(x,t1) = l1 \ \& \ L(x,t2) = l2 \ \& \ T(x,t1) = T1 \ \& \ T(x,t2) = T2 \ \& \ C(x) = \beta) \rightarrow (l2 = l1*(1 + \beta*(T2 - T1)))]$$

The semantic rules for the primitive signs specify a customary interpretation for this statement. So interpreted, the abstract formula states the law of thermic expansion: it means just what the natural-language version of the law means.

Now let's give an example for Carnap's second main thesis.

2- Physical theories need not be visualizable, nor even fully interpreted, to be understood.

Example 1: the "wave function" in quantum mechanics.

The so-called wave function Ψ has no customary interpretation in quantum mechanics. No one is quite sure what it designates in the physical world, or whether it designates anything at all. The relevant law in which the sign appears is the Schrodinger Equation, which you can look up on your own.

Example 2: Maxwell's equations for electricity and magnetism.

There are no "picturable" models for Maxwell's equations. For example, Gauss's law (one of Maxwell's equations) states that the total electric flux through a closed surface is proportional to the charge enclosed. (You can look up the equation on your own.) There's no way to visualize "total electric flux."

Carnap is saying that we don't need to know what the wave function designates in order to understand the Schrodinger equation, and that we don't need to visualize *total electric flux* in order to understand Gauss's law.

So much for examples. Let's finish by trying to figure out the *reasons* Carnap has for endorsing his theses as correct.

1- Any physical theory can be reconstructed as an abstract calculus and an associated customary interpretation of that calculus.

I think Carnap attempts to prove this by giving a procedure for reconstructing physical theories. His idea seems to be that giving a general procedure that applies for all physical theories will thereby show that all physical theories can be reconstructed, even if he doesn't actually attempt to implement the procedure for all theories. Here's the procedure (p.309-312):

> Translate each universal statement from the theory into the language of a formal logic (such as predicate logic).

Carnap's example about the law of thermic expansion is a good example of how this procedure is supposed to work. The symbolic claims obtained by translation

are the abstract calculus, and the "translation key" for the translation provides the customary interpretation for the calculus. If the translation is accurate, the translated statement should have the same meaning as the original.

2- Physical theories need not be visualizable, nor even fully interpreted, to be understood.

Here Carnap's argument seems to be more conceptual and abstract. He specifies the kind of understanding he has in mind, and then he argues that physical theories give that kind of understanding. (This approach has its limits: Carnap admits that there are no *other* kinds of understanding such as "intuitive" understanding; but he does not argue that *his* kind of understanding is better, or that other kinds are not valuable.)

I like to present arguments in "standard form," as a series of numbered premises followed by the conclusion. So I'll do that for Carnap, even though his argument is scattered through his prose; and I'll cite where I locate each premise in his text.

1. Understanding X means being able to use X for description or prediction (p.317).
2. Translating a physical theory into an abstract calculus suffices for yielding descriptions and predictions, even without an interpretation of that calculus (proof by example on p.311).
3. Hence, physical theories need not be fully interpreted in order to be understood.

Positivists on Reference: Laws and Correspondence Rules

Nicholaos Jones

Reading: Carl Hempel, "On the 'Standard Conception' of Scientific Theories"

Scientists explain, predict, and control the natural world using *theories*. There are Newtonian and Einsteinian theories about gravitation, caloric and kinetic theories of heat, *qi*-balance and germ theories of disease, Aristotelian and evolutionary theories of speciation, wave and particle theories of light.

As part of their project to understand the success of scientific inquiry, logical positivists sought to understand how theories work. They did so, in part, by examining *what theories are made of*. Their inquiries converged upon a rough consensus known as the *standard conception* of scientific theories. This is a *conception*, because it purports to say what theories are; and it is *standard*, because it is the conception to which all subsequent alternatives responded.

Hempel presents, assesses, and suggests revisions for the standard conception of theories. His criticisms typically come from attending to details about what's going on in science: logical positivists tended to let logical considerations drive their picture of science, neglecting details of scientific practice, and Hempel is among the first positivists to reverse this tendency. Here I focus on presenting the standard conception (which Hempel presumes we already know) and explaining why Hempel suggests various revisions.

Logical positivism understands scientific theories as *linguistic entities*:

(1) Scientific theories are collections of declarative statements.

This contrasts with understanding theories as, say, ideas in people's heads, or ways of thinking about the world. Saying theories are linguistic means we can find written down in textbooks. It also means they are the sorts of things we can translate into the language of formal logic, and that we can use them to make logical deductions. Hempel makes no objection to this posit.

Logical positivism divides the terminology—the subject- and predicate-terms—scientists use to formulate their theories into two disjoint categories:

(2) Every term in a constituent statement of a scientific theory is either a theoretical term or (exclusively) an observational term.

A term is defined as *observational* whenever there is a specific and public procedure for directly determining, without the assistance of technical instruments, whether and how it applies or holds in the natural world. For example, *distance* is an observational term, because we can use rulers to measure it; *weight* is as well, because we can use scales to measure it. All other terminology that appears in constituent statements of scientific theories is *theoretical*. For example, *mass* is a theoretical term, because we have to perform a calculation in order to measure it: we have to divide an object's weight by the local magnitude of gravitational acceleration (m = w/g).[1] *Temperature* is also a theoretical term, but for a similar reason: we measure it by using a thermometer, but what we *directly* measure with thermometers is (say) the height of liquid in a tube, and we use that height to infer temperature. Hempel makes no objection to this posit, either.[2]

Logical positivism also divides the constituent statements of a scientific theory into two disjoint categories:

(3) Every statement in a scientific theory is either an internal principle *or* (exclusively) a bridge principle.

I'm using Hempel's labels, but others refer to "internal principles" as theoretical laws, and to "bridge principles" as correspondence rules. The "standard conception" defines each kind of statement as follows:

(3a) Internal principles are those statements composed *only* with theoretical terms.

(3b) Bridge principles are those statements composed with *both* theoretical and observational terms.

Newton's third law of motion is a good example of an internal principle: the force one body exerts on a second is equal and opposite to the force the second exerts on the first ($\mathbf{F}_{A \to B} = -\mathbf{F}_{B \to A}$). Newton's law of universal gravitational ($\mathbf{F_g}(m)$ =

[1] A body's mass, or its amount of matter, is constant; but its weight can change, which is why astronauts "weigh" much less on the moon than on Earth. Weight, properly understood, is a measure of how strongly the force of gravity pulls on a body.

[2] Later critics were fond of arguing that there is no sharp boundary between theoretical and observational terms, because we cannot determine by syntax alone the category to which any scientific term belongs. The positivist response is that *vague boundaries* are boundaries nonetheless, and that it is perfectly possible for a term to shift, over time, from being theoretical to being observational. (That's what happened with *temperature*.)

$G \times m \times m_R / d^2$) is a good example for a bridge principle, because *gravitational force* and *mass* are theoretical terms but *distance* is an observational term.[3]

Hempel criticizes the *definitions* of "internal principle" and "bridge principle." He prefers different, *theory-relative* definitions:[4]

> (3a*) The internal principles of a theory are statements that use only theoretical terms *distinctive to the theory*.
>
> (3b*) The bridge principles of a theory are statements that use both theoretical terms *distinctive to the theory* and *either* observational terms *or theoretical terms from antecedently accepted theories*.

Here is Hempel's example: according to the kinetic theory of gases, the temperature of a gas is proportional to the gas' mean kinetic energy ($T = \frac{1}{2}m \times v^2$). To understand why Hempel wants to say this is a bridge principle, you need to know that the kinetic theory of gases explains the measurable properties of gases—their temperature, pressure, and volume—as resulting from the motion of their atomic constituents. For a gas with n atomic constituents, such that atom #j has kinetic energy equal to $\frac{1}{2}m_j \times v_j^2$, the mean kinetic energy of the gas is

$$\frac{1}{2}m \times v^2 = (\frac{1}{2}m_1 \times v_1^2 + \frac{1}{2}m_2 \times v_2^2 + \frac{1}{2}m_3 \times v_3^2 + \ldots + \frac{1}{2}m_n \times v_n^2) / n.$$

Mean kinetic energy ($\frac{1}{2}m \times v^2$) is a theoretical term *in the kinetic theory of gases*: first, because we can't measure the mass or velocity of any individual atom; second, because it is a term *distinctive to the theory*. Temperature, by contrast, is not a theoretical term in the kinetic theory of gases. *But it is a theoretical term*. To mark the difference between *temperature* and *mean kinetic energy* with respect to the kinetic theory of heat, Hempel calls it an *antecedent term*: it is a theoretical term *from a theory accepted prior to the kinetic theory of heat*.

You might think Hempel is being pedantic in revising the standard conception's understanding of bridge principles For example, you might think it doesn't really matter whether we classify the claim $T = \frac{1}{2}m \times v^2$ as a bridge principle (as Hempel prefers) or an internal principle (as the standard conception prefers). But Hempel isn't being pedantic. He's trying to solve a substantive (non-pedantic) objection to

---

[3] The gravitational force a body experiences equals the product of the body's mass and the mass of a reference body "exerting" a gravitational force on the body, divided by the square of the distance between the bodies, all multiplied by the so-called universal gravitational constant.

[4] Here is N.R. Campbell: "A theory is a connected set of propositions which are divided into two groups. One group consists of statements about some collection of ideas which are characteristic of the theory; the other group consists of statements of the relation between these ideas and some other ideas of a different nature" (*Physics: The Elements* (1920), 122).

the standard conception. We need to introduce another posit from the standard conception in order to understand the objection.

(4) The "outputs" of scientific theories are explanations and predictions of *phenomena*.

For example, one output of Newtonian gravitational theory is an explanation for why the tides rise and fall: the gravitational force the moon exerts on our oceans changes as the distance between the moon and the oceans changes.[5] Insofar as theories are linguistic entities, it follows that *phenomena* must be linguistic entities as well: the "outputs" of statements are yet more statements. Insofar as the statements in theories contain *theoretical terms*, it is natural to suppose that *phenomena* contain no theoretical terms—because phenomena are outputs of, rather than parts of, theories. Hence:

(4a) A *phenomenon* is a statement composed *only* with observational terms.

The substantive objection to the standard conception is that, if this is what phenomena are, the standard conception is much too narrow in saying what scientific theories explain and predict. For example, the kinetic theory of gases explains Charles' Law, according to which the volume and temperature of a gas are directly proportional at constant ($V = kT$). But Charles' Law contains a theoretical term, namely, temperature, which disqualifies it from being a phenomena *as the standard conception defines that term*. Hempel's response to this objection is to redefine *phenomena*, as follows:

(4a*) A *phenomenon* for a theory is a statement composed *only* with observational terms and/or *theoretical terms from antecedently accepted theories*.

There's one more posit to consider from the standard conception:

(5) Theoretical terms from scientific theories typically are *not* completely definable with observational terms.

Consider, for example, *gravitational force*. Newton's law of universal gravitation says that the gravitational force on a body $\mathbf{F_g}(m) = G \times m \times m_R / d^2$. This is as close as we can get to a definition. But, while *distance* is an observational term, *mass* is not. The mass of a body is defined as the quotient of its weight and its gravitational acceleration: $m=w/g$. *Weight* is an observational term. But we can't define

[5] This follows from Newton's law of universal gravitation. if $\mathbf{F_g}(m) = G \times m \times m_R / d^2$, reducing the distance d means increasing the gravitational force (because all other terms are constant), and the stronger gravitational force "pulls" the ocean water farther up onto the shores.

gravitational acceleration only with observational terms, because we define a body's acceleration as *acceleration at an instant* and we can't directly measure instantaneous acceleration.

Now recall the logical positivist criterion for cognitive meaning:

> (CM) Declarative statements are (cognitively) meaningful only insofar as they are composed with observational terms or terms *completely definable with* observational terms.[6]

This criterion, together with the prior posits (1) through (5), entails that *scientific theories are cognitively meaningless*. That is: internal and bridge principles are *neither true nor false*. This is counterintuitive. Hempel considers responses:

> (R1) *Models* provide *analogical* meaning for scientific theories.
> (R2) Bridge principles are (cognitively) meaningful because they are *true by definition* (convention, stipulation).
> (R3) Scientific theories are *pragmatically* meaningful, useful rules to endorse for making calculations.

He prefers the third: "We come to understand [theoretical] terms, we learn how to use them properly, in many ways besides definition: from instances of their use in particular contexts, from paraphrases…, and so on."

Questions for Discussion

1- Do you agree with the five posits in the standard conception? What reasons or evidence support your stance?

2- Do you agree with Hempel's definitional revisions?

3- Do you find it *counterintuitive* to say that scientific theories are cognitively meaningless? If so, where does logical positivism go awry—and what alternative posit(s) would you suggest as more plausible?

Supplement: Energy Theory

Examples help with understanding the standard conception of scientific theories. Here's one from what I'll call energy theory.

According to energy theory, all bodies contain *energy*. Energy is that which gives bodies their capacity for movement. It comes in two forms. *Kinetic energy* is the

[6] *Objection*: This posit isn't meaningful, because *meaningful* is a theoretical term. *Reply*: Correct. But it is *pragmatically* meaningful, a useful rule to endorse for understanding scientific inquiry.

energy the body is using to move. *Potential energy* is the energy the body has for moving but is not using. Potential energy also comes in several forms. These include *elastic potential energy*, which is the potential energy bodies have for moving in the way a spring or rubber band move; and *gravitational potential energy*, which is the potential energy bodies have for moving toward other objects.

The central internal principle for energy theory is known as the *Law of the Energy Conservation*. It says that:

> *The total energy of a body at some initial time is equal to the total energy of the body at every other time*.

If we represent kinetic energy as K and potential energy as U, and if we use subscripted 'i' and 'f' to represent different *initial* and *final* times, we can write this law as a mathematical equation:

$$K_i + U_i = K_f + U_f.$$

Energy theory also provides a range of bridge principles that connect each form of energy with "observational" conditions. For example, according to energy theory, a body's kinetic energy is half the product of its mass m and the square of its velocity v:

$$K = (mv^2)/2.$$

This means that bodies moving faster increases a body's kinetic energy. So, for example, when a rock is resting at the top of a hill, its kinetic energy is 0; as it gains speed rolling down the hill, its kinetic energy increases; and as it slows down after reaching the bottom of the hill, its kinetic energy decreases toward 0.

According to energy theory, a body's *potential energy* U equals the product of its mass, its distance from some reference point (such as the surface of the Earth), and it free-fall acceleration g:

$$U = mgy.$$

This means that lifting a body away from the surface of the Earth increases its gravitational potential energy.

There's one more bridge principle we need for our example. It's an import from Newtonian gravitational theory, and says that a body with mass m, at some distance d from a reference point with mass M, has a *gravitational potential energy* U_G, such that

$U_G = -G(Mm)/d.$

G is a constant known as the *universal gravitation constant*.

If we want to understand how energy theory works *according to the standard conception*, it's best to pick see how the standard conception reconstructs the reasoning involved in using the theory to make a prediction. So here's a scenario: Imagine there's an asteroid heading toward Earth. Using some fancy measuring tools, suppose we are able to determine that the asteroid has a speed of 12 kilometers per second when it is a distance of 10 Earth radii ($10R_E$) from the center of the Earth. Suppose we want to use energy theory to make a prediction about the speed of the asteroid when it reaches the Earth's surface (at 1 Earth radius ($1R_E$) from Earth's center).

Here's what we do. First, we specify the observable inputs. To keep our calculations simple, we can make some simplifications. For example, let's imagine that the mass of the asteroid is 0, because it's so negligible compared to the Earth's mass (at 5.98×10^{24} kilograms). So we have, as inputs (in uints m for meters, kg for kilograms, and s for seconds:

$G = 6.67 \times 10^{-11}\ m^3s^2/kg$
$R_E = 6.37 \times 10^6\ m$
$M = 5.98 \times 10^{24}\ kg$
$m = 0\ kg$
$d_i = 10R_E$
$d_f = 1R_E$
$v_i = 12 \times 103\ m/s$

Next, let's work just with energy theory. Let's also imagine that the asteroid only has the gravitational form of potential energy, because the gravitational force on the asteroid is so terrifically large compared to, say, the frictional force it experiences while falling through the atmosphere. Then the Law of Energy Conservation entails:

$$(mv_i^2)/2 - G(Mm)/di = (mv_f^2)/2 - G(Mm)/1R_E.$$

That is: the sum of the asteroid's initial kinetic energy and its initial gravitational potential energy (when measured by us) equals the sum of its "final" kinetic energy and its "final" gravitational potential energy (just before smashing into Earth).

Because we're trying to find v_f, we can use algebra and substitute our input values to solve our theoretical equation for v_f. Doing so yields that

$v_f = 1.60 \times 10^4$ m/s.

So energy theory predicts that the asteroid will hit the surface of the Earth with a speed of 16 kilometers per second.

Ignore the math required to make this prediction, and focus on the procedure: we made some measurements; we used those measurements as input for an abstract mathematical equation; we solved the equation; the solution gave us a measurable prediction.

Note that, at no point in this procedure, did we need to *interpret* the symbols for kinetic energy (K) or gravitational potential energy (U_G). Note, more importantly, that we made the prediction without every *defining* the central notion of *energy*, and without ever providing a way to visualize what energy is. All we needed to do was understand how to manipulate mathematical terms. This is why logical positivists are not too concerned with saying that internal principles, such as the Law of Energy Conservation, are cognitively meaningless: they don't need to have a meaning in order for us to use them.

Positivists on Explanation: Nomological Account

Nicholaos Jones

Reading: Carl Hempel, “Two Models of Explanation"

If, like the positivists, you endorse an instrumentalism about scientific theories—that they are calculating devices that don't make claims about the way the world is—then you might think that scientific theories don't *explain* anything. Here's the line of thinking you might have:

1. Explanations provide understanding.
2. Understanding is a matter of giving answers to "why?" and/or "how so?" questions.
3. Answers to "why?" and "how so?" questions must have cognitive meaning.
4. Scientific theories lack cognitive meaning. (This is the positivism part.)
5. Hence, scientific theories don't explain (because they don't give answers to the right sorts of questions, and so they don't provide understanding).

If scientific theories don't explain anything, that doesn't mean they're useless. Certainly they make predictions, and predictions are useful for purposes of controlling or manipulating nature.

Now you don't *have* to be a positivist to think that scientific theories don't explain anything. Here's John Macmurray, from his *Reason and Emotion* (1935):

> Science, though it may know everything in general, can know nothing in particular, and reality is always something in particular. It follows that science is not knowledge of reality. Is that startling? … Science is descriptive, not explanatory…. Knowledge is by definition the apprehension of the real, not the description of it…. I do not say that it is not cognition, but simply that it is not knowledge in the full sense. For example, you cannot know anybody, your father or your friend, by science (1935, 187).

Macmurray holds that science doesn't explain, but his argument is entirely different, and involves *endorsing* scientific theories are having cognitive meaning:

1. Explanations provide knowledge.
2. Knowledge is a matter of apprehending the real.
3. Descriptions do not apprehend the real.
4. Scientific theories only provide descriptions. (This means they have cognitive meaning.)
5. Hence, scientific theories don't explain.

Macmurray's style of argument doesn't hold much sway with philosophers of science. It doesn't say as much as it seems to say. Macmurray's argument shows that scientific theories don't provide what we might think of as *intimate* understanding—the sort of understanding we have when we "know" a person. It's the sort of argument you often find given by people who think scientific inquiry threatens their worldview. But it is an odd way to argue, because no one ever thought otherwise.

When philosophers of science say that scientific theories explain, they don't mean to say that science provide *intimate* understanding: because intimate understanding requires apprehension of the real, it can't be conceptual, but science is *necessarily* conceptual. Instead, philosophers of science means that scientific theories provide some sort of *conceptual* understanding, the sort of understanding we achieve when we're able to conceptualize particulars as instances of more general patterns. Like all conceptual activity, this kind of understanding involves *deliberately ignoring* whatever makes things *unique* and focusing, instead, on what makes them *similar*.

Hempel, like most other philosophers of science (positivists and otherwise), thinks that scientific theories give (conceptual) explanations about the world. He also thinks that (conceptual) explanations provide (conceptual) understanding.

> [T]here is a second motivation for the scientific quest, namely, man's insatiable intellectual curiosity, his deep concern to *know* the world he lives in, and to *explain*, and thus to *understand*, the unending flow of phenomena it presents to him.

He also thinks that understanding is a matter of giving answers to "why?" questions—or, as he puts it, in an effort to contrast scientific and mythological inquiry, a matter of answering "questions as to the *what* and the *why* of the empirical world."

Because Hempel is a positivist, he denies that scientific theories have cognitive meaning. Hence, in order to avoid the conclusion that scientific theories don't explain, he need to deny that answers to questions must have cognitive meaning. His strategy for doing so involves identify an alternative way to answer questions—a way that doesn't require the answers to have cognitive meaning—and then to show that his alternative is a good approximation to what practicing scientists do when they provide explanations.

Hempel provides two "models" of explanation as alternative wys to answer explanation-seeking questions. The first, known as the Deductive-Nomological Model, or sometimes as the Covering law Model, is so influential among

philosophers of science (past and present) that many refer to it as the *standard conception* of scientific explanation.

Hempel's model presumes familiarity with the distinction between *theoretical* and *antecedent* statements. The theoretical statements of a scientific theory contain at least one theoretical term that's new to that theory. Antecedent statements contain no such terms: all of their terms are either *observational* or else theoretical terms *that are antecedently understood*. In the context of giving models for explanation, Hempel likes to refer to the theoretical statements as *laws*. This can be confusing: for Hempel, laws are *statements*. To avoid confusion, I'll speak about *law-statements* rather than laws, to make explicit their linguistic status.

Hempel's model also interprets explanation-seeking *questions* as questions about why or how *some statement is true*. For example, when someone asks "Why do the planets orbit the sun?" Hempel interprets as a request to explain why *the planets orbit the sun* is true. This means that for Hempel, there's no such thing as explaining something that isn't true: we can explain why it's *true* that something is false, but we can't explain why something false is true, or why something true is false. Hempel calls the *explanandum* the *statement* for which an explanation-seeking question is requesting an explanation; and he calls what the explanandum is *about* the *explanandum-event*. So, for example, the statement *the planets orbit the sun* is an explanandum, and this statement is about the explanandum-event that is the planets orbiting the sun.

Hempel's basic idea is that answering an explanation-seeking question is a matter of *giving an argument*. He means "argument" in the philosopher's sense: a series of statements (called premises) which, taken together, support some other statement. Hempel's model is deductive, because he thinks the argument should be deductive in form: it should be possible to show that the conclusion follows from the premises by using standard rules of deductive logic (of the sort you might learn in Introduction to Logic courses). His model is *nomological*—from the Greek word *nomos*, which means law—because he thinks that some of the premises in the argument must be law-statements.

Questions for Discussion

1- What do explanations provide? Understanding, knowledge, something else?
2- Is explanation a matter of answering explanation-seeking questions? If not, what else might it be?
3- Must answers to questions have cognitive meaning? Can *arguments* (in the philosopher's sense) provide answers to questions?
4- How well does Hempel's model fit scientific practice?

Supplement: Thermodynamic Explanation of Soda Fog

Thermodynamics posits an interdependence between the pressure, temperature, and volume of gases. One of its central internal principles is the *Ideal Gas Law*, according to which the product of the pressure and volume of a gas is proportional to the temperature of the gas. If we represent gas pressure as P, volume as V, and temperature as T; and if we use k to stand for a proportionality constant, we can write the Ideal Gas Law as a mathematical equation:

$$PV = kT.$$

Some other important internal principles from thermodynamics relate pressure, temperature, and volume to *internal energy*. If we represent the internal energy of a gas as E, and if we use k_1 and k_2 as some more proportionality constants, we can mathematize two of these relations:

$$V = k_1/E.$$

$$T = k_2E.$$

The first says that gases use internal energy as they expand: as volume increases, internal energy decreases. For example, if you open the mouth of an inflated balloon, the energy the balloon has to move decreases as the air trapped inside the balloon expands into the environment (so volume and kinetic energy are inversely proportional). The second says that temperature is a source of internal energy: as temperature rises, internal energy rises. For example, if you heat the air trapped in a hot air balloon, the balloon rises—the gas inside moves upward.

We can use these relations to construct Hempel-style nomological explanations. For the sake of concreteness, suppose we have an unopened bottle of carbonated soda, and we want to explain why "soda fog" forms at the lip of the bottle when the bottle is first opened.

We know that the trapped air is a gas composed of water vapor and carbon dioxide (or whatever gas is used for carbonation). We also know that the pressure of this gas is greater than the atmospheric pressure (outside the bottle). So we know that, when we open the bottle, the pressure of the trapped air will decrease. These are the initial conditions for our explanation.

Now let's use thermodynamic theory to explain what will happen to the trapped air. When we open the bottle, its pressure decreases. Hence, according to the

Ideal Gas Law, the volume of the trapped air must increase: when we open the bottle, the trapped air must start to expand beyond the mouth of the bottle.

According to thermodynamics, gases use internal energy as they expand: as their volume increases, their internal energy decreases. But if their internal energy decreases, their temperature does as well. Hence, as the trapped air expands beyond the mouth of the bottle, thermodynamics entails that its temperature will drop. If you recall that making gases colder causes them to condense into liquid, this amounts to predicting that, when we open the bottle, the trapped air will convert itself into a vapor. And this is why the trapped air in a soda bottle forms a "fog" when we open the bottle.

Case Study—The Caloric Theory of Heat

Nicholaos Jones

1 Heat

We are going to develop a sustained case study from the history of science. Our case will not be entirely historically accurate. We will do some conceptual reconstruction and simplification, and some blending of ideas from different times. This sacrifices scholarly rigor for the sake of pedagogical clarity. But this is appropriate: the point of our case study is to illustrate relatively abstract ideas about scientific methodology, rather than to support those ideas against their competitors.

Our case study centers upon the scientific study of *heat*. Rub your hands together quickly and repeatedly. You will notice they get *hot*. Were you able, you'd also notice that the temperature of your hands remains the same. Now boil some water. You'll notice that it evaporates into steam. The boiling water holds steady at 100 degrees Celsius; but the steam is hotter than the water. The study of heat is the study of what it is that makes your hands hot when rubbed and steam hotter than boiling water.

Hellenistic approaches to heat, dominant from Aristotle's time, in the schools of atomism and Epicureanism, and through the scholastic tradition of medieval Europe, understood fire—one of the four elements, along with water, earth, and air—as a physical manifestation of heat. These approaches were never well developed, and they failed entirely to distinguish between heat and temperature. But they left a legacy for thinking of heat as a physical manifestation of some material substance. This *materiality of heat* hypothesis persisted even as the *mechanical* philosophy, positing particles interacting through forces (led by Rene Descartes, Robert Boyle, and Isaac Newton, among others), displaced scholastic philosophy.

Sustained scientific attention to heat began in the late 1700s, especially in Scotland and France. This attention continued through the 1800s, and produced a series of major scientific theories about phlogiston, caloric, and statistical motion. Scientific inquiries culminated in what, today, we study as Thermodynamics and Statistical Mechanics—first in their "classical" forms, later in the "quantum" extensions. But the study of heat was a subject all its own until around the middle of the 19th century (1850-1860).

We will be ignoring more recent treatments of heat, from thermodynamics and statistical dynamics, in favor of earlier—and now superseded—treatments. We also will be ignoring various issues associated with scientific instrumentation. Scientific studies of heat are scientific, at least in part, by virtue of the role of

measurement and experimentation. Such empirical investigations require instruments, and there are significant epistemic issues relating to how we come to know that our instruments measure what we intend them to measure, and how we come to know that such measurements are reliable. We are going to ignore these issues.

1.1 Measuring Heat

For our purposes, we only need to know that there are two kinds of scientific instrument for studying heat, thermometers and calorimeters. Thermometers measure *temperature*. The thermometers in common use throughout the 18th and 19th centuries were typically closed glass tubes.

The tubes were filled with air, alcohol ("spirit of wine"), or mercury ("quicksilver"), and they were marked with intervals of equal length. Ideally, the thermometric fluid inside the tube would rise and fall in accordance with the rise and fall of the temperature of the body being measured, so that the position of fluid inside the tube with respect to the markings on the tube indicates the temperature of the body being measured.

Thermometers, useful as they are, do not directly measure *heat*. It is possible for bodies to gain or lose heat without changing their temperature. For example, adding heat to boiling water, rather than increasing its temperature, converts it into its gaseous state; and percussion or friction can heat solids without increasing their temperature. It is also possible for a body to increase or decrease its temperature without gaining or losing heat. For example, certain chemical reactions and mixtures change temperatures of bodies but not their heat. (In our day, microwave ovens raise temperatures without adding heat.)

In the late 1700s, Antoine-Laurent Lavoisier and Pierre-Simon Laplace devised an "ice calorimeter" to directly measure heat. Their device involved four sections: a hollow and spherical metallic chamber, a hot substance inside the chamber, crushed ice outside of and surrounding the chamber, and a collection dish below the chamber to collect melted ice.

Lavoisier-Laplace's Calorimeter. Credit: Gregory Tobias, Chemical Heritage Foundation

Lavoisier and Laplace's idea was to use the quantity of melted ice as a measure of the hot body's heat. But the ice calorimeter has various practical limitations, chief among them being that melted water often adheres to crushed ice rather than flowing into the collection dish, rendering calorimeter-based measurements of heat unreliable.

1.2 Heat Phenomena

So much for measurements of heat, and the distinction between heat and temperature. For the sake of fixing the referent of our case study, consider several phenomena associated with heat and well-known to scientists of the 18th and 19th centuries. We have mentioned several already:

P1: Adding heat to boiling vaporizable liquids (such as water) changes their state (from liquid to gas) without changing their temperature.

P2: Melting ice maintains a constant temperature.

P3: Friction and percussion add heat to solids without changing their temperature.

P4: Some chemical reactions alter a body's temperature without adding to or detracting from the body's heat.

Some other phenomena include:

P5: Slowly compressing (expanding) a gas increases (decreases) its temperature.

P6: Heat radiates (transmits across macroscopic distances) apparently instantaneously and without requiring a material medium.

P7: Mirrors and other surfaces can reflect and refract radiant heat.

P8: Bodies expand when heated and contract when cooled.

2 Laws and Correspondence Rules in "the" Caloric Theory of Heat

Readings: Sanborn C. Brown, "The Caloric Theory of Heat," *American Journal of Physics* 18.6 (1950), 367-373.

Nicholas W. Best, "Lavoisier's 'Reflections on Phlogiston' II: On the Nature of Heat," *Foundations of Chemistry* 18 (2016), 3-13.

2.1 Nature of Caloric

The caloric theory of heat rose to prominence during the early to mid-1800s. There were in fact many caloric theories of heat, with several variants throughout both France and Scotland. The caloric theory of heat's central and distinctive posit is the existence of *caloric* ("calorique," as coined by the Frenchman Lavoisier in the 1780s, and formerly referred to as "igneous fluid / *fluide igne*," "matter of fire / *matiere du feu*," or "matter of heat") as the cause or ground of heat.

> [W]e have distinguished the cause of heat, or that exquisitely elastic fluid which produces it, by the term of caloric.[1]

Combining the *materiality of heat* hypothesis with the *mechanical* approach to substances as collections of interacting particles, the theory treats caloric as an aggregate material *substance* composed of *interacting material particles*.

As a *substance* (*corps*), caloric is an object rather than an event.

---

[1] Antoine Lavoisier, Elements of Chemistry: In a New Systematic Order, Containing All the Modern Discoveries, translated by Robert Kerr (University of Adelaide: eBooks@Adelaide, 2014), Part 1, Chapter 1, Paragraph 8.

As an *aggregate*, caloric is a whole composed of particles. This follows from the mechanical philosophy, according to which the fundamental elements of the world are particles of some sort. We shall refer to the constituents of caloric (the substance) as *caloric particles*.

As *material*, caloric exists as some quantity of matter. This means it has a mass. By caloric is "special" by virtue of being weightless. (Contrast this with "ordinary" substances, which have weight.) Because aggregates of weighted elements themselves have weight, it follows that caloric particles must also be weightless. Insofar as weight is a product of mass and gravitation force, caloric particles must either lack mass or experience zero gravitational force. But lacking mass means lacking any quantity of matter, which in turn means being immaterial. So we should think of caloric particles as not being subject to gravity. (Incidentally, the weightlessness of caloric marks it as distinct from the fire substance, which Hellenic philosophers understood as having weight.)

Just as water transitions among various states (solid, liquid, gas), caloric does too, always occupying exactly occupy one of two states.

> Free caloric, is that which is not combined in any manner with any other body. But, as we live in a system to which caloric has a very strong adhesion, it follows that we are never able to obtain it in the state of absolute freedom.
>
> Combined caloric, is that which is fixed in bodies by affinity or elective attraction, so as to form part of the substance of the body, even part of its solidity.[2]

When caloric is in its latent (combined) state, caloric particles chemically combine with "ordinary" particles to form part of the substance of a body composed of both particle kinds. By contrast, when caloric is in its free state (as *calorique sensible*), caloric particles do not combine in any manner with any other particles.

2.2 Correspondence Rules for Caloric

> Heat, considered as a sensation, or, in other words, sensible heat, is only the effect produced upon our sentient organs, by the motion or passage of caloric, disengaged from the surrounding bodies.[3]

---

[2] Antoine Lavoisier, Elements of Chemistry: In a New Systematic Order, Containing All the Modern Discoveries, translated by Robert Kerr (University of Adelaide: eBooks@Adelaide, 2014), Part 1, Chapter 1.

[3] Antoine Lavoisier, Elements of Chemistry: In a New Systematic Order, Containing All the Modern Discoveries, translated by Robert Kerr (University of Adelaide: eBooks@Adelaide, 2014), Part 1, Chapter 1.

The caloric theory enshrines a strict correspondence between caloric and heat:

> *Caloric-Heat Correspondence*: A substance's heat rises (falls) if, and only if, caloric particles accumulate in (withdraw from) the substance.

"Free" caloric, moreover, is supposed to track temperature:

> *Temperature-Free Caloric Correspondence*: Temperature is a measure of the density of free caloric.[4]

This is a Correspondence Rule for Caloric Theory, connecting the observable term *temperature* with the theoretical term *free caloric*. There is a similar Correspondence Rule for the theoretical term *latent caloric*:

> *Volume-Latent Caloric Correspondence*: Changes to the amount of latent caloric in a substance are directly proportional to changes in the volume of the substance (and do not affect the temperature of the substance).

So latent caloric, unlike free caloric, tracks volume but not temperature: it adds to the volume of a substance but does not alter the substance's temperature.

2.3 Laws for Caloric

Caloric particles, while "special" by virtue of being immune to gravity, themselves have a repulsive capacity.

> [S]trictly speaking, we are not obliged to suppose [caloric] to be a real substance; it being sufficient, as will more clearly appear in the sequel of this work, that it be considered as the repulsive cause, whatever that may be, which separates the particles of matter from each other; so that we are still at liberty to investigate its effects in an abstract and mathematical manner.[5]

This is a Theoretical Law for Caloric Theory.

> *Law of Caloric Self-Repellence*: Free caloric particles are self-repelling.

---

[4] Density is mass per volume. So positing that temperature tracks density of free caloric is yet another reason for supposing that the weightlessness of caloric particles results from their exemption from gravity rather than from their masslessness.

[5] Antoine Lavoisier, Elements of Chemistry: In a New Systematic Order, Containing All the Modern Discoveries, translated by Robert Kerr (University of Adelaide: eBooks@Adelaide, 2014), Part 1, Chapter 1, Paragraph 8.

There is a converse for the relation of caloric particles to ordinary particles, namely,

> *Law of Caloric Other-Attraction*: Free caloric particles attract nearby "ordinary" particles and shift state to become latent caloric particles.

The caloric theory maintains that the strength (magnitude) of attraction between free caloric and "ordinary" particles varies across substances. The *specific heat* (*chaleur specifique*, as coined by the Frenchman Jean Hyacinthe de Magellan in 1780) of a substance tracks this variation, being defined as the quantity of heat required for raising the temperature of a substance by one degree.[6] Early studies of heat measured specific heats using the *technique of mixtures*:

> For a body at a relatively high temperature a, and a reference body at a lower temperature b, if immersing the former into the latter yields a mixture with temperature m, the first body's specific heat $c = [(m - b)/(a - m)]*c_0$, where $c_0 = 1$ for one "unit" of water.

The notion of specific heat, incidentally, connects calorimetry to thermometry.

> For a reference body with specific heat c, if adding heat increases its temperature by amount t, then the amount of heat added is equal to ct.

This supports using thermometers as devices for indirectly measuring heat.

Because caloric is material, it is subject to the *Law of the Conservation of Mass* according to which, for any closed system (into which nothing enters and from which nothing exits), the mass of the system must remain constant over time. Because caloric is a substance, it follows that caloric itself is conserved in closed systems: caloric is neither created nor destroyed. Finally, because caloric is composed of particles, it follows that caloric particles likewise are conserved in closed systems, neither created nor destroyed. Because caloric particles are always either free or latent it follows that

> *Law of Caloric Conservation*: For any closed system, the amount of its free and latent caloric is constant.

---

[6] Only later do chemical physicists distinguish specific heat from heat capacity (the quantity of heat required to raise the temperature of one gram of a substance by one degree Celsius), and only later do they distinguish between specific heat *under constant pressure* and specific heat *under constant volume*. Incidentally, the *calorie* is the metric unit for the quantity of heat required for raising the temperature of one gram of water through one degree Celsius, and the BTU (British Thermal Unit) is the imperial unit for the quantity of heat required to raise the temperature of one pound of water through one Fahrenheit degree.

This is a Theoretical Law for Caloric Theory.

3 Explanations with "the" Caloric Theory of Heat

The laws and correspondence rules for the caloric theory support deductive, nomological (law-based) explanations of various heat phenomena.

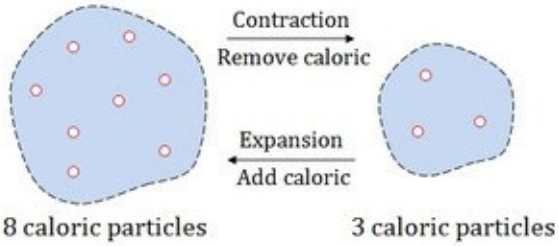
Together with the *Caloric-Heat* and *Temperature-Free Caloric Correspondences*, the *Law of Caloric Self-Repellence* entails that hot liquids cool when left to themselves.

1. Hot liquids are hotter than the surrounding air; that is, they contain more heat.
2. Hence, there is more free caloric in a hot liquid than its surrounding air (1, *Caloric-Heat Correspondence*).
3. So, over time, the free caloric particles in a hot liquid are pushed away into the surrounding air (2, *Law of Caloric Self-Repellence*).
4. When a substance loses particles of kind K, the density of K-particles in the substance decreases.
5. Hence, over time, the density of free caloric in a hot liquid decreases (3, 4).
6. Therefore, over time, the temperature of a hot liquid decreases (5, *Temperature-Free Caloric Correspondence*).

Together with the *Caloric-Heat* and *Volume-Latent Caloric Correspondences* as well as the *Law of Caloric Other-Attraction*, the *Law of Caloric Conservation* entails that heated bodies expand.

> When an ordinary body—solid or fluid—is heated, that body increases in dimensions in all directions, it occupies a larger and larger volume. If the cause of heating ceases, the body retreats through the same degrees of extension that it reached, at the same rate as it cools. Finally, if it is returned to the same temperature that it had at the first instant, it will clearly return to the same volume as it had before.[7]

[7] N.W. Best, "Lavoisier's 'Reflections on Phlogiston' II: On the Nature of Heat," *Foundations of Chemistry* 18 (2016), 4.



http://image.wikifoundry.com/image/1/AEoMU3l6jW-HsROxIso69g26160/GW300H135

1. When a body is heated, its caloric content increases (*Caloric-Heat Correspondence*).
2. When an increase of caloric content accompanies a temperature increase, the body's caloric content increases by virtue of an addition of free caloric (*Temperature-Free Caloric Correspondence*).
3. Hence, when heating a body increases its temperature, the body receives into itself free caloric (1, 2).
4. When a body receives into itself free caloric, it converts those free caloric particles into an equal number of latent caloric particles (*Law of Caloric Other-Attraction*, *Law of Caloric Conservation*).
5. Hence, when heating a body increases its temperature, the body comes to contain more latent caloric (3, 4).
6. So when heating a body increases its temperature, the body expands in volume (5, *Volume-Latent Caloric Correspondence*).

Similar reasoning entails that cooled bodies contract.

The *Laws of Caloric Conservation* and *Caloric Other-Attraction*, together with the *Volume-Latent Caloric* and *Temperature-Free Caloric Correspondences*, entail that compressing gases raises their temperatures.

1. Compressing a gas decreases its volume to some new, constant amount.
2. Hence, compressing a gas thereby "squeezes out" latent caloric into a fixed volume (1, *Volume-Latent Caloric Correspondence*).
3. When a body "squeezes out" latent caloric, the latent caloric particles convert to an equal number of free caloric particles, lest they "squeeze back in" to the body (*Law of Caloric Conservation, Law of Caloric Other-Attraction*).
4. Hence, compressing a gas increases its density of free caloric (2, 3).
5. So compressing a gas increases its temperature (4, *Temperature-Free Caloric Correspondence*).

Similar reasoning entails that decompressing gases lowers their temperatures.

All of these explanations fit Hempel's Nomological Model of scientific explanation. Insofar as good explanations also provide good predictions, these explanations also show that there is a good fit between caloric theory and some central phenomena about heat.

4 Competing Theories of Heat

Reading: Stephen Brush, "The Wave Theory of Heat: A Forgotten Stage in the Transition from the Caloric Theory to Thermodynamics," *British Journal for the History of Science* 5.18 (1970), 145-167.

Caloric theory is not the only available theory of heat from the history of science. Consider, for example, the wave theory of heat, according to which heat is the vibration of an ethereal fluid that fills all of space, transmitting vibrational motion among atoms. Here is T.S. Traill, writing about heat for the 1856 edition of *Encyclopaedia Britannica*:

> [H]eat is produced … by the vibrations or undulations of a very subtile matter existing in all bodies.

Traill is clear that the wave theory of heat is an alternative to the caloric theory. For he goes on to say that

> [T]he phenomena of caloric are owing to the movements of a subtile fluid, the particles of which are strongly repellant of each other, and have an affinity for those of all other bodies, different in force according to each kind of matter.

Here, also, is Thomas Young, in his 1907 *Lectures on Natural Philosophy*:

> [H]eat consists in a minute vibratory motion of the particles of bodies, [and] this motion is communicated through an apparent vacuum, by the undulations of an elastic medium, which is also concerned in the phenomenon of light… It is easy to imagine that such vibrations may be excited in the component parts of bodies, by percussion, by friction, or by the destruction of the equilibrium of cohesion and repulsion.

Between 1830 and 1832, the Italian physicist Macedonio Melloni had, through delicate experimental work, established that radiant heat—heat traveling through empty space in the apparent absence of matter—shares all the qualitative properties of light, such as reflection, refraction, diffraction, polarization,

interference. His result led many to *identify* heat as a form of light. Because many took the qualitative properties of light as evidence that light is a wave (rather than particles), it followed that heat was a wave as well.

The waves of wave theory differ from the caloric of caloric theory. Caloric is a material substance composed of interacting material particles; waves, by contrast, are immaterial events that propagate across material particles in material substance (ether). But, despite these ontological differences, the theories agree in positing unobservable particles. They disagree on whether these particles, or their undulatory motions, cause heat. Wave theory, but not caloric theory, posits in addition an unobservable substance, ether, that permeates through all of space. Caloric theory makes no such posit, and one reason physicists came to favor wave theory over its caloric alternative is its better fits with the wave theory of light (which posits the same ether).

As it turns out, Ampere shows that it is possible to use the wave theory to make exactly the same predictions as the caloric theory, under the following assumptions (which I provide in qualitative form):

1. The undulatory activity of a physical system is a matter of *vis viva* ("living force") associated with its constituent atoms.
2. The total *vis viva* of any closed system is constant.
3. *Vis viva* flows between groups of atoms, at a rate proportional to the difference between the *vis viva* in the groups.
4. The difference in *vivs viva* between two groups of atoms is proportional to the difference in their temperature.

Here is Ampere's central insight:

> We find manifestly the same result by considering the subject [of heat], according to the system of emission [of caloric], or according to the system of vibrations, substituting for quantity of caloric in the first system, the *vis viva* of the vibratory motions of the molecules in the second.[8]

Questions for Discussion

1- What are the major differences between caloric and wave theories of light? What are the similarities?
2- Which theory provides a *better* explanation of heat phenomena?
3- When two theories are "empirically equivalent" (make exactly the same prediction), which (if any) provides a more correct representation of reality?

---

[8] From Ampere's *Note on Heat and Light considered as the Results of Vibratory Motion*.

Case Study—Phlogiston Theory of Combustion

Nicholaos Jones

1 Ancient Fire Theories

Ancient conceptions of matter often posit four fundamental elements (or kinds of matter): air, earth, fire, and water. They also tend to posit that fire is a material that combustible bodies *release* during combustion. We can represent this posit as a chemical equation.

Ancient Theory of Combustion

$$X\ (\text{heated}) + \text{air} \Rightarrow (X - \text{fire}) + (\text{air} + \text{fire})$$

The left side of this equation uses the variable X to represent some combustible body. Values for the variable include wood, paper, and gasoline. It uses the parenthetical "(heated)" to represent that the body is undergoing combustion. It also shows the heated body "added to" air to represent that the combustion is occurring in air.

The right side of the equation uses "subtraction" to represent fire leaving the combustible body X, and it uses "addition" to represent this same fire entering the air. This is meant to capture the theorized outcome of combustion: the heated body releases fire, and the released fire enters the air.

The main virtue of the ancient theory of combustion is that it easily explains why heated bodies become inflamed: prior to heating, there is fire trapped in bodies, and after heating the bodies become inflamed by virtue of releasing this fire.

The main vice of the ancient theory of combustion is that combustible bodies, prior to being heated, need not be warm. Because the ancient theory of combustion entails that combustible bodies are laden with fire, and because fire is always hot, the ancient theory seems to predict that combustible bodies should always have some warmth to them.

2 Stahl's Phlogiston Theory

There are ways to reconcile the warmth of fire and the lack of warmth in some combustible (but unheated) bodies. For example, one might posit that the fire latent in combustible (and unheated) bodies has its warmth somehow obstructed by the earth and water in those same bodies. Or one might reject the

natural assumption that fire is always hot, theorizing instead that fire is *potentially hot* when mixed with earth or water and that fire actualizes this potential only when liberated from these elements.

George Ernst Stahl is known for developing an alternative to the ancient theory of combustion. Stahl's theory posits that combustible bodies *release phlogiston* during combustion.

Stahl's Theory of Combustion

X (heated) + air ⇒ (X – phlogiston) + (air + phlogiston)

Phlogiston, for Stahl, is an inflammable material that, when mixed with air, generates fire. Stahl is led to posit the existence of phlogiston from the observable facts that some substances burn and inflame.

Because Stahl's theory entails that heating a body releases phlogiston from the body into the air, Stahl's theory fits with ancient conceptions of fire as a substance in bodies. But Stahl's theory avoids the main vice of those conceptions. For Stahl, the error in the ancient conceptions is supposing that combustible bodies release *fire* during combustion. Stahl avoids this error by positing that combustible bodies release *phlogiston*. Because phlogiston itself is not hot, Stahl's theory does not entail that combustible bodies should always be warm (even when not heated). And so Stahl's theory does not need to explain why some combustible bodies are not warm.

3 Boyle's Phlogiston Theory

Because Stahl endorses the natural assumption that all matter has mass, he is committed to conceptualizing phlogiston as a substance that has positive weight. But if phlogiston has positive weight, Stahl's theory entails that combustible bodies should *lose* weight when burned.

Stahl's Theory of Combustion (with positive-weight phlogiston)

X (heated) + air ⇒ (X – phlogiston^{+}) + (air + phlogiston^{+})

Robert Boyle makes an observation that is difficult to reconcile with this consequence of Stahl's theory. The observation is that combustible bodies *gain* weight when burned. (This is especially evident with metals, sulfur, and phosphorus.)

Although Boyle's observation seems to refute Stahl's theory, Boyle does not use his observation to infer that Stahl's theory is false. Instead, Boyle infers from his observation that phlogiston has *negative* weight.

Boyle's Theory of Combustion (with negative-weight phlogiston)

$$X\ (\text{heated}) + \text{air} \Rightarrow (X - \text{phlogiston}^{-}) + (\text{air} + \text{phlogiston}^{-})$$

If phlogiston has negative weight, then Stahl's theory *correctly* entails that combustible bodies *gain* weight when burned (because losing negative weight is equivalent to gaining weight). Boyle's theory of combustion thereby rescues the core posit of Stahl's theory by modifying an auxiliary assumption association with the theory.

4 Priestley's Phlogiston Theory

Stahl's theory of combustion dates to 1703. This dating is important, because it means that there is pressure to make revisions of Stahl's theory consistent with Isaac Newton's theory of matter.

Newton publishes *Principia Mathematica* in 1687, and one of the central results of Newton's work is that matter is gravitational. Newton establishes that the mass of a body is directly proportional to the gravitational force it exerts upon other bodies. This means, for example, that more massive bodies exert stronger gravitational forces. It means that bodies without mass (if any exist) exert no gravitational force. It also means that bodies with *negative* mass exert *negative* gravitational force—they are *anti-gravitational*.

Because Newton's theory of matter entails that bodies with negative mass are anti-gravitational, conjoining Newton's theory with Boyle's phlogiston theory entails that phlogiston is an anti-gravitational material. The problem with this entailment is that combustible bodies (prior to heating) do not seem to have anti-gravitational properties.

Joseph Priestly takes Newton's theory of matter to justify preserving the natural assumption that phlogiston has positive weight. This assumption makes Stahl's

phlogiston theory of combustion is inconsistent with Boyle's observations (about the weight of burned bodies). But rather than reject Stahl's theory as false, Priestley infers that Stahl's theory is incomplete. According to Priestley, what Stahl failed to notice is that heating combustible bodies not only *releases phlogiston* from those bodies but also *adds* to the bodies matter from the surrounding air.

Priestley's (Incomplete) Phlogiston Theory

$$X\ (\text{heated}) + (\text{air} + Y) \Rightarrow (X - \text{phlogiston} + Y) + (\text{air} + \text{phlogiston} - Y)$$

According to Priestly, the phlogiston that combustible bodies release into the air during combustion displaces some substance from the air into the bodies. Because this substance is unknown to Priestley, the resulting phlogiston theory of combustion is incomplete. (The variable Y represents Priestley's unknown substance.) Even so, provided that the released phlogiston weights less than the substance it displaces from the air, Priestley's theory correctly entails that combustible bodies *gain* weight when burned.

5 Priestley's Experiments

Priestley's phlogiston theory of combustion reconciles phlogiston theory with Newton's theory of matter, and it accommodates Boyle's observations about the relative weight of burned and unburned bodies. But these virtues are conditional. The success of Priestley's theory depends upon his assumption that the substance that phlogiston displaces from air weighs more than the phlogiston that displaces it. Because this assumption is difficult to evaluate without specifying exactly what the displaced substance is, Priestley's theory motivates research into identifying the components of air.[1]

In 1774, Priestley conducts his most famous experiment for identifying the components of air. Priestley begins by burning mercury. Priestley infers, from his phlogiston theory of combustion, that the reside that remains after burning mercury is "calx of mercury," where "calx of mercury" is mercurial substance (red powder) that is emptied of phlogiston but laden with some unknown substance from the air.

[1] This research program for identifying the components of air is closely connected to a concurrent research program that was aiming to identify the components of water—and the attendant debate about whether water even has components.

mercury (heated) + (air + Y) ⇒ "calx of mercury" + (air + phlogiston – Y),

where "calx of mercury" = (mercury – phlogiston + Y)

Priestley places his "calx of mercury" in an inverted glass container that is itself placed in a pool of mercury. After using a lens to focus sunlight on the "calx of mercury," Priestley captures the gas that remains in the glass container. He names this gas "dephlogisticated air."

Priestley performs a series of tests on his dephlogisticated air. When he inhales it, he observes that it is "five or six times as good as common air" because it makes his chest feel "peculiarly light and easy for some time afterwards." When he compares the lifespans of mice who breathe dephlogisticated air and mice who breathe common air, he observes that the mice who breathe dephlogisticated air live about four times longer. When he strikes a match in dephlogisticated air, he observes that the flame burns more intensely than flames burning in common air.

Because his dephlogisticated air has remarkable effects, Priestley infers that the unknown substance in his phlogiston theory of combustion—the substance represented by the variable Y—is "pure air." Priestley contrasts this "pure air" with "common air." "Pure air" combines with metals and supports respiration. "Common air" is an asphyxiant (substance that causes suffocation) that supports neither combustion nor respiration.

Priestley's (Completed) Phlogiston Theory

X (heated) + (common air + pure air) ⇒

(X – phlogiston + pure air) + (common air + phlogiston – pure air)

The result is a complete phlogiston theory of combustion that coheres with Newton's theory of matter and accommodates Boyle's observations about the weight of burned bodies.

6 Lavoisier's Oxygen Theory

Inspired by Priestley's experiments on "dephlogisticated air," Antoine-Laurent Lavoisier develops a competing theory of combustion. This competing theory arises from Lavoisier's efforts to develop a theory of chemical reactions. The theory of combustion posits that combustible bodies *combine with oxygen* when burned.

Lavoisier's Oxygen Theory

X (heated) + (oxygen + inflammable air) ⇒

(X + oxygen) + (inflammable air – oxygen)

Like Priestley, Lavoisier divides air into two kinds. He identifies Priestley's "pure air" with oxygen, and je identifies Priestley's "common air" with "inflammable air." Unlike Priestley, Lavoisier denies that there is any such substance as phlogiston. Phlogiston, as Lavoisier sees it, is entirely imaginary.

Lavoisier's oxygen theory of combustion is, in effect, the result of deleting from Priestley's theory any mention of phlogiston. For this reason, Lavoisier's theory has the advantage of simplicity over Priestley's theory. It also conveniently explains an experimental result from Henry Cavendish.

Cavendish's experiment involves enclosing Priestley's "common air" and "pure air" in a closed vessel. Cavendish observes that, when the air inside the vessel is ignited with a spark, a small amount of "dew" forms on the inner walls of the vessel. Through repeated trials, Cavendish discerns that this "dew" is water.

Because Cavendish endorses Priestley's phlogiston theory, he infers that the water must have been present in the "common air" and "pure air" prior to ignition with the spark. Lavoisier has a much simpler explanation: the spark causes the two "airs" to combine, and the combination is water. Lavoisier subsequently confirms his explanation by decomposing water into oxygen and "inflammable air."[2] The outcome is that Lavoisier's oxygen theory of combustion triumphs over Priestley's phlogiston theory.

[2] Lavosier's explanation is obvious once one identifies his "inflammable air" as hydrogen. Lavosier's insight, in effect, is that water is a compound of hydrogen and oxygen—namely, H_2O.

Case Study—The Chemical Nature of Water

Nicholaos Jones

Reading: P.K. Basu, "Theory-ladenness of Evidence: A Case Study from the History of Chemistry," *Studies in History and Philosophy of Science* 34 (2003): 351-368.

The Debate

Antoine Lavoisier, the leading French chemist of his era, claimed that water is a chemical compound by arguing for three separate hypotheses. The first hypothesis is that hydrogen is a constituent of water, because iron reacts with steam to produce iron oxide and hydrogen. The second hypothesis is that the explanation of iron oxide being produced in the iron-steam reaction is that the oxygen must have come from the water, because iron reacts with oxygen to form iron oxide and the product of the iron-oxygen reaction is chemically similar to the solid product in an iron-steam reaction. The third hypothesis is that water can be formed from oxygen and hydrogen by performing the combustion of hydrogen in oxygen. Given the methodological assumption that the chemical constituents of a substance are established when the substance has been divided into chemically distinct parts and formed by combining those parts, Lavosier infers that water is a chemical compound of hydrogen and oxygen.[1]

Joseph Priestley objected that the black powder produced in the reaction of iron and steam is not iron oxide, contrary to Lavoisier's first hypothesis. Priestley agrees that the reaction produces black powder, but denies that the powder is iron oxide. Priestley also argued that iron is not a simple substance, because in the iron-oxygen reaction both iron oxide and a gas (fixed air-carbon dioxide) are produced.

The Iron-Oxygen Reaction

Lavoisier supports his second hypothesis with two claims:

C_1: Whenever iron is heated in oxygen, iron oxide is produced;

C_2: Nothing else is produced in the iron–oxygen reaction.

[1] Incidentally, Lavosier's theory of water is also an example of ontological reduction driven by explanatory reduction: Lavosier *reduces* water to hydrogen and oxygen (thesis: water *is nothing more than* oxygen and hydrogen), and he justifies this reduction by using oxygen and hydrogen to explain some chemical properties of water.

Both claims, C_1 and C_2, employ theoretical language, including terms such as 'iron', 'oxygen', 'being heated', 'iron oxide'.

The evidence for the hypotheses, C_1 and C_2, are claims of the sort:

E_1: Iron oxide is produced;

E_2: On the occasion when iron oxide is produced, nothing else is produced.

Thus, apparently, E_1 confirms C_1, E_2 confirms C_2.

The Raw Data

The set of (raw) data for the iron–oxygen reaction acceptable to both the Lavoisians and Priestley is:

1. A *pure* sample of *iron* was employed.
2. A *pure* sample of *oxygen* was employed.
3. Balance reading of the weight measurement of the pure sample of iron: M_1.
4. Balance reading of the weight measurement of the pure sample of oxygen: M_2.
5. A black powder is formed in the vessel. (Call this claim r_1.)
6. Balance reading of the weight measurement of the black powder: M_3.

Note: the evidence, E_1, is not expressed by the sentence, 'there is a black powder in the vessel' or 'a black powder is formed in the vessel'.

Let r_1 be the sentence 'There is a black powder in the vessel'.

The sentence r_1 is the (raw) data as expressed in a "theory neutral" language (neutral with respect to Lavoisier and Priestley, at least). The other (raw) data include the balance reading of the weight measurement of black powder and the balance readings of the weight measurements of the samples of iron and of oxygen. These data do not constitute any evidence by themselves. One constructs evidence out of these raw data. How this evidence is constructed depends on what theories or background assumptions are employed.

Let r_2 be the sentence 'The balance reading of the weight measurement of the iron sample is M_1.'.

Let r_3 be, 'The balance reading of the weight measurement of the oxygen sample is M_2'.

Let r_4 be the sentence, 'The balance reading of the weight measurement of the black powder is M_3'.

The four statements r_1, ..., r_4 taken jointly do not yet form the evidence for Lavoisier's hypothesis. These statements about (raw) data are not, *prima facie*, evidence for or against Lavoisier's hypothesis. If they were the evidence for Lavoisier's hypothesis, Priestley could not have challenged Lavoisier's theory. This is because Priestley accepted these statements about the (raw) data. If they were uncontroversial as evidence, Priestley would have to accept the way the evidence was constructed from the (raw) data. Once the evidence was constructed in the way mentioned above, it would be hard for Priestley to reject the Lavoisian argument.

Priestley did not accept these statements as evidence for Lavoisier's hypothesis. What he objected to was the employment of the data to construct evidence for Lavoisier's theory.

Background (Theoretical) Assumptions

The data become evidence in light of (at the time controversial) theoretical principles. Consider, for example,

The Principle of Conservation of Mass (PCM): Mass is conserved in any chemical reaction.

PCM entails that the mass of the black powder must be equal to the mass of iron and of oxygen if and only if nothing else is produced in the reaction. Lavoisier computed and established the equality of the masses, i.e. $M_B=M_J+M_O$. That equality constituted evidence for Lavoisier for the claim that only the black powder is produced and nothing else.

The plausibility of PCM rests on a specific kind of theory of matter. Priestly rejects this theory. And so he could have rejected this construction. However, Priestly did not reject Lavoisier's construction for this reason. For, in order for a comparison of theories to be as fair as possible, one needs to grant one's opponents the premises the opponents require for their argument.

Consider again the (raw) data expressed in the claim r_1: 'A black powder was formed in the vessel'. *Prima facie*, r_1 is not an evidence for C_1 or C_2.

The claim r_1 can be transformed into E_1 by making use of some auxiliary assumptions. These auxiliary assumptions include the general Stahlian theses concerning determining the chemical nature and the specific reactivities of iron oxide. One such thesis, call it *S*, has the form: if the black powder undergoes specific reactions (characteristic of iron oxide), then it is iron oxide.

The data claim e_1 and the auxiliary assumption *S* imply E_1. The claim E_1 is also arrived at from H_1 by means of a different set of auxiliary assumptions, namely the assumption A_1 that both the iron and oxygen samples were pure, and the assumption A_2 that the iron was heated in oxygen. The evidence statement, E_1, is thus constructed in two different ways and these constructions are not interchangeable. Thus,

(1) e_1 and *S* imply E_1;

(2) H_1 and A_1 and A_2 imply E_1.

However, H_1 and A_1 and A_2 and *S* do not imply e_1; and e_1 and A_1 and A_2 and *S* do not imply H_1. This is because the construction of E_1 in (1) is asymmetrical. The fact that iron oxide is produced does not entail (along with *S*) that a black powder is produced.

This asymmetry gives us some idea regarding how classification of objects into different kinds can and does proceed, as follows. The (raw) data about an object to be classified is employed along with some auxiliary assumptions to determine if the object falls under a certain class. It should be obvious that terms like 'iron' and 'oxygen' are classificatory terms that are, at least one level removed from the terms employed to express the (raw) data. It should also be apparent that what counts as agreed upon (raw) data and what counts as disputed evidence has a multi-layered nature. That a solid metallic sample was iron and was pure, could only be determined (in the eighteenth century) by subjecting the sample to Stahlian tests. Thus, strictly speaking, the existence of a solid metallic body was the (raw) data that both the Lavoisians and Priestley could agree upon. However, they also agreed that this solid metallic body was a sample of pure iron. This represented thus a first level of evidence construction. Because of the agreement on all sides, this premise—that one was using a sample of pure iron—could be considered (uncontroversial) (raw) data at that time.

Scientists can and sometimes do disagree whether a certain sentence expressing a raw datum leads to a sentence expressing evidence. Their disagreement might be basic, at the data level, but may also be at the level of auxiliary assumptions.

The Lavoisians employed a set of assumptions, A_3, about the nature of matter (derived from a Newtonian view of matter), about the identity of the inertial mass and gravitational mass of any object, and about the balance being a reliable instrument to determine the gravitational mass of an object, in order to construct first order evidence from the balance readings. Thus,

> r_2 and A_3 imply SM_J (the mass of the particular sample of pure iron is M_J);
>
> r_3 and A_3 imply SM_O (the mass of the particular sample of pure oxygen is M_O);
>
> r_4 and A_3 imply SM_B (the mass of the black powder produced in the reaction is M_B).

Priestley could have objected to this construction of evidence from the (raw) data on the grounds that he did not accept the set of assumptions (A_3). However, he did not.