

Introduction to History and Philosophy of Science

Introduction to History and Philosophy of Science

Hakob Barseghyan, Nicholas Overgaard, and Gregory Rupik



Introduction to History and Philosophy of Science by Hakob Barseghyan, Nicholas Overgaard, and Gregory Rupik is licensed under a Creative Commons Attribution 4.0 International License, except where otherwise noted.

Contents

Introduction to History and Philosophy of Science	vii
Table Of Contents	ix
1. Introduction -- Introduction to History and Philosophy of Science	1
2. Absolute Knowledge -- Introduction to History and Philosophy of Science	11
3. Scientific Method -- Introduction to History and Philosophy of Science	31
4. The Laws of Scientific Change -- Introduction to History and Philosophy of Science	49
5. Scientific Progress -- Introduction to History and Philosophy of Science	69
6. Science and Non-Science -- Introduction to History and Philosophy of Science	87
7. Aristotelian-Medieval Worldview -- Introduction to History and Philosophy of Science	99
8. Cartesian Worldview -- Introduction to History and Philosophy of Science	119
9. Newtonian Worldview -- Introduction to History and Philosophy of Science	135
10. Contemporary Worldview -- Introduction to History and Philosophy of Science	151
11. Worldviews: Metaphysical Components -- Introduction to History and Philosophy of Science	171

Introduction to History and Philosophy of Science

Hakob Barseghyan, Nicholas Overgaard, and Gregory Rupik



Introduction to History and Philosophy of Science by Hakob Barseghyan, Nicholas Overgaard, and Gregory Rupik is licensed under a Creative Commons Attribution 4.0 International License, except where otherwise noted.

This book was produced using Pressbooks.com.

Table Of Contents

Contents

- Part I. Chapter 1: Introduction
- Part II. Chapter 2: Absolute Knowledge
- Part III. Chapter 3: Scientific Method
- Part IV. Chapter 4: The Laws of Scientific Change
- Part V. Chapter 5: Scientific Progress
- Part VI. Chapter 6: Science and Non-Science
- Part VII. Chapter 7: Aristotelian-Medieval Worldview
- Part VIII. Chapter 8: Cartesian Worldview
- Part IX. Chapter 9: Newtonian Worldview
- Part X. Chapter 10: Contemporary Worldview
- Part XI. Chapter 11: Worldviews: Metaphysical Components

1.

Introduction -- Introduction to History and Philosophy of Science

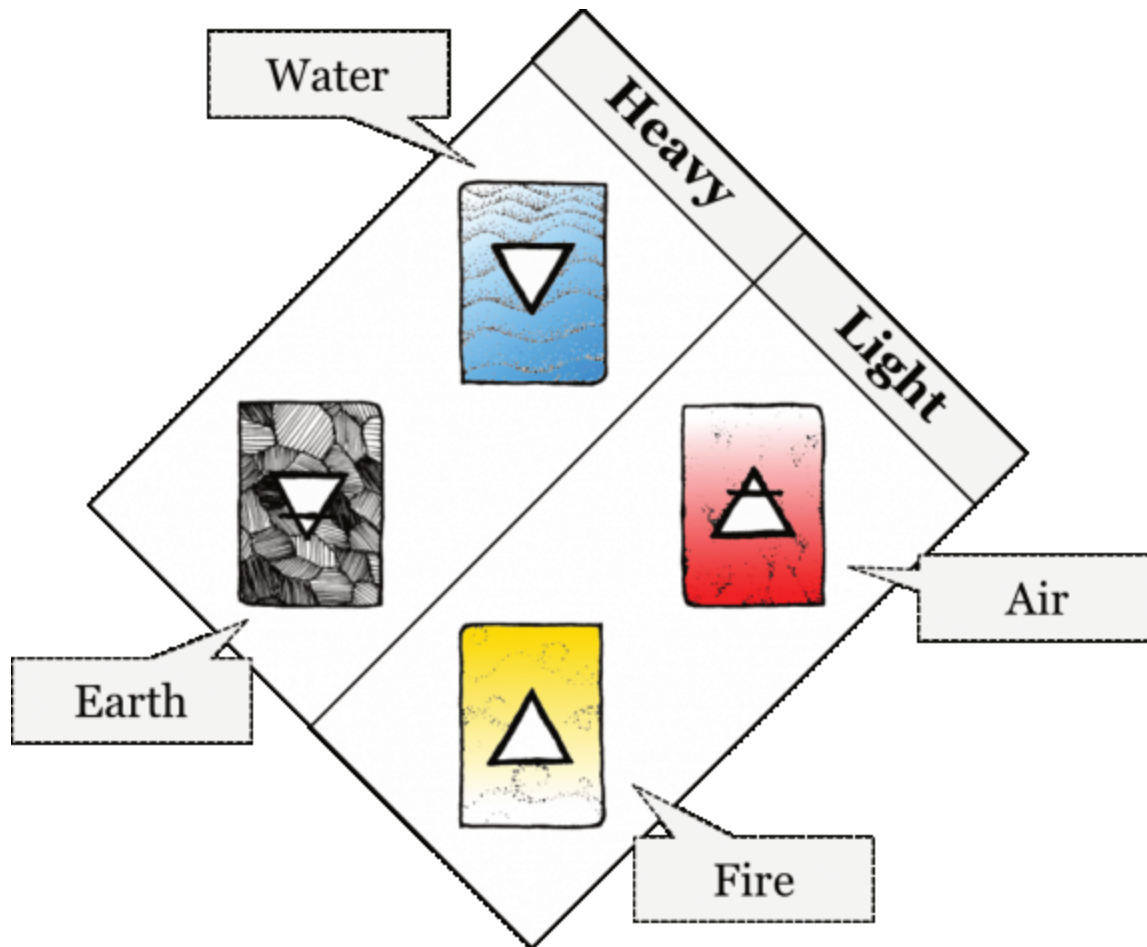
Chapter 1: Introduction

Imagine that you are not reading this textbook. Imagine instead that you are lying on your back in some soft grass on a warm summer night, far from city lights, staring into the vast, dark night sky. As you continue to gaze at the stars, you would likely notice that over the course of hours they all slowly move – in unison – in the same direction. From the Northern Hemisphere, you will always see the constellation *Canis Major* near *Orion*, or the constellation of the Celestial Bear flanked by its seven hunters, but all of them will seem to rotate around *Polaris*, the North Star. If you are incredibly perceptive, however, you may notice that not *all* points of light in the night sky move together. Some of them follow their own path, wandering through the sky with the stars as a backdrop. The ancient Greeks called them *asteres planetai*, meaning *wandering stars*, which is where we get the word *planet* from. If you were to carefully track the path of a planet over the course of a few nights, you would realize that – even though its movement is different from that of the stars – it is far from random. It follows a certain path through the night sky. Indeed, while different planets follow different paths, you could begin to notice similarities between the motions of all the planets as they wander through the heavens. Observed from the Earth, they all appear to move in an eastward direction, and their paths are roughly on the same plane.

But why? What kind of explanation could we give for why planets' paths differ from those of the stars? Why do planets seem to behave in very similar ways to one another? What are the best scientific theories we have to explain planetary motion?

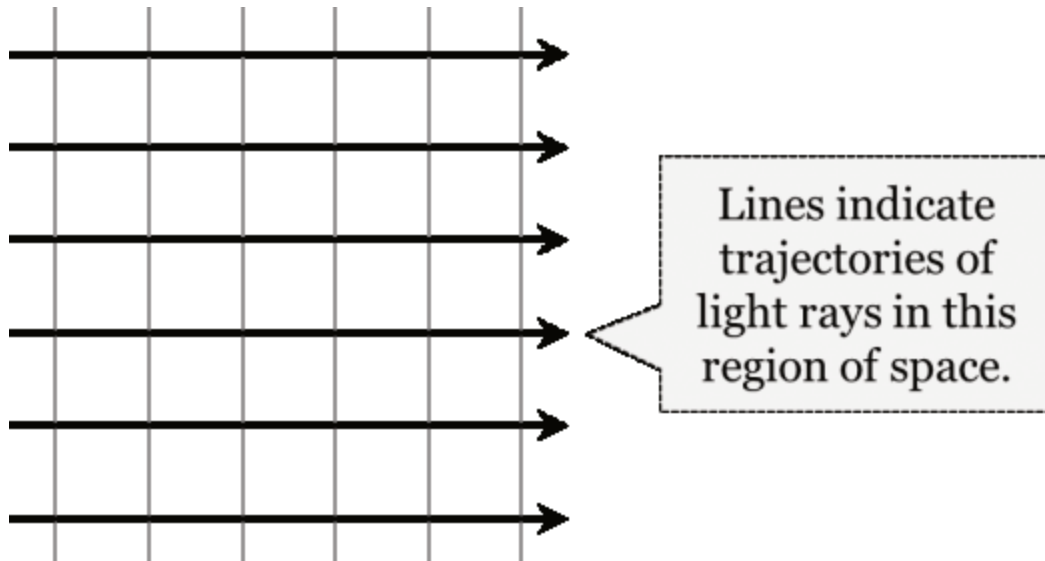
Let's try a familiar explanation. Those planets are actually no different from the Earth: they are large massive objects, all orbiting around a much more massive object in the centre of our solar system – the Sun. Isaac Newton showed in his law of universal gravitation that the very same force which pulls an apple to the ground, and which causes the parabolic paths of projectiles, also causes planets and moons to take the precise paths they do through space. The speed of the planets and the force of gravity keep planets like the Earth and Mars in orbit around the Sun. From the vantage point of Earth, therefore, planets seem to wander through the night sky because they are following their own, elliptical paths around our nearby Sun. Meanwhile, the constellations and positions of the stars remain relatively fixed because they are so far away from the solar system, and they rotate together due to the rotation of the Earth on its axis. This is the answer you would receive if you were able to travel back to the year 1800 and ask a member of the scientific community at the Royal Society in London, England to give you their *best, agreed-upon* scientific theories about planetary motion.

But what if we were to travel even further back in time, say 500 years? What accepted theories would an astronomer from the University of Paris in the year 1500 use to explain the wandering of the planets? A late-medieval astronomer would explain planetary motion by referencing Aristotelian natural philosophy. This set of theories accounted for the motion of objects by considering the movements that are *natural* to different elements. It was believed at the time that the universe is made of two completely distinct regions – terrestrial and celestial. Everything in the *terrestrial region* was thought to be composed of a certain combination of the four terrestrial elements – *earth, water, air, and fire*. The elements earth and water were believed to be *heavy*, while the elements air and fire were believed to be *light*.

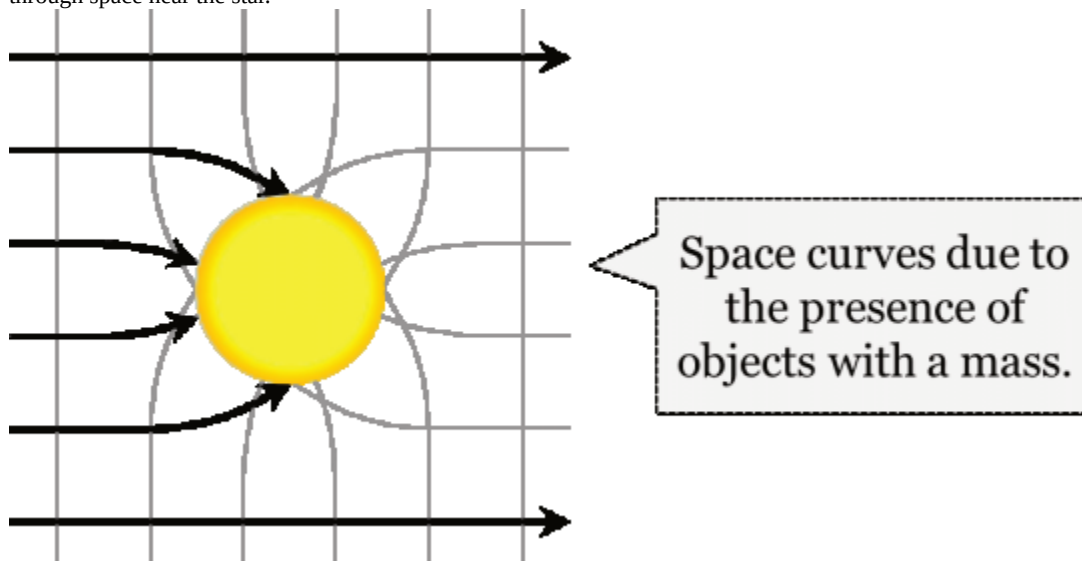


Each of the four elements was thought to have a natural position to which it is predisposed. For heavy elements, the natural position is the centre of the universe, which explained why everything made of elements earth and water has a tendency to fall down. This is why, they would say, when you drop a rock it goes straight down. This would also suggest that the terrestrial globe, which is predominantly a combination of the elements earth and water, should necessarily be at the centre of the universe. In the *celestial region*, in contrast, everything, including the planets and the stars, was believed to be made of a completely different element, *aether*. The natural tendency of aether is to revolve in a circular path around the centre of the universe. The planets, being between the stationary sphere of the Earth and the slowly-rotating stars, naturally follow their own circular paths through the night sky, accounting for their apparent “wandering” in front of the distant stars.

Tired of all this hypothetical time travel, let’s say you made an actual voyage to the Mauna Kea Observatories in Hawaii, USA, and – after a relaxing day at the beach – asked a modern-day astronomer to explain planetary motion using the best, agreed-upon scientific theories. The astronomer would not give you the Aristotelian-Medieval answer, *nor would they give you the Newtonian answer* you may be familiar with from basic physics or astronomy classes. The accepted view today is that the paths of the planets, like the Earth, are best explained by Albert Einstein’s theory of general relativity, *not* Newton’s law of universal gravitation. Today, the elliptical paths of planets around the sun are not taken to be due to a force called gravity but are rather due to the fact that the mass of our Sun bends the fabric of space-time itself. Imagine a region of space-time without any material objects. Such a region would be completely flat. What this means is that in such a space, light rays would travel along straight lines, and the geometry we learned in secondary school, Euclidean geometry, will hold exactly.

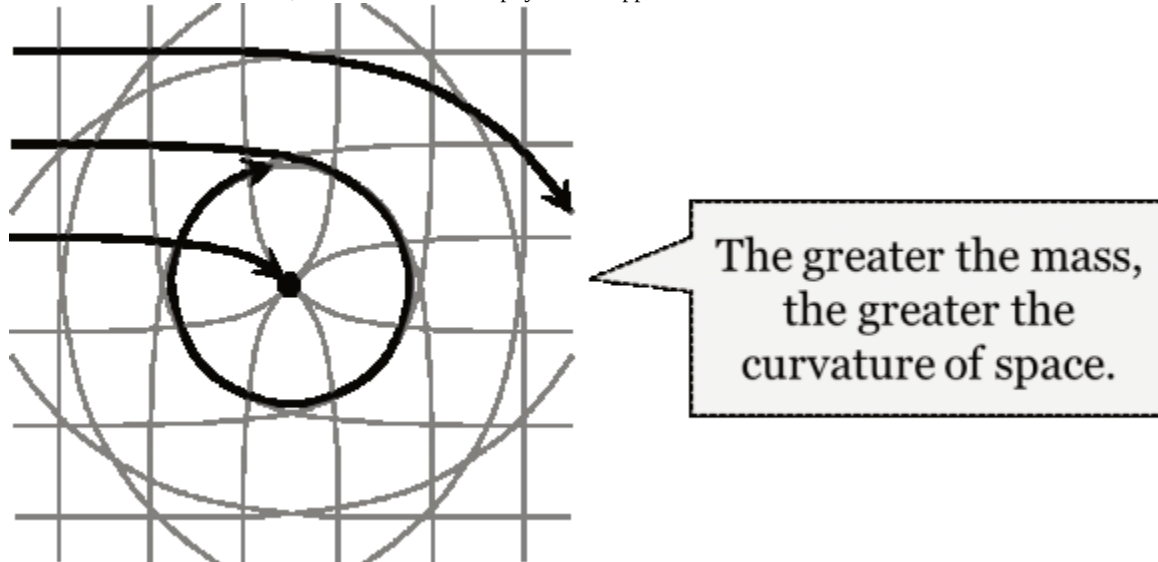


Now, let's add a star to this region of space. According to general relativity, this star will curve the space-time around it, affecting the motion of all other material processes in its vicinity, including light rays. The space will no longer be exactly describable by Euclid's geometry, but rather by a geometry developed by the German mathematician Herman Minkowski and incorporated by Einstein into his theory. This geometry treats time as a fourth dimension, perpendicular to the familiar three dimensions of length, width, and breadth, which is why we speak of *space-time*. Even physicists can't really picture all this. They can represent the situation using mathematical equations and make predictions by solving them. They understand these mathematical models by using analogies that involve fewer dimensions. As an example of such an analogy, let's imagine a stretched bedsheet with a basketball placed in the middle of it. The basketball will make a dip in the bedsheet. The two-dimensional bedsheet represents four-dimensional space-time. The dip in the bedsheet in the third dimension produced by the ball represents the curvature of four-dimensional space-time produced by an object with mass, like a star. Now, let's roll a tennis ball across the bedsheet. Because the fabric of the bedsheet is curved by the basketball, the tennis ball will not move in a straight line, but rather will have a curved trajectory along the bedsheet. It will appear as though the tennis ball is attracted by the basketball, while in fact it is merely following the curvature of the bedsheet. According to general relativity, something like this happens when a region of space-time is curved by a massive object, such as a star, with the tennis ball being something like a planet moving on a curved trajectory through space near the star.

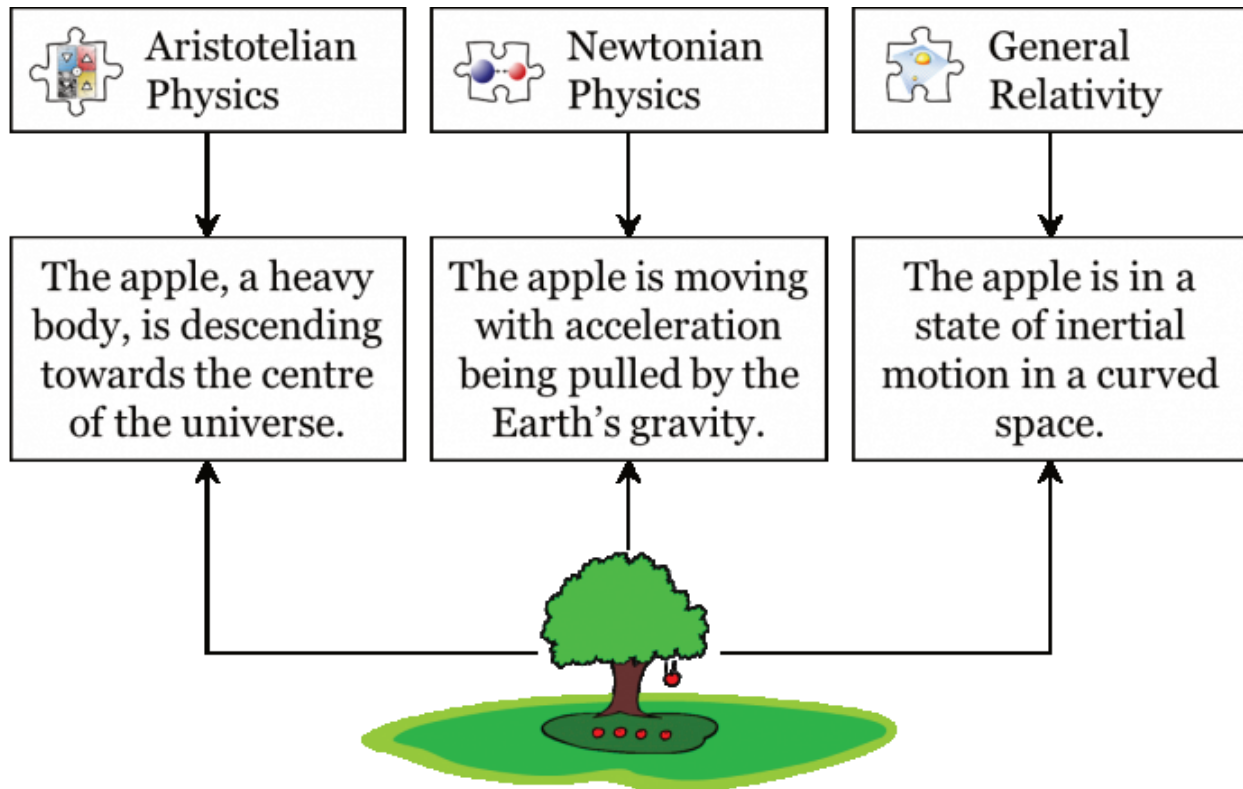


The same goes for any object with mass. The reason the Moon or a spacecraft continues to revolve around the Earth is that the Earth, as a massive object, bends the space-time around it to capture the Moon in a sort of dip in space-time. Similarly, the Sun, a significantly more massive object, bends a larger region of space-time than the Earth and captures the Earth, the planets, and many other celestial objects in its larger dip in space. The degree of curvature of space-time around an object depends on both the mass of an object and on how compressed that mass is into a small region of space. If any object is compressed within its

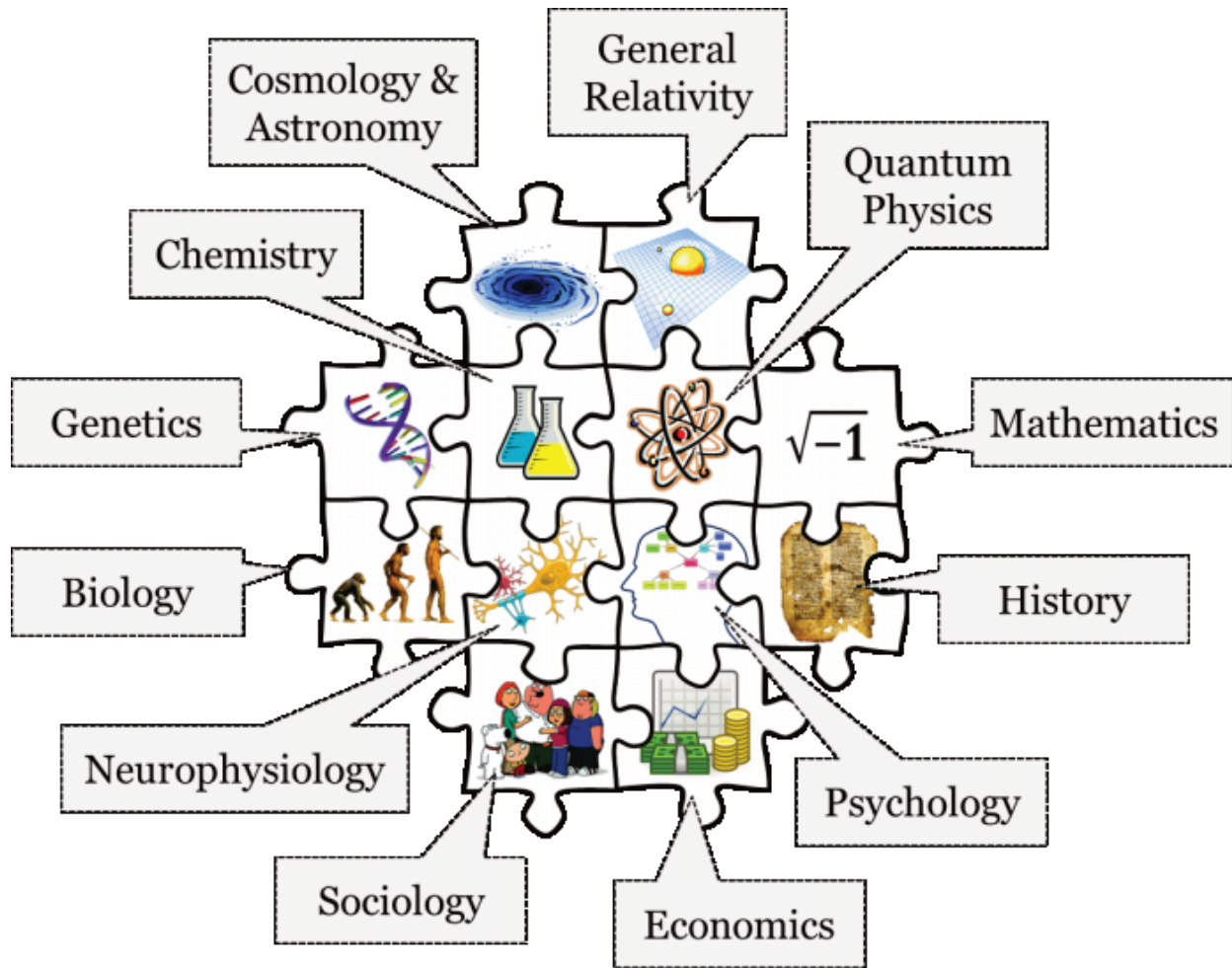
Schwarzschild radius, named for the German physicist Karl Schwarzschild, the curvature of space will become so great that even light rays cannot escape it. It will become a black hole. The Schwarzschild radius of the Earth is 8.7 millimetres. If the Earth were somehow compressed to this tiny radius, it would become a black hole. Although in the 2009 *Star Trek* movie, malevolent aliens destroyed the planet Vulcan this way, astrophysicists don't know of any natural process that would crush a planet to such densities. The only known natural process that can crush an object to within its Schwarzschild radius is the collapse of a massive star that has exhausted its nuclear fuel, and this is how astrophysicists suppose black holes form.



What this means is that the seeming attraction between two material objects is nothing but an inertial motion in a curved space. According to general relativity, there is no force of gravity; all material objects curve the space around them to a greater or lesser degree and this curvature affects the motion of other material objects which happen to be in the vicinity, just like the motion of the tennis ball on a suspended bedsheet. Note that physicists still use the word "gravity" as a shortcut for "motion in a curved space-time". Yet, strictly speaking, in general relativity the seeming gravitational attraction among material objects is understood as merely a motion in a space-time which is not flat but curved by material objects. A planet's elliptical orbit, therefore, is not because of gravity, but because of a combination of that planet's own momentum and the shape of the space-time bent by the Sun's mass.



In short, the best astronomical theories from different historical periods explained the motion of the planets in very different ways. But the scientific communities in any one of these periods didn't just agree on astronomy! In addition to astronomical and physical theories, the scientific communities of each of these periods also accepted a variety of theories on different natural, social, formal, and artificial objects. Considered together, these individual theories from any one historical moment can be shown to make up a complex, interwoven tapestry of theories, constituting that historical community's best available description of the world. Take, for instance, the theories that we accept nowadays. If we were to ask a scientist what theories best describe the world, they would probably mention several theories from *natural science*, such as general relativity, quantum physics, big bang cosmology, contemporary chemistry, the modern evolutionary synthesis in biology, etc. They would also likely mention some theories from *social science*, such as those from psychology, economics, or sociology. In addition, they would probably mention a few theories that pertain to *formal science*, including mathematics and logic. Here is a snapshot of some of the theories accepted these days:



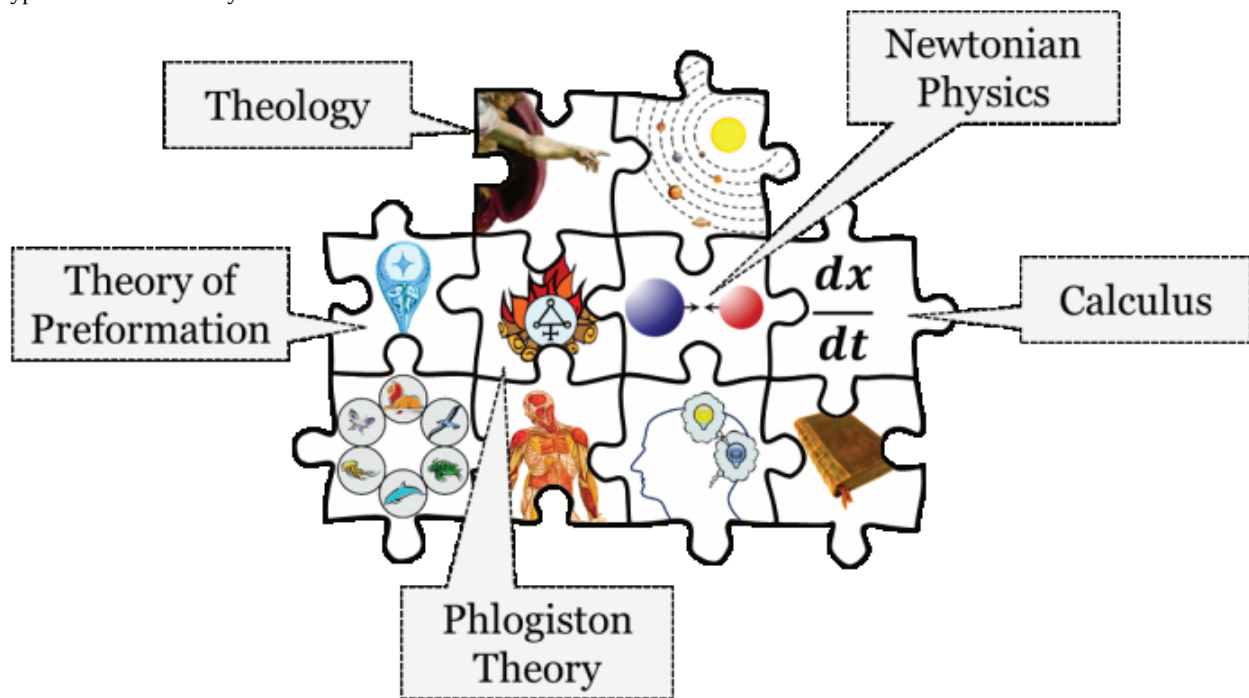
This interlocking jigsaw puzzle represents many of this community's best available descriptions of the world – their accepted *theories*. Taken together, we call this complete set of a community's accepted theories its *mosaic*.

Scientific Mosaic \equiv
 A set of all accepted *theories*.

Theory \equiv
 A set of propositions that attempt to describe something.

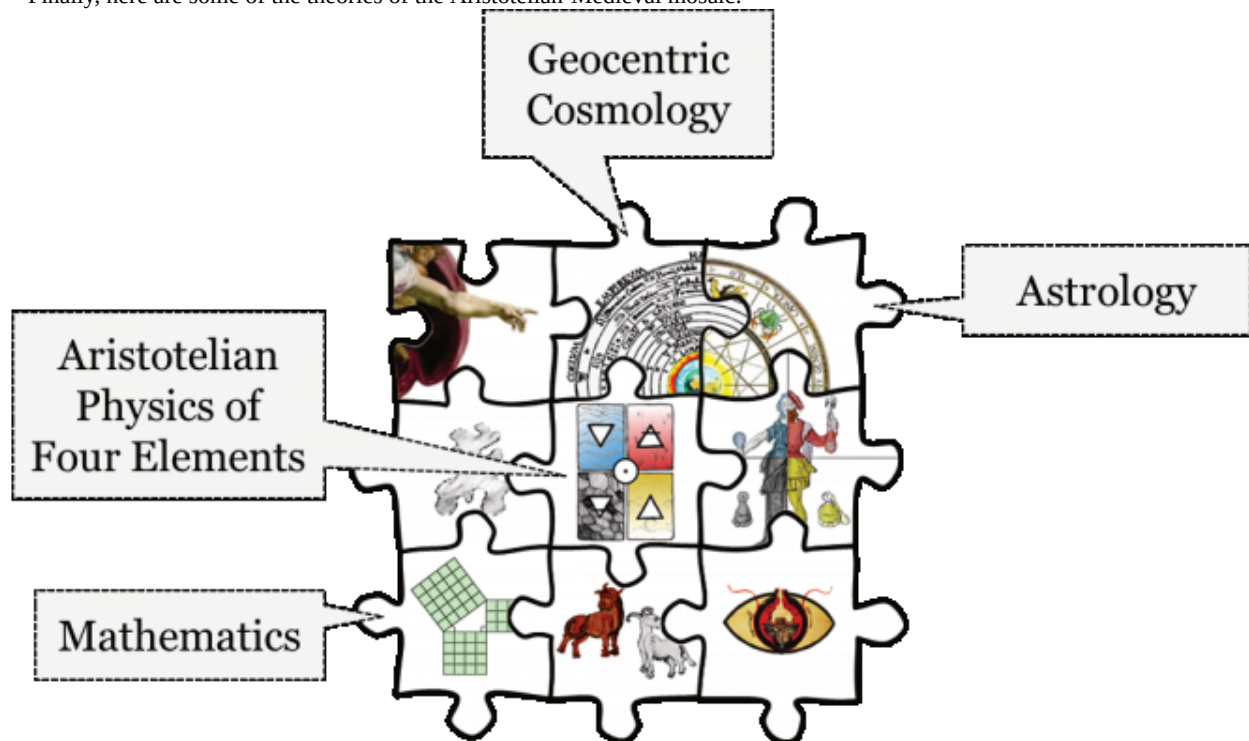
Note:
 This is the sign of *aggregation*, i.e. a mosaic consists of theories.

Mosaics *change* through time as scientific communities accept new theories and reject old ones. Here is a quick snapshot of a typical mid-18th-century Newtonian mosaic:



Among other things, this mosaic included Newtonian physics with its three laws of mechanics and the law of gravity, the chemical theory of phlogiston, and even theology, the study of God and his works.

Finally, here are some of the theories of the Aristotelian-Medieval mosaic:



Notice the presence of Aristotelian natural philosophy with its theory of four terrestrial and one celestial element, theology, and, interestingly, astrology, the study of celestial influences on terrestrial affairs.

Mosaics and changes in them are the central focus of this textbook. For any point in history, a community's mosaic showcases its best attempt to understand reality, in all its dynamism and complexity. This drive to understand reality is a human one, which

means that individuals – in their social and institutional contexts – are important parts of the story of how certain theories come to be, and how those theories come to change. For instance, to better understand the theories of planetary motion we sketched earlier, we could trace the histories of institutions like the Royal Society in England, the University of Paris, or the Mauna Kea Observatories in the U.S.A., or of individual investigators at any one of those institutions. One can certainly come to appreciate science’s history by approaching it this way! But a different kind of appreciation for science can be gained by taking a step back and seeing these astronomical theories as one facet of the total set of the best theories accepted by the scientific community of the time. Focusing on the mosaics of general scientific communities allows us to get the “big picture” of how scientific knowledge – writ large – has changed through time, without necessarily excluding or sacrificing that particular data of *who* was producing that knowledge, *where*, and under what specific circumstances.

Considering science from this broad historical perspective sheds light on some of the perennial questions in the *philosophy of science*. Philosophy of science asks questions that attempt to clarify exactly what science is, how it is different from other human endeavours, and how it works. We will take the opportunity to engage with the following pivotal questions of the philosophy of science in the first half of this textbook, all the while drawing from the rich history of science:

- **Absolute Knowledge:** Is there anything we can know with absolute certainty? That is, are there any theories in the mosaic that will never be replaced, that are established beyond any reasonable doubt? (Chapter 2)
- **Scientific method:** How do scientists evaluate competing theories? What are the criteria they employ to assess theories? (Chapter 3)
- **Laws of Scientific Change:** What is the mechanism of scientific change? How do mosaics and their elements change through time? Is there a pattern to those changes? (Chapter 4)
- **Scientific Progress:** Are our theories becoming better descriptions of reality? Is there such thing as scientific progress? (Chapter 5)
- **Science and Non-science:** What’s the difference between science and non-science? What differentiates scientific theories from unscientific theories and scientific changes from unscientific changes? (Chapter 6)

Our culture is saturated by scientific claims and the technological results of scientific investigations. We are used to hearing about scientific discoveries in popular news media. Careers increasingly require a level of scientific literacy, and you may even be pursuing a career as a practicing scientist. This is all to say that we talk about “science” all the time. But have we really taken the time to think about what science is, how scientific theories are accepted and rejected, and the degree of certainty we can have about scientific claims? Engaging with the aforementioned questions in the first chapters is an opportunity to do just that – to look at science from a new perspective, with fresh eyes.

After having done so, we will trace the genealogy of our contemporary scientific worldview by examining the mosaics of four key historical moments and addressing two key historical questions:

What was the content of the mosaic at each of these four moments?

How do these mosaics change over time?

That is, what theories did that community actually consider to be the best available in each scientific field/discipline (astronomy, physics, biology, etc.) in any given historical period and what led to the eventual replacement of those theories? In chapter 7, we will dive into the beautiful systematicity of the Aristotelian-Medieval worldview, shedding light on the theories which made up their cosmology, physics, and informed their medical practices. In chapter 8, we will turn to a lesser-known and underappreciated worldview in the history of science – the Cartesian worldview. We will turn to the Newtonian worldview in chapter 9, with a focus on how its mosaic eventually replaced the Cartesian worldview in Europe and represented the final shift away from the Aristotelian-Medieval worldview. Finally, in chapter 10, we will consider how the acceptance of a key set of new theories led to the shift from the Newtonian to the Contemporary scientific worldview, returning us to the present.

With a greater grasp of science’s history and philosophy under our belts, we will bring them both together in our concluding chapter 11. Here, we will discuss the *metaphysics* of science and how mosaics shape metaphysical assumptions and, concomitantly, worldviews. Additionally, chapter 11 will be an opportunity to consider the limitations of an introductory textbook to so vast a topic, but we will do so by directing you to fascinating avenues of further research.

2.

Absolute Knowledge -- Introduction to History and Philosophy of Science

Chapter 2: Absolute Knowledge

Intro

“Science has established beyond any reasonable doubt...”, “physicists have proven that...”, “it is absolutely true that...” – such phrases are abundant not only in popular science, but even in academic literature. But is there such a thing as *proof* in science? Can we ever establish anything *beyond any reasonable doubt*? The history of science shows us that scientific theories change through time. What was accepted 500, 250 or even 10 years ago may or may not still be accepted nowadays. We’ve seen how drastically our theories of free fall have changed over the last several centuries. Similarly, in less than half a millennium, we moved from the theory of a finite geocentric universe to our contemporary cosmology. The theory of evolution by natural selection has been accepted for no more than a century. Our phylogenetic tree of life is constantly changing as we discover new species or new relations between species. Theories in all fields of inquiry change over time. But if our theories change through time, then is there anything unchangeable in the mosaic? In other words:

Can we know anything with absolute certainty?

Or, alternatively:

Is there absolute knowledge?

This is the central question of this chapter.

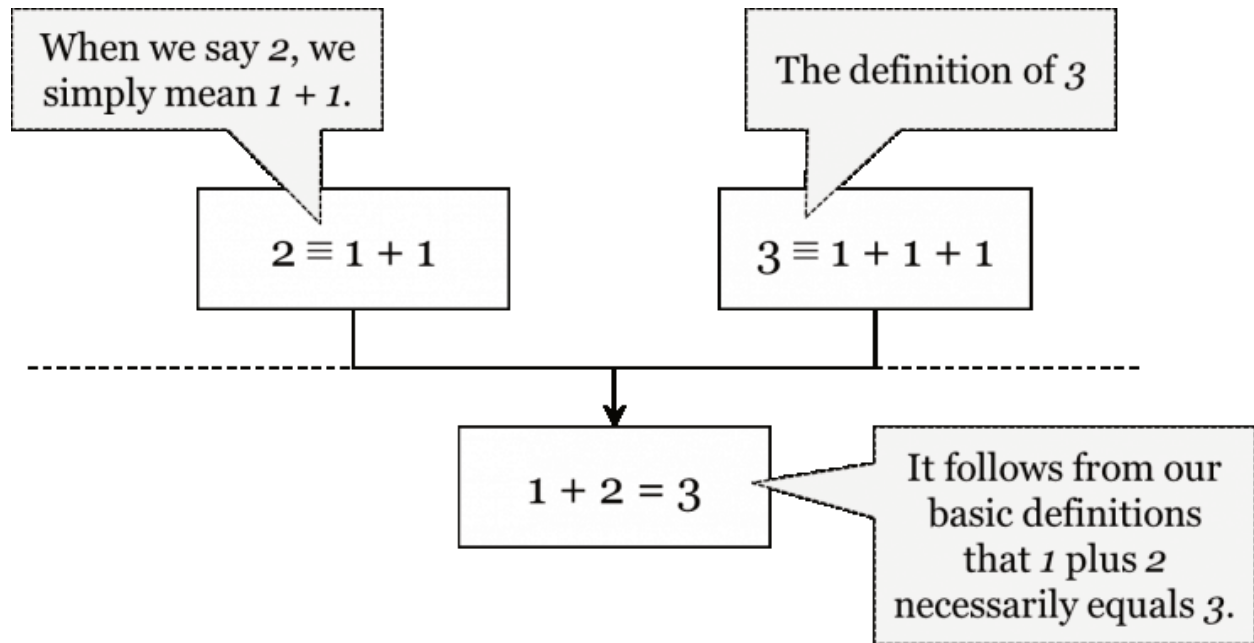
Three Cases

Let’s consider some examples. Here is the first one:

$1+2=3$

Question: how can we show that this proposition is actually true? Keep in mind, what we are asking here is not how we as human beings historically came to discover that one plus two equals three. Instead, the question is: how do we *know* this proposition is true? How is it *justified*?

While it might be argued that it probably took centuries of human experience with the world before this proposition occurred to our ancestors, it is equally clear that we don’t need any experiments or observations to *justify* this proposition. Indeed, we don’t need to conduct any experiments or observations in order to establish that one plus two actually equals three at all times. But if it is not based on experience, then what is it based on? The answer is: the truth of this proposition stems from the very meanings of the terms “two”, “three”, “plus”, and “equals”. What do we mean by “two”? Setting aside the question of a precise mathematical definition of the concept, we can say that “two” roughly means something like “one and another one”. The same goes for the concept of “three”, which is essentially short for “one, another one, and another one”. But if that’s what we mean by “two” and “three”, then it follows deductively that one plus two equals three. In other words, this proposition holds true as far as “two”, “three”, and other terms are understood in the way that they are usually understood. Importantly, we don’t need to conduct any experiments or observations to ascertain that one plus two equals three at all times and in all places, for the proposition is true by virtue of the definitions of its terms.



Now consider another example:

All swans are white.

How do we know this? Once again, the question is not when and under what circumstances humans first came to appreciate the truth of the statement. Rather, the question is: how can this proposition be justified? What makes it true?

The intuitive answer is that this should have to do with experience; surely, one cannot know anything about the colour of swans unless one goes out and observes some actual swans! To establish that all swans are white, one should first observe the colours of individual swans and record the results of these observations. Now, suppose we go out and find a white swan. We formulate this in a singular proposition that describes our experience:

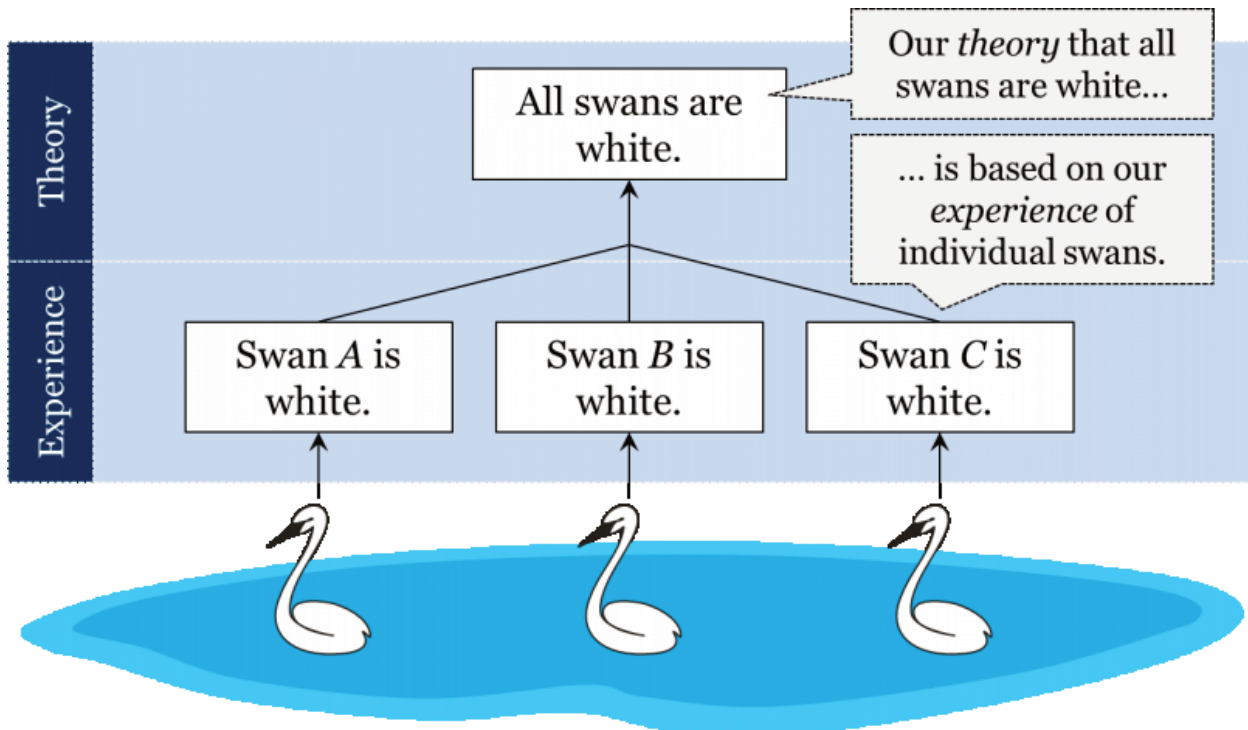
Swan a is white.

We then find two more swans, observe that they are also white, and record this in the following two propositions:

Swan b is white.

Swan c is white.

Based on the results of these observations, we generalize and conclude that all swans are white. This inference from individual instances to a general proposition is called *induction*.



In short, we seem to agree that the proposition “all swans are white” is justified by *experience*: it is an inductive generalization of the results of our *observations* of individual swans.

Our third and final example is from physics – the law of universal gravitation:

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

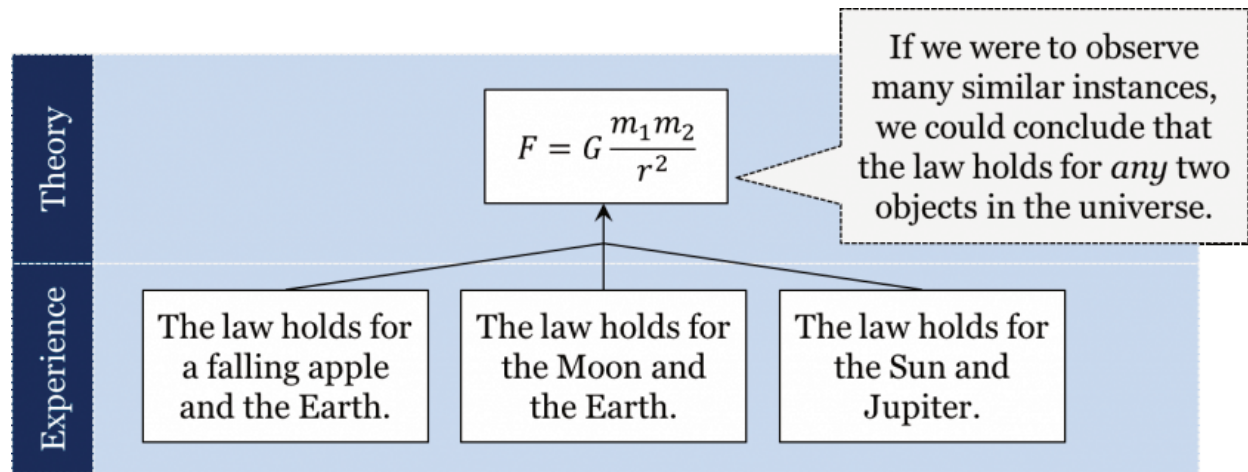
It doesn’t take any extensive knowledge of math to grasp the gist of the law. Here, F stands for the force of attraction between two material objects, m_1 and m_2 are the respective masses of these two objects, r is the distance between them, and G is a constant that we can ignore here. The law says that there is a force of attraction between any two objects in the universe; the greater the masses of these objects, the greater the force of attraction between them; the greater the distance between them, the smaller the force. In other words, the attraction between two objects increases with mass and decreases with distance. The same idea can be expressed in more technical terms: the force of gravity is proportional to the product of the two masses and is inversely proportional to the square of the distance.

Now, how can we justify this proposition? How do we know it’s true? More specifically, can the law of gravity be justified independently of experience simply by virtue of the definitions of its terms, or does it take some experiments and observations to justify it? Is it similar to “one plus two equals three”, or is it similar to “all swans are white”? In other words, is it possible to show that this proposition is true simply by analysing the definitions of “force”, “mass”, and “distance”, or should we actually go out there and see how objects in the world behave in order to justify the law?

If we pay attention to the concepts of “force”, “mass”, and “distance”, there is nothing there that suggests that all objects should attract each other with a force proportional to the product of the masses and inversely proportional to the square of the distance. In fact, we could conceive of an infinite number of different propositions which use the exact same concepts of “force”, “mass”, and “distance”. For instance, we could say that the force of gravity is proportional to the *sum* of the masses, or that it is inversely proportional to the *cube* of the distance, or that it *increases* with the distance:

$$F = G \frac{m_1 m_2}{r^3}$$

Thus, there is no way to deduce the law of gravity merely from the definitions of its concepts, for there is clearly more than one way in which these concepts could be put together. But then how do we know which one of these numerous possibilities is true? The intuitive answer is that we have to go out and observe: there is simply no other way to know which one of these is the case without conducting experiments and observations. How might we do it? Perhaps, if we were to observe that it holds for the Earth and a falling apple, and if we were also to observe that it holds for the Earth and the Moon, the Sun and Jupiter, as well as for any two planets, then we could generalize and conclude that it holds between any two objects in the universe.



In brief, intuitively the mode of justification of the law of gravity appears to be similar to that of “all swans are white”. The fact that the law of gravity can be *expressed* as a mathematical equation shouldn’t confuse us here. It makes no difference how the proposition is expressed: the same proposition can be expressed mathematically or in a natural language, but that doesn’t change the gist of the proposition. The same proposition can be expressed as “ $1+2=3$ ” or as “one plus two equals three” – its truth follows from the definitions of its concepts regardless of how we express it. Similarly, we’ve seen how the law of gravity could be expressed both as a mathematical equation and as “the force of gravity is proportional to...” – it makes no difference. The important point here is that one needs to conduct experiments and observations to check if the law is true.

Analytic vs. Synthetic

Let us now appreciate that the three propositions we’ve discussed belong to two different categories. We say that the proposition “one plus two equals three” is an *analytic* proposition, while the law of gravity and “all swans are white” are *synthetic* propositions. Analytic and synthetic propositions are two distinct types of propositions.

Now, what makes a proposition analytic? Analytic propositions are either definitions or deducible from definitions. Since analytic propositions are true by definition, they can never contradict the results of experiments and observations. For instance, the proposition “ $1+2=3$ ” will hold at all times and places regardless of what we happen to observe. Consider another typical analytic proposition:

All bachelors are unmarried.

This proposition is true by definition, since “bachelor” is *defined* as someone unmarried. If someone were to claim to have observed a married bachelor, we would probably conclude that she uses the word “bachelor” differently, for it is impossible to observe a married bachelor, insofar as we stick to our common definition of “bachelor”. To express the same idea differently: an analytic proposition necessarily holds in all possible worlds, i.e. its opposite is inconceivable. Thus, the proposition expressed as “ $1+2=3$ ” is true in all possible worlds, in the sense that it’s impossible to conceive of a world where this is not the case. Of course, we can always decide to change the definitions of our concepts and, for example, agree that “2” stands for “one, another one, and another one”. But that would effectively change the meaning of “ $1+2=3$ ”; it would no longer be the same proposition. So as long as “2”, “3”, “+”, and “=” are defined the way they are commonly defined, the proposition “ $1+2=3$ ” holds, come what may.

This is clearly not the case for synthetic propositions. Synthetic propositions are ones that are *not* deducible merely from the definitions of their concepts. Because they are not deducible from the definitions of their concepts alone, they can potentially contradict the results of observations and experiments. This is the same as to say that synthetic propositions do not hold in all possible worlds, for their opposites are conceivable/imaginable. Thus, we can easily imagine a world where at least some swans are not white, or a world where the law of gravity is different.

Analytic Propositions	Synthetic Propositions
<p>Deducible from definitions.</p> <p>Cannot contradict the results of experiments or observations.</p> <p>Necessarily hold in all possible worlds: the opposite is inconceivable.</p>	<p>Not deducible from definitions.</p> <p>Can contradict the results of experiments or observations.</p> <p>Do not necessarily hold in all possible worlds: the opposite is conceivable.</p>

Here are a few additional examples of analytic propositions:

$$a(b + c) = ab + ac$$

A centaur is a creature that is part human and part horse.

$$p \rightarrow (p \vee q)$$

The sum of the squares of the lengths of the sides of a triangle is equal to the square of the length of its hypotenuse: $a^2 + b^2 = c^2$.

All of the above propositions are such that their opposites are simply inconceivable, since they are either definitions or logically follow from definitions.

And here are some examples of synthetic propositions:

In combustion, methane combines with oxygen to form carbon dioxide and water: $CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$.

The game of football is played in four 15-minute long quarters.

Paris is the capital of France.

There are no centaurs on planet Earth.

It is easy to grasp that all these propositions are attempting to say something about the world we happen to inhabit and, consequently, they may or may not hold in other possible worlds. We can easily conceive of worlds where these propositions are false. Thus, we can imagine a strange world full of centaurs, or a bizarre world where Paris is the capital of England. We can also conceive of an absolutely ridiculous world where the game of football is played in two 45-minute halves.

This brings us to a useful rule of thumb for distinguishing analytic from synthetic propositions: if the opposite of a proposition is conceivable, i.e. if it doesn't lead to logical contradictions, then the proposition is synthetic. Conversely, if the opposite of a proposition is inconceivable, i.e. if it contains logical contradictions such as the idea of a married bachelor, then the proposition is analytic.

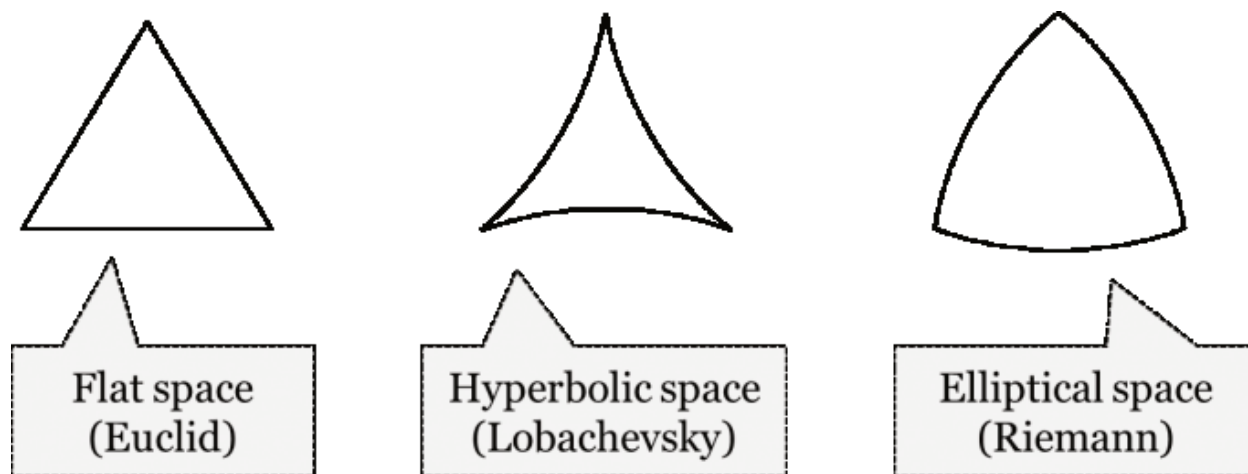
Formal vs. Empirical Sciences

The distinction between analytic and synthetic propositions comes in handy when distinguishing between *formal* and *empirical* sciences. While in *empirical* sciences, such as physics, chemistry, biology, psychology, sociology, or economics, there are both analytic and synthetic propositions, in *formal* sciences, such as logic or mathematics, *all* propositions are analytic. It is the absence of synthetic propositions that characterizes *formal* sciences; indeed, all propositions of math or logic are either definitions or follow from definitions.

Even those mathematical propositions that don't really strike us as analytic are essentially analytic. Consider, for instance, Euclid's famous postulate:

The sum of the angles of a triangle is two right angles.

It is not quite obvious that this postulate is analytic, since it is possible to replace it with its negation. In fact, Euclid's postulate has no place in the non-Euclidean *hyperbolic* geometry of Lobachevsky, where the sum of the angles of a triangle is less than 180° , as well as in the non-Euclidean *elliptical* geometry of Riemann, where the sum of the angles of a triangle is greater than 180° :



But if it is possible to replace some mathematical propositions with their negations, then in what sense are they said to be *analytic* propositions? The answer is this: while some mathematical propositions don't usually strike us as definitions, they essentially *define* properties of mathematical objects that are studied by that specific mathematical theory, and thus are veiled definitions. What makes this possible is the fact that the objects of mathematical theories are all *formal*, i.e. they do not exist *independently* of those mathematical theories that define them with their postulates and definitions. Thus, Euclid's postulate remains an analytic proposition in the context of Euclidean geometry, whereas some other postulates would be analytic in Lobachevsky's and Riemann's non-Euclidean geometries. The same goes for the objects of any formal science.

In contrast, theories of empirical science don't have the luxury of *defining* the properties of their objects. The objects of empirical science – natural or social – are all assumed to exist independently of the empirical theories that attempt to describe them. Thus, physical processes described by a physical theory are not defined/created by the physical theory. Similarly, a biological theory attempts to *describe* plants and animals that exist independently of that theory – the biological theory does not *define/create* organisms. By the same token, theories in political science try to *describe* political processes, like the succession of a new monarch, which would have occurred even if there were no political science.

Because the objects of empirical theories are not defined by empirical theories themselves, empirical theories cannot consist only of analytic propositions. They must contain at least some synthetic propositions that are hypotheses about the objects under study. For example, a physical theory cannot merely give us the definitions of “force of gravity”, “mass”, or “distance”, without telling us how mass and distance affect the force of gravity. Similarly, a biological theory doesn't merely tell us what it means by “species” or “evolution”, but, importantly, tells us how species evolve through time. It is these relations between different features of natural or social objects (structures, systems, processes, etc.) that any empirical theory attempts to uncover. Thus, hypotheses about these objects, i.e. synthetic propositions, are inevitable in empirical science.

The rule of thumb for distinguishing formal from empirical theories is this: if a theory contains no synthetic propositions, it is a formal theory; if a theory contains at least one synthetic proposition, it is an empirical theory. If we consider any theory from any one of the plethora of empirical sciences, we will see that it contains some synthetic propositions, i.e. propositions the opposite of which is conceivable. We've already seen this in the example of the law of gravity. The same applies to the chemical formula for the combustion of methane – we can easily conceive of a world with a different set of chemical laws that make methane combust differently, or even not combust at all. The same applies to all synthetic propositions of all empirical sciences.

Two Questions of Absolute Knowledge

Now that we understand the difference between analytic and synthetic propositions, we must specify the main question of this chapter for the two types of propositions:

Can analytic propositions be absolutely certain?

Can synthetic propositions be absolutely certain?

Let's first consider the former. Is it possible to establish the truth of an analytic proposition beyond any doubt? For instance, can the propositions of math or logic be proven with absolute certainty? The short answer is “yes”, because analytic propositions are either definitions or follow from definitions and, thus, they unfold what is already in a sense implicit in the definitions. Insofar as we stick to a given set of definitions, any analytic proposition that follows from them will be absolutely true. For example, as long as “bachelor” is defined as “an unmarried man”, the proposition “all bachelors are unmarried” will remain absolutely certain. Similarly, as long as we stick to the definitions (including the postulates) of Euclid's geometry, its propositions will always hold.

There is a common misconception in the literature concerning the status of Euclidean geometry nowadays. It is often said that Euclid's geometry was rejected and replaced by a non-Euclidean geometry when general relativity became accepted ca. 1920. The error here is in the confusion of *pure geometry* with *physical geometry*. As a *purely formal* theory, any geometry defines its own

space as a *formal* object and then provides a description of that space; it doesn't attempt to say anything about the actual space of our universe. Naturally, the spaces of different geometric theories can be very different: some geometries may define a flat space (i.e. a space where the sum of the angles of any possible triangle is 180°), while others may define a curved space (i.e. a space where this relationship for triangles does not hold true). The resulting geometries will be very different but, importantly, each of them will only contain analytic propositions and will provide an absolutely true description of *its own formally defined space*. Thus, Euclid's geometry is still and will always be an absolutely certain description of space *as it is defined in Euclid's geometry*. As a formal theory, it doesn't attempt to accomplish anything more.

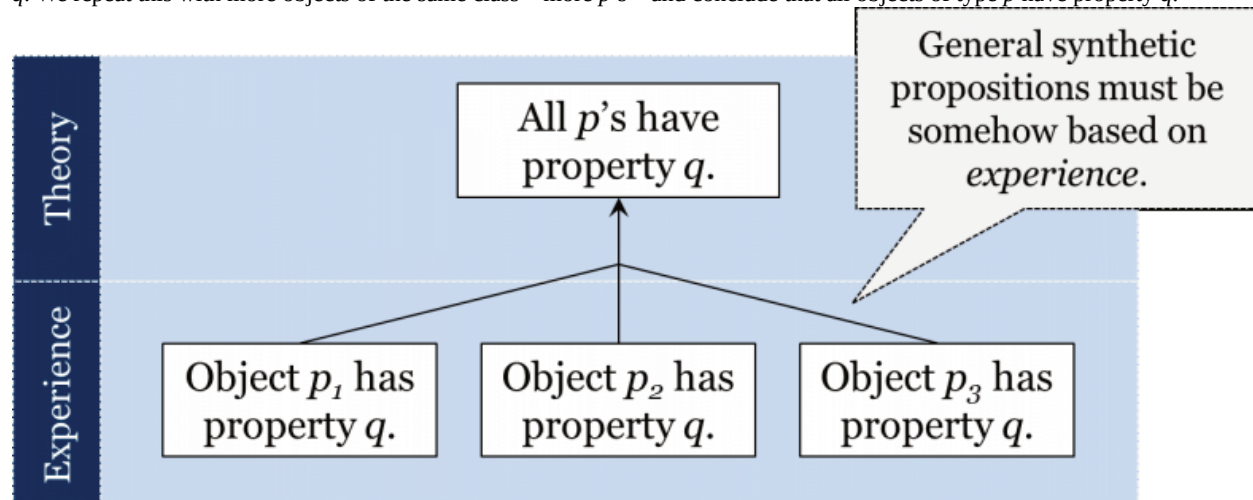
In contrast, *physical* geometry attempts to describe the properties of the *physical* space of our universe. For instance, it attempts to find out whether the space of our universe is flat or curved. Any theory of physics has some physical geometry built into it: for instance, Euclid's geometry was implicit in Newtonian physics, and Riemann's non-Euclidean geometry is built into general relativity. But when any geometry becomes part of a physical theory, it ceases to be a formal theory and becomes physical geometry. It now makes hypotheses about the space of the universe and, thus, no longer consists merely of analytic propositions. Such a physical geometry is an empirical theory and is not to be confused with a pure geometry which is a formal theory. Thus, when we say that Euclidean geometry was rejected, we have to keep in mind that it was only rejected as an empirical theory of the space of our universe. As a formal theory, it's never been and could never be under any threat. This is because it consists exclusively of analytic propositions which, by definition, can never be in conflict with any observational results.

This is equally true for any analytic proposition and, consequently, any theory that consists of only analytic propositions. Therefore, the answer to our first question is "yes": there can be absolutely certain analytic propositions. As a result, we can legitimately speak of mathematicians or logicians *proving* a certain theorem beyond any reasonable doubt.

Now, what about *synthetic* propositions: can they ever be absolutely certain? Can our hypotheses about the workings of the world ever be proven beyond any reasonable doubt? In other words, can we have absolutely certain theories in physics, chemistry, biology, psychology, sociology, economics, etc.? The short answer to this question is "no". To appreciate this, we have to consider how synthetic propositions are justified and what problems their mode of justification faces.

As we have seen with the example of white swans and the law of gravity, synthetic propositions must somehow be based on experience. So intuitively, we are inclined to think that the justification of a synthetic proposition requires some experience – some experiments and/or observations. This goes for both *singular* synthetic propositions, such as "swan *a* is white", and *general* synthetic propositions, such as "all swans are white". Indeed, how else can we justify any of these synthetic propositions, unless we go out there and observe?

Intuitively, we are inclined to accept the following general template. In order to justify a synthetic proposition, we find an object, let's call it p_1 , and observe that it has some property q . We then find a similar object p_2 and observe that it has the same property q . We repeat this with more objects of the same class – more p 's – and conclude that all objects of type p have property q .

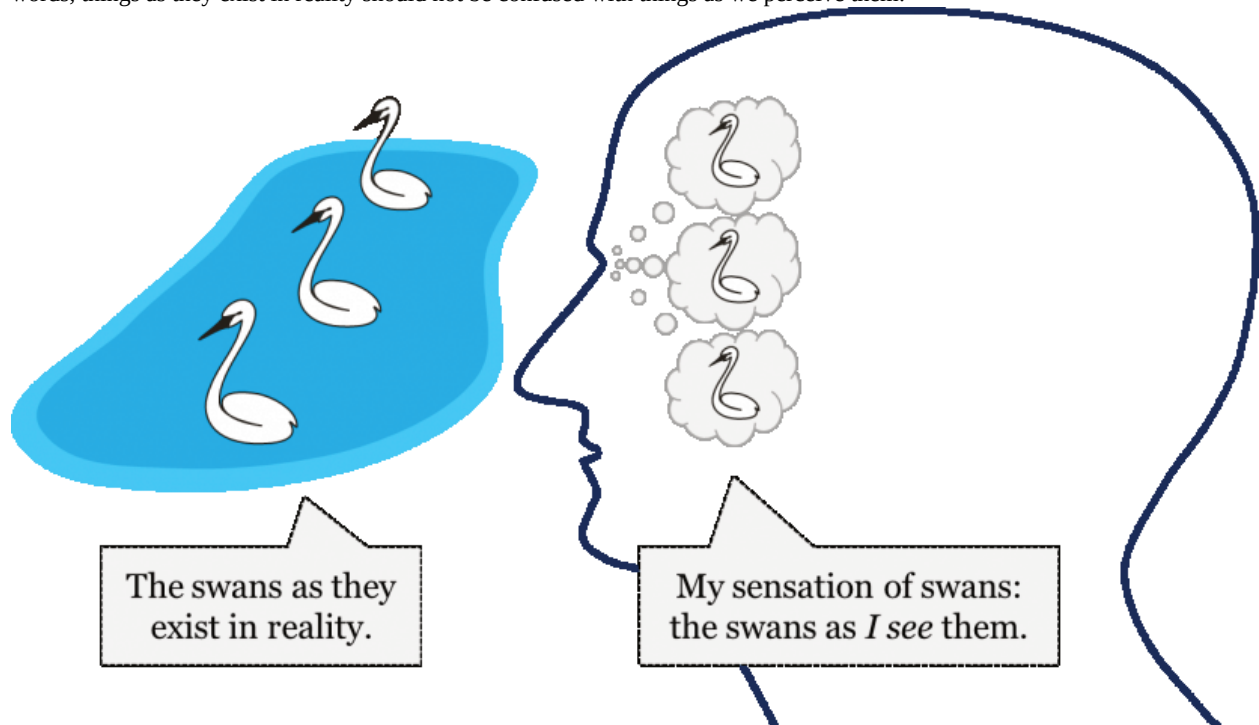


But precisely because all synthetic propositions must eventually be based on experience, they cannot be absolutely certain. There are three giant obstacles that prevent synthetic propositions from being established beyond any reasonable doubt: *the problem of sensations*, *the problem of induction*, and *the problem of theory-ladenness*.

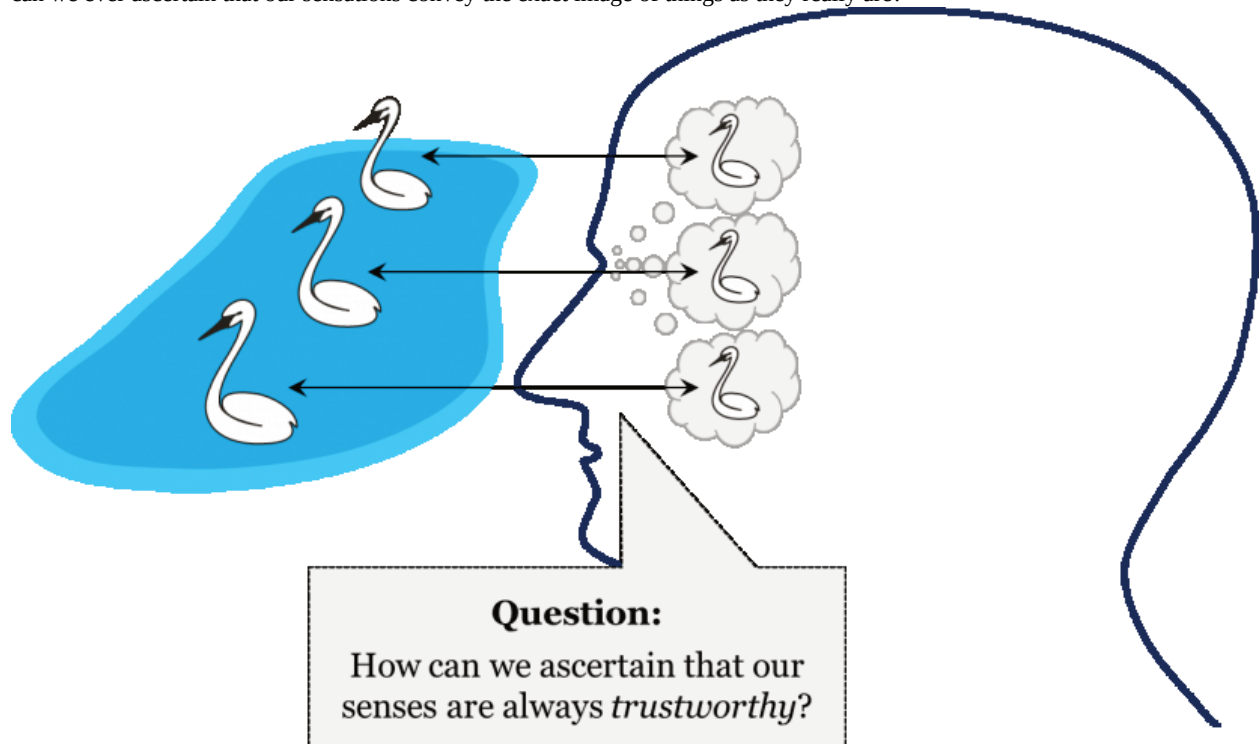
Problem 1: Sensations

First, let us distinguish between the human mind and the world outside of the human mind. Our feelings, emotions, sensations, theories, thoughts, beliefs, ideas, etc. are all in the mind. However, plants, animals, other people, and the Earth itself, as they exist in reality and independently of the human mind, are in the external world. So, our sensations of swans will always be in our minds.

Swans as they exist “out there”, objectively, should not be confused with my sensations of swans, which exist in my mind. In other words, things as they exist in reality should not be confused with things as we perceive them:



In order to claim that the proposition “all swans are white” is absolutely certain, we have to first establish that our sensations of swans are absolutely trustworthy: we have to make sure that swans as we see them are exactly the same as swans as they exist independently of the human mind. But how can we make sure that swans as we see them are exactly as they are in reality? How can we ever ascertain that our sensations convey the exact image of things as they really are?



There are reasons to suspect that our senses are not always trustworthy; surely, we don't normally trust everything we perceive when we are not sober! So, the question of the trustworthiness of our senses is far from idle; if we are going to claim that any of

our hypotheses about the external world are absolutely certain, we have to prove that our sensations convey to our minds exact images of the things as they really are.

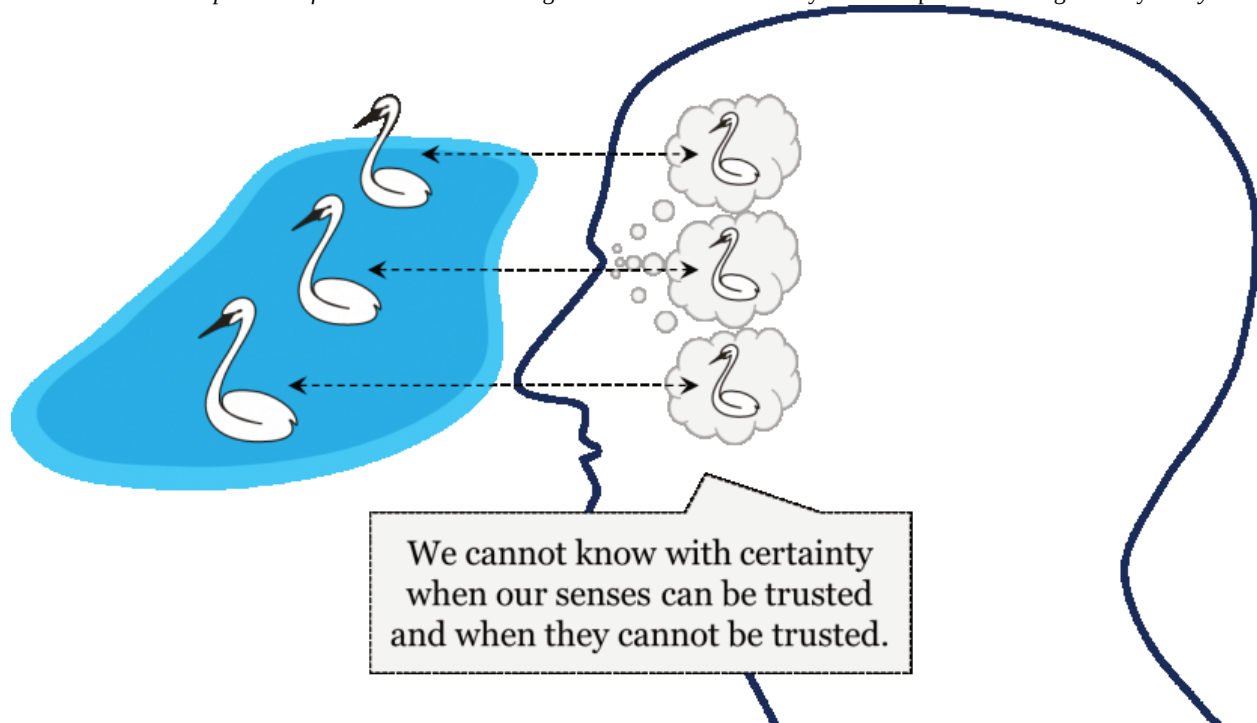
Suppose I see a cup of coffee in front of me. This is my visual sensation of a cup of coffee. How can I possibly make sure that there is *really* a cup of coffee in front of me? I could approach the cup and pick it up, smell it, and taste it. In other words, I could try to confirm my visual sensations with my sensations of touch, smell, and taste. Now, suppose that my sensations of touch, smell, and taste also suggest that there is a cup of coffee in front of me. Will this guarantee that there is *really* a cup of coffee in front of me? Of course, not! All it would demonstrate is that my visual, tactile, olfactory, and gustatory sensations *cohere* with one another. That's all! But the question was not whether my sensations are in agreement with one another. The question was whether my sensations are absolutely trustworthy.

In order to be absolutely sure that the cup of coffee is real, I would somehow have to grasp what the cup is like using a means that *bypassed* my senses. But that is impossible! We don't have a way of getting in touch with the world without using our senses. So even when *all* my senses seem to suggest that there is a cup of coffee in front me, all it really tells me is that *I sense* that there is a cup of coffee in front of me.

What if we asked other people to confirm the presence of the cup? Would that solve the problem? No, it would not. After they observed the cup, they would have to somehow convey the results of their observations to me. There is no way for me to receive their message that does not somehow involve my senses. On top of the fact that other people are relying on their senses, I have to rely on my senses to communicate with them about the information they gathered through their senses. So, all this would tell me is that I have an additional sensation that agrees with my other sensations.

The same goes for any tools we might want to use to confirm our sensations. Suppose I was to use some kind of a coffee detector to ascertain that there is a cup of coffee in front of me. Say, the coffee detector had a screen that would display the message "this is coffee". Would that solve the problem? No, since I would have to read the screen or otherwise interact with the detector and the only way I can do that is through my senses – visual, auditory, etc. Therefore, I would only have one more sensation that is coherent with my other sensations.

In short, in order to be in a position to say that there is a cup of coffee in front of me, I have to be absolutely sure that my senses conveyed the exact picture of the cup as it really is. Yet, this cannot be guaranteed. All I can check is that my sensations are in agreement with one another, but that doesn't necessarily guarantee that the world out there is exactly as my sensations suggest. This is known as *the problem of sensations*: there is no guarantee that senses convey the exact picture of things as they *really* are.



It is even possible that we might be living in a computer simulation; we just don't have an absolutely certain way of knowing that we aren't! While there have been many attempts to solve the problem of sensations, it is accepted nowadays that strictly speaking it cannot have a solution.

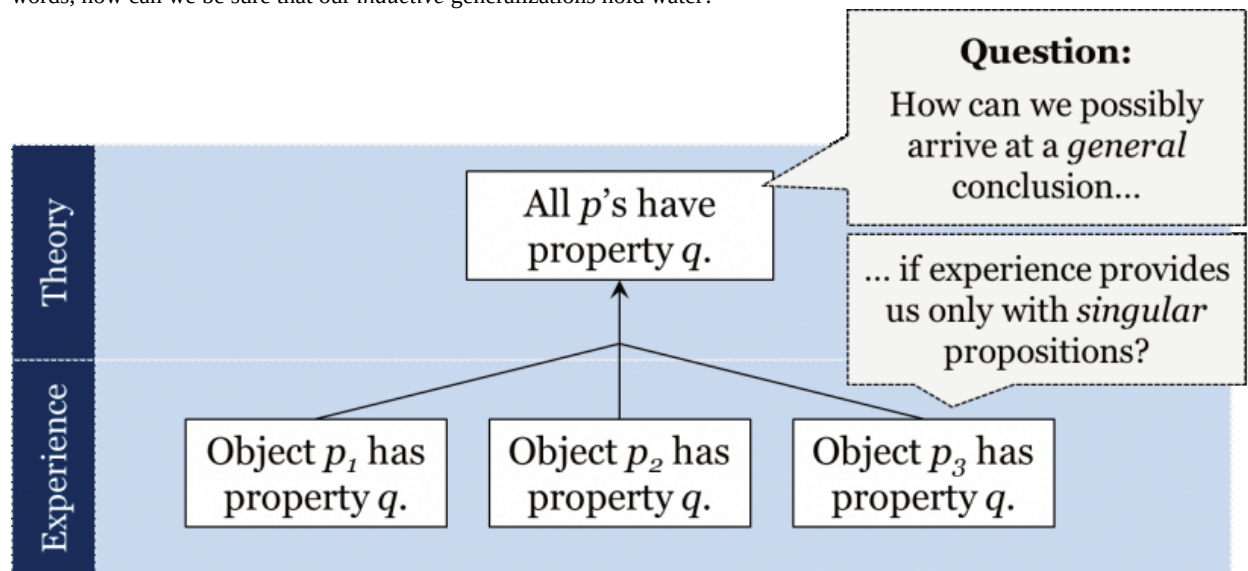
One suggested solution was to argue that because our senses are a product of evolutionary processes, they have adapted to perceive the world correctly, for otherwise humans wouldn't survive. Is this a good solution? No, it's not, because it is based on the theory of natural selection, which is itself an empirical theory; the proposition "our senses adapt to their environment" is a synthetic proposition which can only be justified by invoking knowledge obtained with the use of our senses. Thus, the solution is

itself circular: in order to show that our senses are trustworthy it invokes the theory of natural selection, but in order to justify the theory of natural selection we need to rely on evidence obtained with our senses.

Importantly, the problem of sensations in general doesn't mean that we cannot trust our senses; not trusting our senses wouldn't be particularly conducive to our survival. The point here is *not* that our senses are *deceiving* us – no! The point is that we just cannot be *absolutely sure* that they are not deceiving us. There is a very important difference here: it's one thing to claim that senses are not trustworthy, it's another thing to claim that we cannot know whether they are trustworthy or not. The problem of sensations is all about the latter: it tells us that we simply cannot be absolutely sure that our senses are not deceiving us.

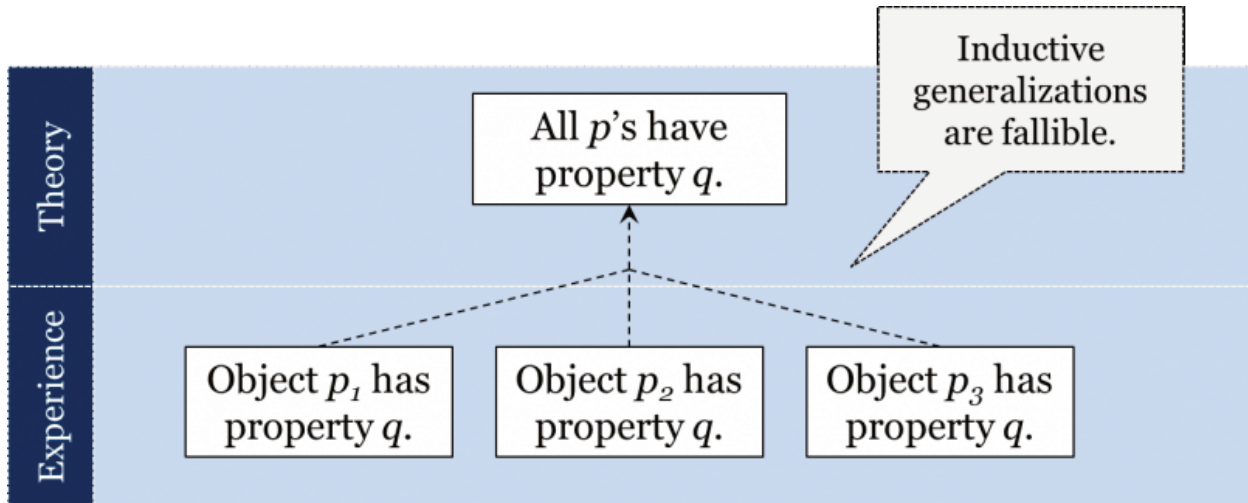
Problem 2: Induction

Suppose for the sake of argument that there is no problem of sensations; suppose that our senses are absolutely trustworthy and convey the picture of the world as it really is. Even if we were to make this assumption, we would still be facing a serious problem. Anytime we observe something, the outcome is a *singular* synthetic proposition, such as “swan *a* is white” or “swan *b* is white”. Question: how can we ever prove any of our *general* synthetic propositions if the outcomes of our observations are always *singular* propositions? In other words, how is it possible to claim that *all* swans are white if our observations only tell us that *individual* swans we have observed so far are white? Is there any guarantee that there won't be any non-white swans out there? Similarly, how can we ever be sure that the law of gravity has been proven beyond any doubt? As a general synthetic proposition, it must be based on observations, but all our observations concern individual cases. We can observe that the law holds between the Earth and the Sun, or the Sun and Jupiter, and so on, but does this prove that the law holds for *any* two objects in the universe? In other words, how can we be sure that our *inductive* generalizations hold water?



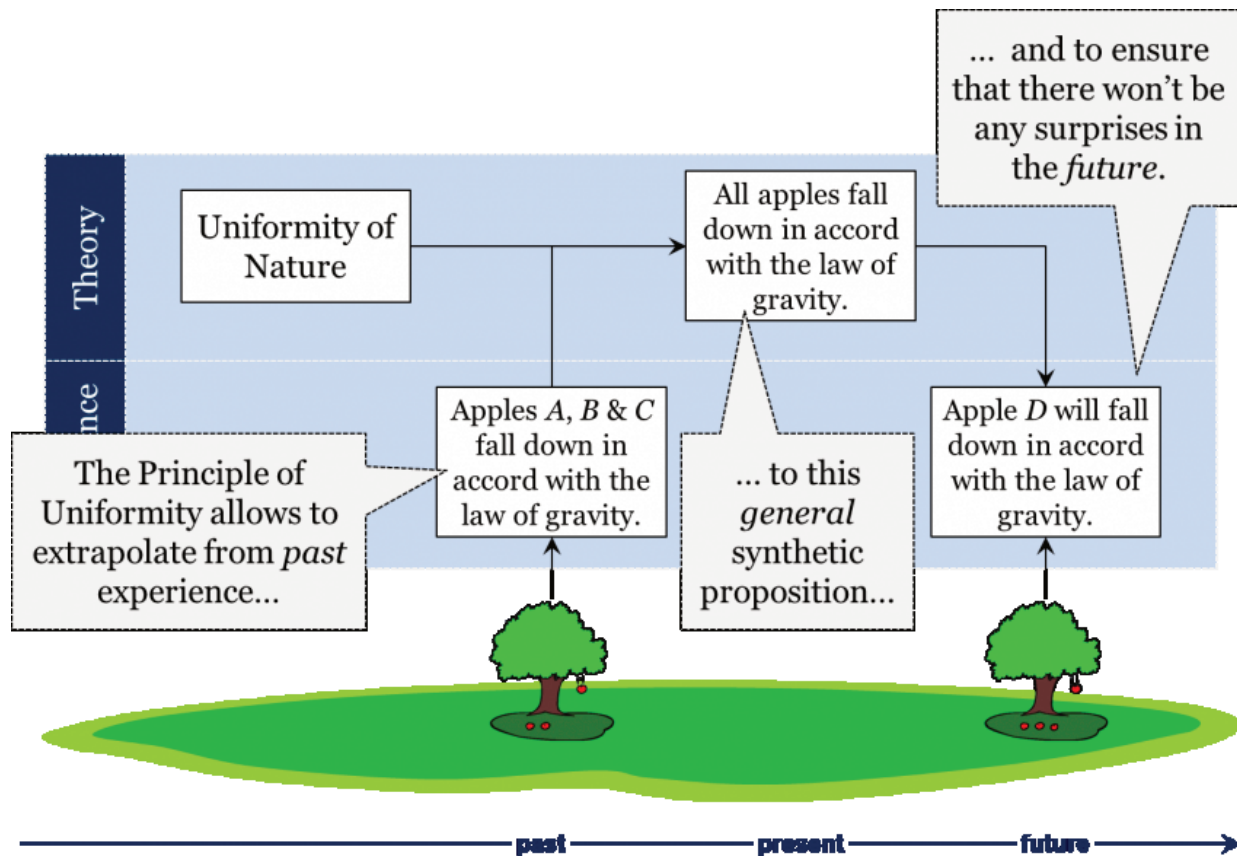
Because general propositions refer to all objects within a given class, they attempt to say something not only about those objects that we have already observed, but *any* object of that type. To establish “all swans are white” beyond any reasonable doubt we should make sure we have observed each and every swan out there, including all swans that have ever lived and all swans that will be born in the future. Surely, this is impossible! But if we are never in a position to observe all swans, how can we be absolutely sure that all swans are white? It would suffice to observe one counterexample – one non-white swan – and the whole generalization would fall apart. Thus, no matter how many millions of white swans we have managed to observe, there is no logical proof that the swans we will observe in the future will also be white. Similarly, we can observe millions of instances of objects acting in accord with the law of gravity, but this doesn't prove that *all* objects in the universe obey this law. Even if all the physicists in the world were to collect their observational results in a huge database, the outcome would still be limited; it would never cover *all* objects in the universe.

This is the gist of *the problem of induction*: because our experience is always limited, our inductive generalizations are inevitably *fallible*.



The only exceptions here are those rare cases, where the number of objects within a class is *finite* and we have managed to observe each and every one of them. For example, the proposition “all capitals of Rus’ were located between 45° and 60° northern latitude” doesn’t face the problem of induction, since Rus’ had a finite number of capitals and all of their locations are known. However, in most cases, there are so many objects within a class that we simply cannot observe all of them. In those numerous cases, our generalizations remain fallible because of the problem of induction.

Now, is there a way to solve the problem of induction? One classic attempt to solve the problem was to invoke the so-called *principle of the uniformity of nature* – the idea that in similar circumstances objects of the same class behave similarly. According to this principle, there are certain regularities in nature and identical causes always lead to identical effects. But since nature is uniform, so the argument goes, we don’t need to observe millions of objects of the same class. It would suffice to observe a few objects of the same class, to note what they have in common and then safely generalize that to *all* objects of that class. This generalization would hold because nature is uniform, i.e. because identical initial conditions always produce identical effects. According to this line of reasoning, we don’t need to observe all swans in the world to prove that all of them are white. They are all white because the ones we have observed so far have all been white and because the principle of uniformity tells us that all swans are similar as they are all products of the same biological mechanism. By the same token, there is no need to test the law of gravity for every pair of objects in the universe, since according to the principle of uniformity all objects with mass must act similarly. In short, the principle allows to safely extrapolate from past experience to all cases of the same class and to ensure that there won’t be any surprises in the future:



With this wonderful principle of the uniformity of nature, we are no longer required to make an infinite number of observations to prove a general synthetic proposition.

Now, does this solution hold water? It is true that if the principle of uniformity were established beyond any reasonable doubt then it would solve our problem. But how do we know that nature is regular and uniform? How is the principle of uniformity itself justified? There are two options here.

Option 1: We can try to argue that the principle itself is an *analytic* proposition, i.e. that it is true by definition. But this is not a viable option: the principle of uniformity is not an analytic proposition, since its opposite is conceivable. For instance, we can easily conceive of a world where regularities allow for exceptions. We can equally conceive of a world with no regularities whatsoever. Clearly, the principle of uniformity itself is a synthetic proposition – it is a hypothesis about the workings of *our* world. This leaves the second option.

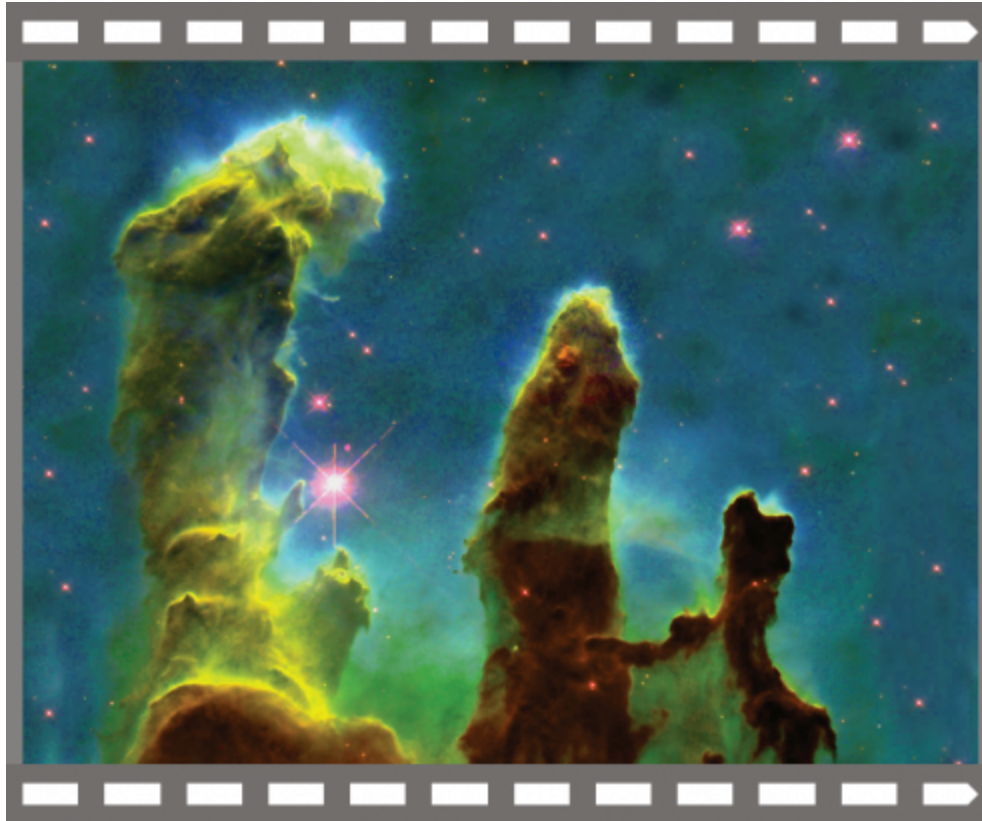
Option 2: We can try to argue that, as a general synthetic proposition, the principle of uniformity is justified through thousands of years of human experience. We see regularities everywhere around us and we see that these regularities don't change through time: apples have been falling to the ground since time immemorial and they still fall down just fine. Thus, the principle of uniformity is based on our observations that in similar conditions objects of the same type have always behaved similarly. In other words, the principle is itself an inductive generalization of our past experience: it assumes that objects of the same type will always behave similarly because that's how they have behaved so far. Thus, the principle of uniformity faces the same problem as any other general synthetic proposition – the problem of induction.

Therefore, we cannot solve the problem of induction by means of the principle of uniformity, because that would lead to a vicious circle: we would use the principle of uniformity to guarantee the absolute certainty of our inductive generalizations, but we would also use induction to justify the principle of uniformity itself, which we would then use to justify induction, which would then justify the principle of uniformity, and so on. We would end up in a vicious circle!

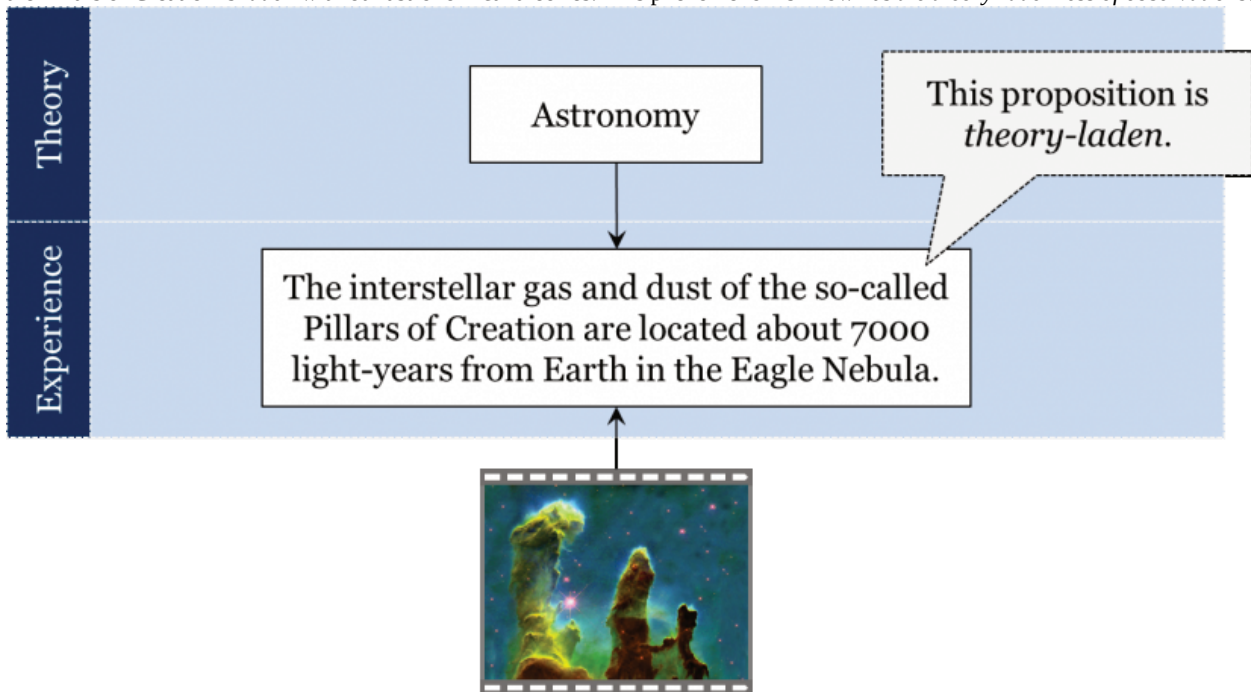
It is accepted nowadays that the problem of induction doesn't have a perfect solution and, therefore, our inductive generalizations based on limited experience are fallible. Of course, this doesn't mean we have to stop making inductive generalizations; all it suggests is that these generalizations are not absolutely certain. We cannot and should not avoid drawing general conclusions; we just have to appreciate that they are not proven beyond any reasonable doubt.

Problem 3: Theory-Ladenness

Let's now proceed to our third problem – *the problem of the theory-ladenness of observations*. Consider the following image:



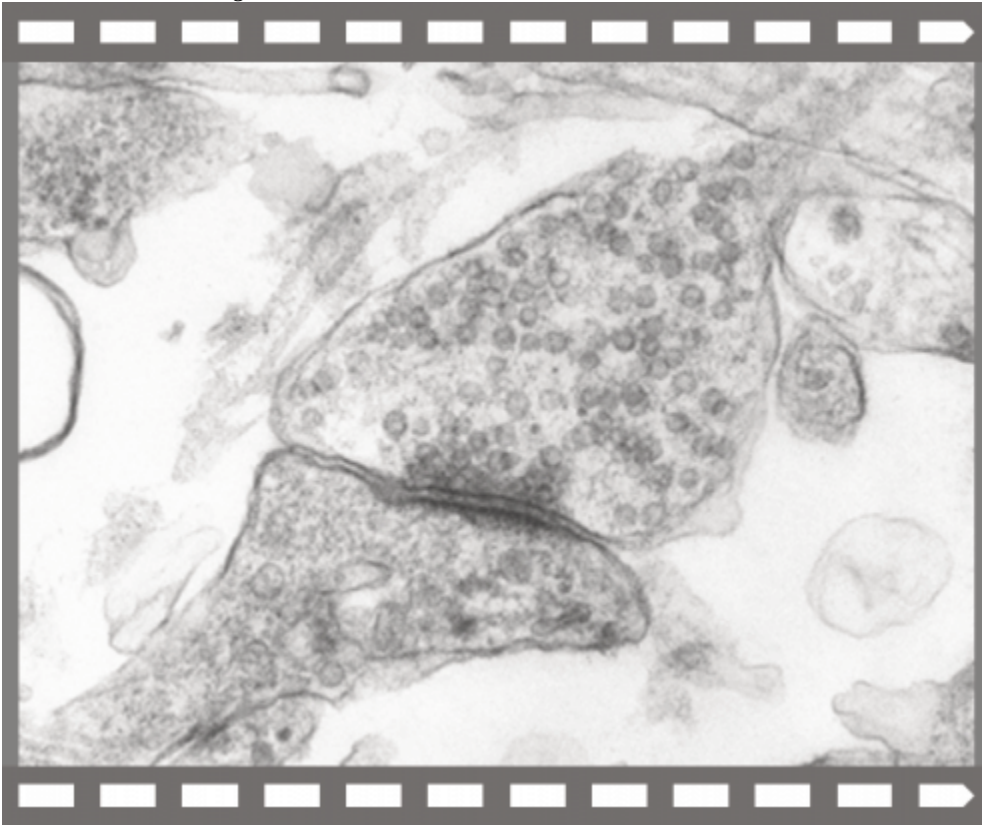
What do we see here? Those of us without astronomical training will probably see some stars obscured by what seems to be a cloud of poisonous gas. However, people trained in astronomy would most likely say that this is the Pillars of Creation, a photograph of interstellar gas and dust about 7000 light-years from Earth in the Eagle Nebula. In other words, our observation of the Pillars of Creation is *laden* with our astronomical theories. This phenomenon is known as *the theory-ladenness of observations*.



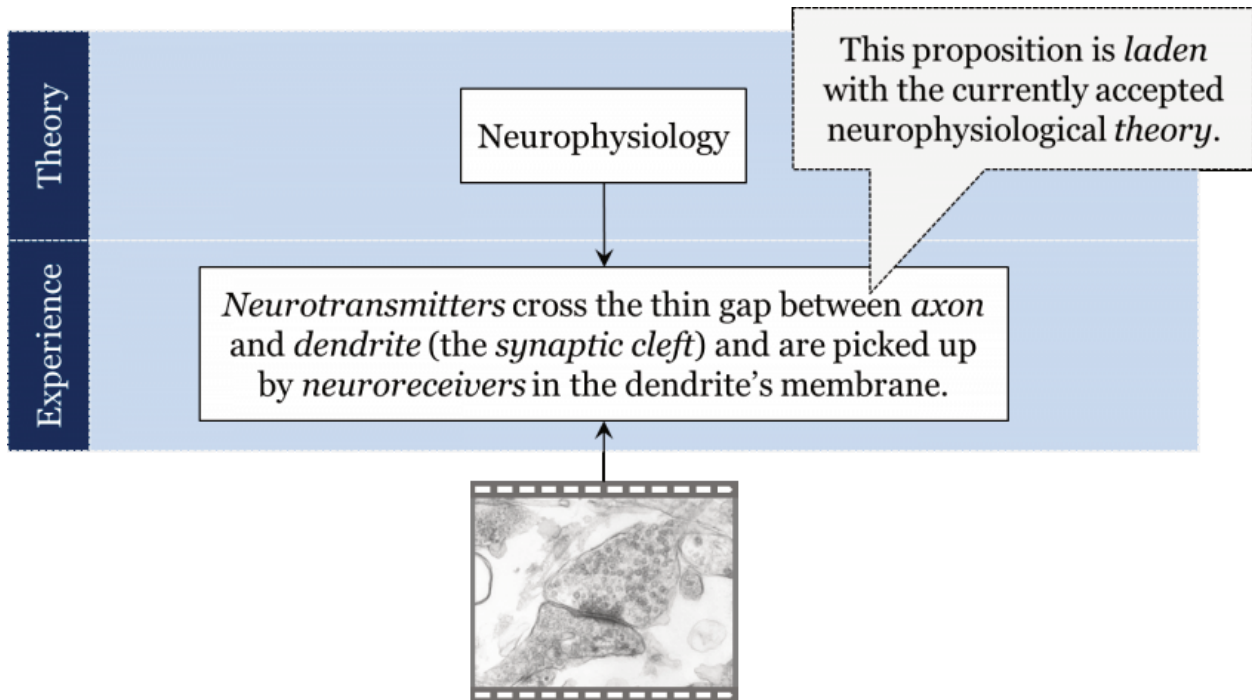
It is nowadays accepted that all observations are *theory-laden*, i.e. they depend on some accepted theories. Another way of saying the same thing is that there are no pure statements of fact, in the sense that all statements of fact presuppose one theory

or another. Note that here “pure” means “unaffected by any theory”. The point is that no observational statement is pure: all observational statements are shaped by one theory or another.

Consider another image:

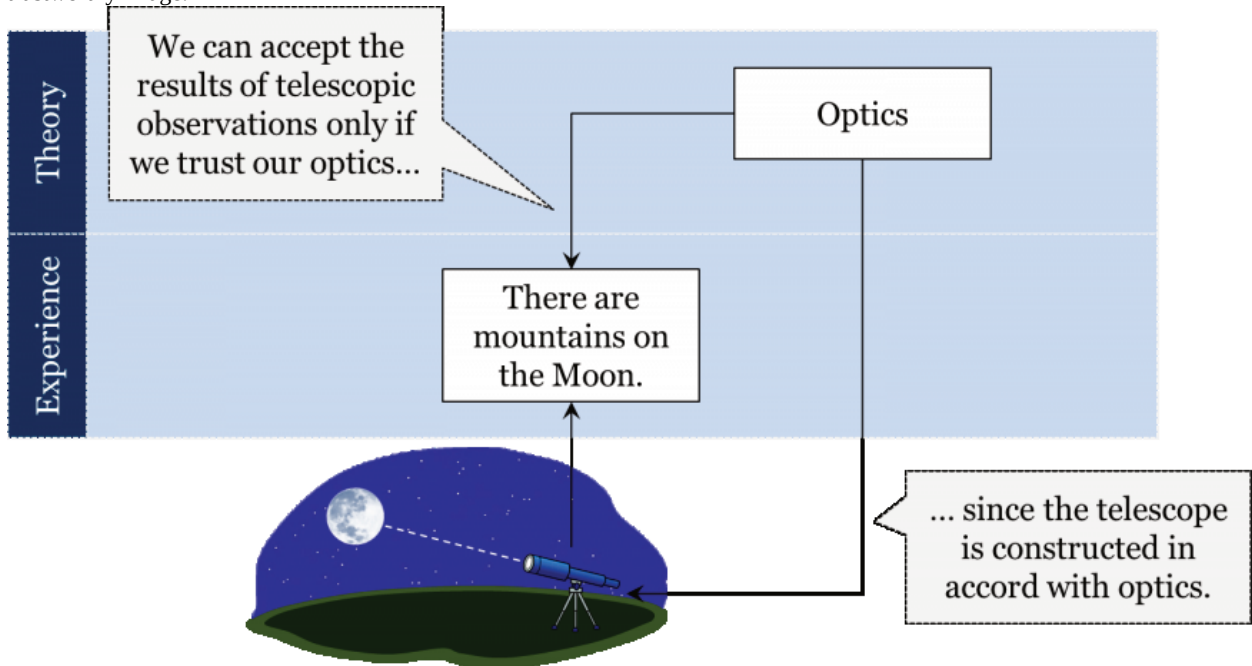


What is this? While a layman will probably see a fuzzy greyish picture, a person educated in neurophysiology will see a *synapse*. More specifically, an educated person will see a *presynaptic terminal* in the upper right portion of the image and a *postsynaptic terminal* in the lower left. They will recognize the presynaptic terminal by the small dark circles it contains. They will see these circles as packets of neurotransmitter molecules called *synaptic vesicles*. These vesicles can release neurotransmitter into the thin gap between the presynaptic terminal and the postsynaptic terminal, which is called the *synaptic cleft*. Neurotransmitter molecules can cross this gap and bind chemically to receptor molecules on the postsynaptic terminal. But this understanding is only possible if we accept some of the theories of neurophysiology. This is another example of our accepted theories shaping the results of our observations.



The general point here is that all propositions that describe our experience are inevitably theory-laden: there are no pure statements of fact.

To appreciate this point, let's take another example. Suppose we look at the Moon through a telescope and observe that it has mountains. Why is this not a pure statement of fact? The proposition "there are mountains on the moon" is not a pure statement of fact because it assumes the reliability of the telescope. But the reason we rely on the telescope is because we know that it was constructed in accord with our accepted optical theory which suggests that a certain combination of magnifiers produces a trustworthy image.



Generally speaking, any observational instrument presupposes one theory or another and, therefore, the observational results obtained by means of an instrument depend on the theories in accordance with which the instrument is constructed.

But what about those cases where we don't seem to use any instruments? Suppose I look out the window with my naked eye and say, "it's raining". How is this proposition theory-laden? Surely, I am not using any instruments, so what theories are presupposed here? Even in this most simple case, my observation is based on some basic assumptions about my own physiology and about

optics. In fact, we do not trust our observations at all times, but only when they satisfy certain basic criteria, such as proper illumination, the absence of visual obstacles such as fog, not having taken a hallucinogenic drug, nor subject to hallucinations due to malnourishment or sickness, etc. For instance, when a teaspoon appears bent or broken in a glass of water, we don't trust this sensation, for we have some prior knowledge of how these things work. We essentially rely on some basic experiential knowledge of optics.

Once again, it doesn't matter that these theories often remain unarticulated; what matters is that any observation – even a naked-eye observation – is based on some theoretical assumptions. At the very minimum, there are some physiological and optical assumptions that allow us to determine which sensations to trust and which not to trust. Thus, there are no pure statements of fact; all observations are theory-laden.

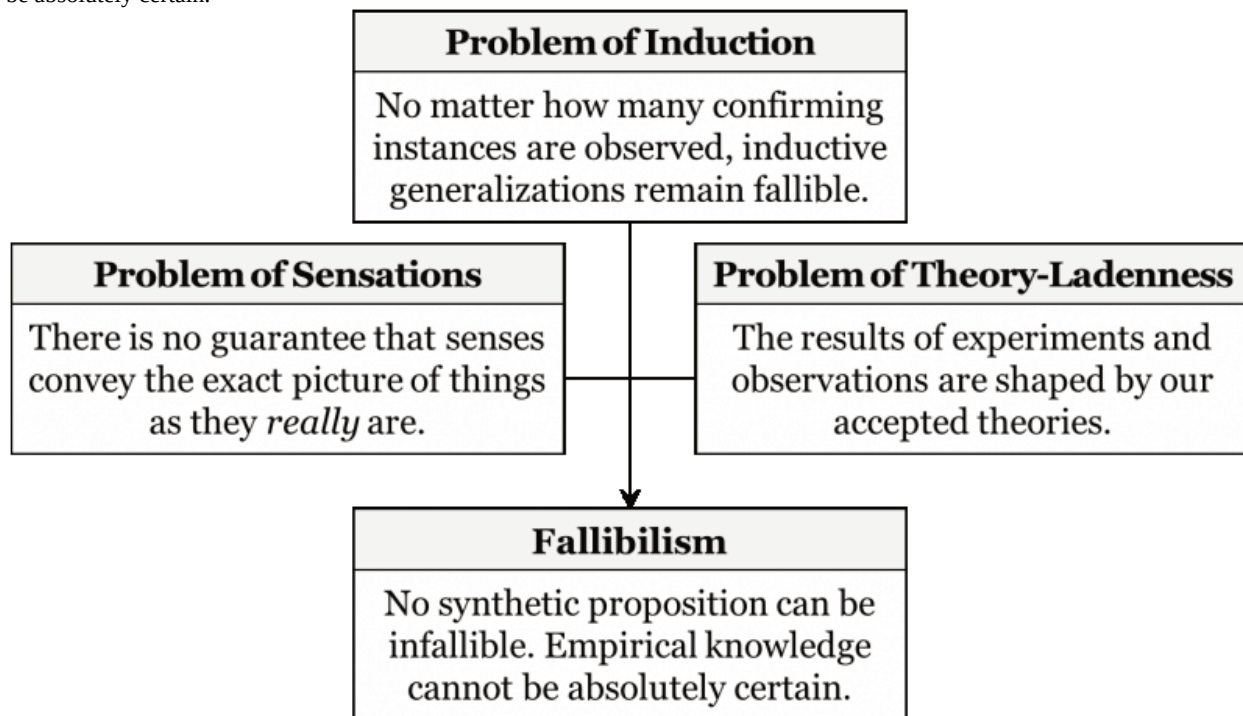
The problem of theory-ladenness becomes especially daunting once we appreciate that the theories through which we see the world change through time. Effectively, the same image can be interpreted differently depending on which set of theories we use to interpret it. Consider the example of a falling apple. Where Aristotelians would see a heavy object composed predominantly of the elements earth and water descending towards the centre of the universe, Newtonians would see a gravitational attraction between the Earth and the apple. In contrast, nowadays we would interpret this as a case of inertial of motion in the curved space-time continuum produced by the Earth's mass. Similarly, where chemists of the mid-18th century would see dephlogisticated air, we would see oxygen. The same goes for any observation in any field of inquiry.

Summary

Let's recap. We've discussed three problems:








- *the problem of sensations*: there is no guarantee that our senses convey the exact picture of things as they really are;
- *the problem of induction*: no matter how many confirming instances are observed, inductive generalizations remain fallible;
- *the problem of theory-ladenness*: the results of experiments and observations are shaped by our accepted theories.

Together, these three problems suggest that synthetic propositions cannot be established beyond any reasonable doubt. Thus, we arrive at the conception of *fallibilism*, which holds that no synthetic proposition is infallible and, thus, empirical knowledge cannot be absolutely certain.



Fallibilism is the currently accepted conception concerning the possibility of absolute empirical knowledge; it opposes the conception of *infallibilism* which holds that empirical science can provide absolute knowledge, i.e. that there can be absolutely certain synthetic propositions. Roughly speaking, infallibilism was accepted until about the early 20th century. Ever since the times of Aristotle and all the way into the early 20th century it was generally assumed that at least some parts of empirical science

had been established once and for all. Philosophers would debate as to how exactly there can be absolutely certain synthetic propositions and which propositions of science have or haven't been established beyond any reasonable doubt, but it was generally assumed that absolutely certain empirical knowledge is possible in one way or another. In the 18th and 19th centuries, the often-cited example of absolutely certain empirical theory was Newtonian physics. Around the time of its replacement with general relativity ca. 1920, it became clear that no empirical theory could ever be safe. That marked the transition to fallibilism. The following table will help summarize the debate between fallibilism and infallibilism:

 Can there be absolutely certain synthetic propositions?	
Yes	No
<p>Infallibilism:</p> <p>Synthetic propositions can be infallible. Empirical knowledge can be absolutely certain.</p> <p> Problem of sensations</p> <p> Problem of induction</p> <p> Problem of theory-ladenness</p>	<p>Fallibilism:</p> <p>No synthetic proposition is infallible. Empirical knowledge cannot be absolutely certain.</p> <p> Problem of sensations</p> <p> Problem of induction</p> <p> Problem of theory-ladenness</p>

The table begins with a formulation of the *question* at issue (the topic): can there be absolutely certain synthetic propositions? It then presents the two opposing answers to the question – the two *conceptions* (positions, viewpoints). Finally, the table lists the *reasons* (pros and cons) for and against these conceptions. The table shows that there are good reasons to accept the conception of fallibilism, which is exactly what contemporary philosophy does.

It is important to repeat that fallibilism doesn't suggest that we have to stop relying on our senses, making inductive generalizations, or describing our observations by means of our accepted theories. In fact, all sciences still do this and that doesn't impede their advancement. The point here is that because we make observations and generalizations, our synthetic propositions are inevitably fallible, i.e. they are not absolutely true.

This brings us to our answers to the two questions we formulated at the outset of the chapter.

Can analytic propositions be absolutely certain? Yes.

Can synthetic propositions be absolutely certain? No.

This outcome has serious consequences for both empirical and formal sciences. Since only analytic propositions can be absolutely true, absolute knowledge is only achievable in formal sciences, such as mathematics or logic. In contrast, absolute certainty is unachievable in empirical sciences, since empirical theories always contain synthetic propositions which, as we know, cannot be established beyond any reasonable doubt. Importantly, this is the reason why there cannot be such a thing as a proof in empirical science. Often scientists speak of having proven something in physics, or biology, or even social science, but strictly speaking proof is only achievable in formal sciences; empirical sciences don't prove anything:

Formal Science	Empirical Science
<p data-bbox="212 304 748 436">E.g. Logic, Mathematics, Systems Theory, Game Theory, Information Theory</p> <p data-bbox="284 478 677 569">Contains only <i>analytic</i> propositions</p> <p data-bbox="305 604 656 688">Absolute certainty is achievable</p> <p data-bbox="274 730 686 821">Propositions can be demonstratively <i>proven</i></p>	<p data-bbox="867 304 1403 436">E.g. Physics, Chemistry, Biology, Psychology, Sociology, Economics, Political Science</p> <p data-bbox="972 478 1299 569">Contains <i>synthetic</i> propositions</p> <p data-bbox="959 604 1310 688">Absolute certainty is unachievable</p> <p data-bbox="924 730 1346 821">Propositions can only be <i>confirmed</i> by experience</p>

Now, if there is no such thing as a strict proof in empirical sciences, does this mean that every empirical theory is as good as any other? Are all competing empirical theories equally good, or is there a way to *compare* competing theories and find out which one of them is better? It is true, that all empirical theories are affected by the problems of sensations, induction, and theory-ladenness; all empirical theories without exception are fallible. Yet, we still find *some* of these empirical theories *acceptable*. Does this mean we have a way of knowing which of the fallible empirical theories are acceptable and which are not? This is the key question of the next chapter.

3.

Scientific Method -- Introduction to History and Philosophy of Science

Chapter 3: Scientific Method

Intro

In the previous chapter, we established that all synthetic propositions are fallible. We discussed three main reasons – three problems – that explain why there can be no absolutely certain synthetic propositions. But since, by definition, all empirical theories contain synthetic propositions, no empirical theory can be proven beyond any reasonable doubt. Yet, despite the fact that all empirical theories are fallible, we seem to believe that some empirical theories are better than others. For instance, today we don't think that Aristotelian physics is as good as that of Newton, just as we don't think that Newtonian physics itself is as good as our contemporary physics. We don't teach Aristotelian physics as the best available description of physical processes; if we teach it, we do so out of *historical* interest, but not as an accepted physical theory. We present general relativity and quantum physics as the best available descriptions of physical processes. A question arises:

If no empirical theory is absolutely true, then why do we think that our current theories are better than the theories of the past?

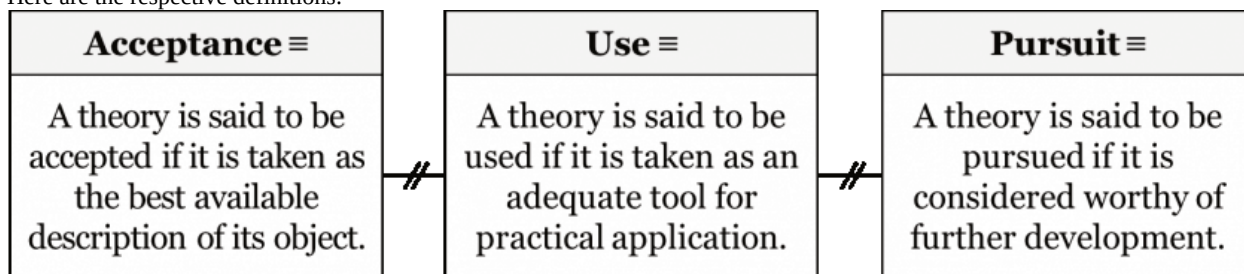
In other words:

How do we decide which theories should become accepted?

This is the central question of this chapter. But before we proceed to the question itself, we need to clarify what we mean by *acceptance*.

Acceptance, Use, and Pursuit

A community can take at least three different epistemic *stances* towards a theory – *acceptance*, *use*, and *pursuit*. A theory is said to be *accepted* if it is taken as the best available description of its object. In contrast, a theory is said to be *used* if it is taken as an adequate tool for practical application. Finally, a theory is said to be *pursued* if it is considered worthy of further development. Here are the respective definitions:



Let's take a closer look at each of these stances separately.

When we say a theory is accepted by a community, what we mean is that the community considers the theory to be the best available description of whatever objects it attempts to describe. Different sciences deal with different types of objects. An object under study can be *natural* (e.g. physical, chemical, or biological processes), *social* (e.g. demographic, political, or economic processes), or *formal* (e.g. mathematical or logical relations). For each of these objects, we can, in principle, have a theory which we take as providing the best available description of that object. It is in this sense that we nowadays accept *the standard model of particle physics* as the best available classification of subatomic particles. Similarly, we accept *the modern evolutionary synthesis* as the best available description of the process of biological evolution. By definition, the scientific mosaic of a community consists of all theories *accepted* by that community.

Importantly, to accept a theory doesn't necessarily mean to consider it absolutely true. Yes, it is the case that historically there have been many communities that considered the theories they accepted as absolutely true, but this is by no means required. Any fallibilist community, including our contemporary scientific community, only accepts theories as the best on the market, not as infallibly true.

Now, *acceptance* is not to be confused with *use*. A theory is considered *useful* when we find a practical application for the theory. A theory can be used in *physical* engineering, such as the construction of bridges, spaceships, or tablets. A theory can also be used in *social* engineering, for instance, to engineer the victory of a certain political party in elections, or to provide a steady income to the owner of a small business. In any case, a used theory may or may not also be *accepted* by the community as the best available description of its object. It is possible to accept one theory but use another unaccepted theory in practice. The current status of classical physics is the best illustration of this point. While no longer accepted as the best available description of its object, classical physics is still used in a clear majority of technological applications. For instance, when the task is to build a

bridge, we will most likely use the equations of classical physics rather than general relativity, because they are much simpler. It is of course possible for the same theory to be both *accepted* and *used*, but one does not entail the other. One doesn't necessarily need to believe that a theory is the best available description of its object to find it useful.

Finally, *acceptance* and *use* should not be confused with *pursuit*. We say that a theory is *pursued* if we see some promise in its advancement and work on elaborating it. We often pursue ideas which are far from acceptable or useful because we believe that they may one day become so. In science, this happens all the time. For instance, no current physical theory explains all the particles and fundamental forces of nature. Physicists are attempting to do so by elaborating different *superstring theories*, which explain the fundamental forces and particles as vibrations of strings much smaller than subatomic particles. However, it is understood that none of these pursued theories is currently accepted. Scientists pursue these theories with a hope that they may one day become accepted. Importantly, a pursued theory is not necessarily accepted or useful. The opposite is also true: we can accept a theory as the best available description of its object without committing ourselves to further pursuing that theory.

Each of these three stances also has its negation. The opposite of *use* is *disuse* or *unuse*: when we don't think a theory is of much use in a specific application, we say it's useless in that respect and as a result it remains *unused*. The opposite of *pursuit* is *neglect*: when nobody works on elaborating a theory, we say it is neglected. The opposite of *acceptance* is *unacceptance*: when we don't think that a theory is the best available description of its object, we say the theory is *unaccepted*. Importantly, we shouldn't confuse *unacceptance* with *rejection*, for in order to be rejected a theory needs to have been previously accepted in the mosaic. In contrast, a theory can remain unaccepted without ever being accepted in the first place. Take, for example, the M-theory (a version of string theory): it is currently unaccepted, but we cannot say that it was rejected, since it has never been accepted to begin with. Phlogiston theory, on the other hand, a once-accepted chemical theory, was rejected more than two centuries ago and is currently unaccepted. Here are the three stances with their respective negations:

	Yes	No
Is a theory taken as the best available description of its object?	Accepted	Unaccepted
Is a theory considered useful in practical applications?	Used	Unused
Is a theory considered worthy of further development, elaboration?	Pursued	Neglected

In brief, it is possible to *accept* one theory, to *use* another theory in practice and, at the same time, to *pursue* some other promising theory. Keep in mind that these stances are not mutually exclusive: the same theory can be pursued, used, and accepted at the same time. For instance, it is possible to pursue an already accepted theory by attempting to apply it to previously unexplained phenomena. Thus, Newton's theory of universal gravitation was already accepted when astronomers continued pursuing it and successfully predicted the existence of Uranus from anomalies in the orbits of other planets. It is also possible to use and pursue the same theory without accepting it, as well as use and accept the theory without pursuing it. Any combination of these three stances is possible.

While it is very interesting to trace the transitions in both used and pursued theories, the question of how and why scientists come to accept their theories seems the most intriguing. This is partially due to the fact that when we acquire a new useful theory, we often also continue using our previous theories; it is possible to use a number of incompatible theories even within the same practical application. For instance, one can use both Ptolemaic and Copernican astronomy to calculate the positions of different planets for a given time. The same goes for pursuit: scientists usually pursue many different competing theories at the same time. Any field of science is full of examples of this phenomenon. *Acceptance*, however, is different; as we only accept the best available theories. Two incompatible theories can be simultaneously *used*, *pursued*, but not *accepted*; only *one* of the competing descriptions can be accepted at a time. For instance, it is impossible to believe that the Earth is both flat and spherical at the same time. Importantly, only accepted theories constitute the scientific mosaic of the community. Thus, the central question of this chapter is how scientists decide which theories are the best available descriptions of the world, i.e. which theories are to be accepted.

Method

So how do we decide whether a theory is acceptable or not? How can we tell that it is the best among the plethora of competing theories? To answer this question, we have to see how scientific communities evaluate theories.

Suppose we take a couple of competing empirical theories that attempt to describe the same object. Suppose that one of these

theories is currently accepted as the best description of its object and that the other theory is viewed as a contender theory that could potentially replace the accepted theory as the best available description. Now, what would it take for the contender theory to become accepted and replace the previously accepted theory in the mosaic? Because we are dealing with *empirical* theories, chances are we are going to need some evidence, i.e. some results of observations and experiments to help us decide if the contender theory is indeed better than the theory we currently accept. Yet, evidence alone won't suffice. What we will also need is some set of *rules for theory assessment*, some *criteria* that a new theory should satisfy in order to become accepted. In other words, we will need some *method* of theory evaluation. *Method* is defined as a set of *requirements* (criteria, rules, standards, etc.) for employment in *theory assessment* (evaluation, appraisal, comparison, etc.):

Method ≡
A set of <i>requirements</i> for employment in <i>theory assessment</i> .

Here are some examples of methods:

A theory is acceptable if it is simpler than its competitors.

A theory is acceptable if it solves more problems than its competitors.

A theory is acceptable if, given the available evidence, it is the most probable among the competitors.

A theory is acceptable if it explains everything explained by the previously accepted theory and also has confirmed novel predictions.

Keep in mind that these are just examples of what sorts of criteria there might be; at this stage we haven't considered which methods scientists *actually employ* when evaluating theories. Before we proceed to explicating the criteria scientists actually employ in theory evaluation, we need to make two important clarifications.

First, it is important not to confuse *methods* with *methodologies*. We define *methodology* as a set of openly prescribed rules of theory assessment:

Methodology ≡
A set of <i>explicitly</i> formulated rules of <i>theory assessment</i> .

Most communities have some idea as to exactly what they think an acceptable theory should look like. Many communities openly state their requirements concerning new theories in their field. This is what we call *methodology*; these methodologies are usually explicitly stated in textbooks, encyclopedias, and research guidelines. But it must be obvious that the rules openly prescribed by a community may or may not coincide with the rules actually employed by that community in theory assessment.

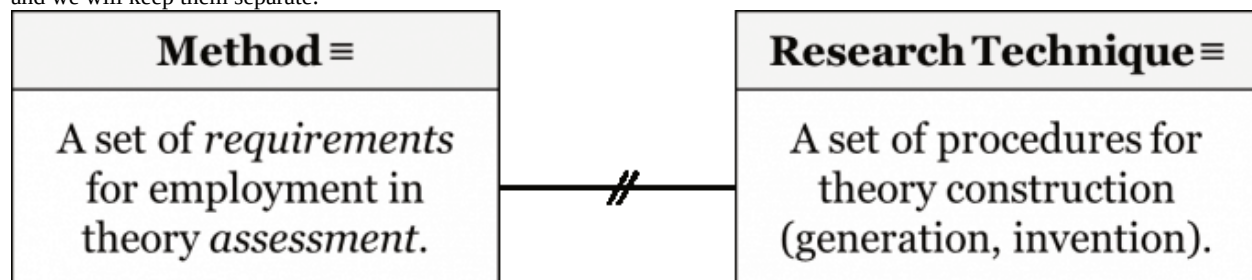
To appreciate the difference between the two, let's look at a simple example: how do we choose our favourite books? We clearly like some books better than others. This shows that we have some expectations as to what a great book should be like. In other words, we have an implicit *method* for book evaluation. Now, let us try to explicate our expectations; let's try to write down what a decent book should be like. These openly stated criteria would be our *methodology* of book evaluation. It is quite possible that prior to this exercise we didn't even have any openly stated methodology for evaluating books. It is also possible that our openly stated requirements differ from our actual expectations. After all, explicating why we choose what we choose is not an easy task – be that for books, foods, or life-partners. Yet, that doesn't seem to stop us from ranking some books higher than others. What this tells us is that implicit expectations (methods) are there regardless of whether we have any idea what those expectations are, or whether we have ever attempted to state those expectations explicitly.

This is similar to what takes place in science. As Steven Weinberg has pointed out, “most scientists have very little idea of what scientific method is, just as most bicyclists have very little idea of how bicycles stay erect”. Thus, if we are to learn how scientists choose their theories, we have to follow Albert Einstein's advice and look at what scientists do (i.e., their *method*) rather than what they *say* they do (i.e. their *methodology*). This is confirmed by the history of science, which provides many examples when the rules of the *methodology* openly stipulated by a community were very different from the actual expectations of that community, i.e. from the *method* they employ to evaluate theories. For instance, the empiricist-inductivist methodology of the late

18th century prescribed that a theory should merely *generalize* the results of experiments and observations without postulating any *unobservable* entities. Yet, it is a historical fact that virtually all of the accepted theories of the time depended crucially on the postulation of unobservable entities. For instance, the fluid theory of electricity posited that electrical phenomena are due to the presence or absence of an *electric fluid*, which repels itself but attracts matter. Similarly, the theory of preformation postulated that men’s semen contains *homunculi*, fully formed tiny people who grow in size into human beings after being deposited into the womb. Finally, Newton’s theory itself assumed the existence of such unobservables as absolute space, absolute time, and the force of gravity. This suggests that the actual expectations of the community of the time differed considerably from the methodological dicta explicitly stated in their textbooks and encyclopedias.

Thus, *method* and *methodology* shouldn’t be confused. Keeping in mind the difference between the two is especially important when reconstructing the state of the mosaic of a certain community at a certain time. To locate the methodologies of the community, we usually look into their textbooks, encyclopedias, and research guidelines. Methods, on the other hand, are much more elusive, as they are normally not on the surface, but are extracted by historians who analyse a series of transitions in a mosaic and attempt to explicate the actual expectations of the community. Now, when reconstructing the state of a given mosaic, it is important to be able to identify the methodologies that were prescribed by the community. Yet, it is even more important to be able to extract the actual expectations of the community, i.e. their methods of theory evaluation. Indeed, it is methods, not methodologies that do the actual job of theory evaluation. In order to convince a community that a certain theory is acceptable, we must make sure that the theory meets their actual expectations, regardless of whether it meets their openly proclaimed methodological rules.

Second, it is important not to confuse *methods* with *research techniques*. In scientific practice, *method* has two related but distinct meanings. Yes, *method* is often used to denote the rules of theory evaluation, just the way we use it in this textbook. But *method* is also often used to denote different *research techniques*, i.e. a set of procedures for generating ideas, constructing theories, or performing experiments, as in a *Materials and Methods* section of a typical scientific paper. The two are by no means the same and we will keep them separate:



It is one thing to ask how we *generate* (construct or invent) our theories; it’s quite another thing to ask how we *evaluate* (assess or appraise) them. Suppose we are trying to come up with an answer to an open question and we decide to sit down with our colleagues and brainstorm to generate some possible answers. Brainstorming is an example of a *research technique*, as it aims to generate as many interesting ideas as possible. Importantly, the outcome of a brainstorming session may or may not be *acceptable*. More likely, some of the outcomes of a brainstorming session will be deemed worthy of pursuit by an individual or a research group, with much further work needed before community acceptance becomes an issue. Whether a theory that results from this becomes accepted or not will be decided by the respective *method* of theory evaluation.

Separating methods from research techniques may sometimes be challenging even for professionals. There is a useful rule of thumb for distinguishing between the two. Research techniques normally contain certain steps or activities that are believed to be conducive to research. Here is a typical example of a research technique:

- Step 1: Pose a question.
- Step 2: Write down ideas.
- Step 3: Discuss.

This research techniques tells us what to do and in what order. Methods, on the other hand, don’t tell us what to do, but always tell us which conditions a theory should satisfy in order to be acceptable. That’s why methods typically begin with “a theory is acceptable if...”. Here is a typical example:

A theory about a drug’s efficacy is acceptable if the drug’s effect has been shown in a randomized controlled trial.

So, if we want to check whether something is a method or a research technique, the general rule is: if it can be formulated as “a theory is acceptable if...”, then it is a method; otherwise it is a research technique. As a quick exercise, consider the following example:

- Step 1: Ask a question.
- Step 2: Conduct background research.
- Step 3: Propose a hypothesis.
- Step 4: Design and perform an experiment.
- Step 5: Record observations and analyse data.

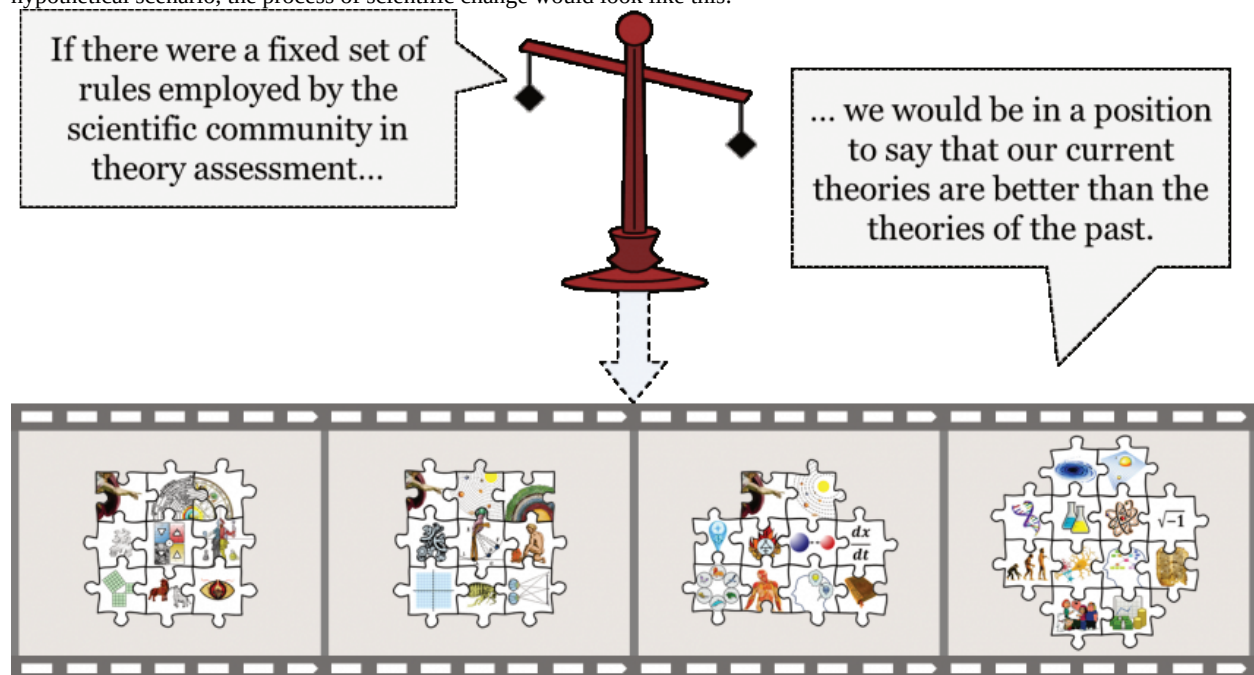
Step 6: Evaluate the hypothesis.

In most literature, this six-step process is called “the scientific method”. But does this really sound like a *method* as we have defined it here? To use our rule of thumb, this formulation does not tell us how a theory is to be evaluated, but rather, outlines a set of research steps. In other words, it tells us how to proceed with our scientific research, but it doesn’t tell us exactly how the fruits of this research are to be evaluated. So, we do not call this a *method*, we call it a *research technique*.

Now, which of the two, methods or research techniques, are we going to focus on? While it might be interesting to inquire into the ways scientists of a certain period generated their ideas, for the purposes of our discussion, what matters is how these ideas were evaluated by the community and whether they were accepted as a result of that evaluation. This reflects the general attitude of scientists who don’t really care about the particular circumstances of a theory’s creation; what they care about is whether the theory actually holds water, i.e. whether it is the best available description of its object. Thus, it makes little difference where exactly a certain theory comes from: I might have seen it in my dreams; I might have used a high-end heuristic technique to generate it; I might have had contact with aliens who told me all about it; I might have even stolen it from a colleague. The provenance of a theory might be an interesting topic when studying the intellectual biography of the theory’s author. Alternatively, the provenance might become relevant when it plays some role in the criteria of the method employed at the time. If, for instance, it turns out that, other things being equal, a community gives preference to theories created by the members of the same nationality, ethnicity, religion, etc., then the provenance of theories becomes part of the method of that community. In any event, what matters is the criteria which are employed to decide whether a theory is acceptable or not. Thus, we will be focusing not on research techniques, but on methods of theory evaluation.

Explicating the Method of Science

Now that we’ve established that theories are evaluated by a method, we can consider what those expectations, rules, or criteria actually are. That is, what specific criteria should a theory satisfy in order to become accepted? If there were a *fixed* set of criteria that scientists of all times and places employed in evaluating their theories, then we would have a good argument that transitions from one accepted theory to the next are not random, but rational. If we could only show that there is an unchangeable method of theory evaluation, then all changes in any mosaic would be governed by this fixed, transhistorical method of science. In that hypothetical scenario, the process of scientific change would look like this:



Naturally, this cosy picture would only make sense if there indeed were a fixed scientific method. But is there such a thing? In other words:

Is there an unchangeable (fixed, transhistorical) method of science?

For most of the history of knowledge, it was generally assumed that while scientific theories change through time, the criteria that scientists employ when assessing competing theories remain somehow fixed. Since the days of Aristotle, if not earlier, philosophers have attempted to openly formulate the requirements of this fixed method of science. Over the centuries, there have been a great many attempts to unearth this elusive scientific method. These attempts can be grouped into several traditions with very diverse formulations of the scientific method. Here are very brief summaries of some of the most notable attempts:

Inductivist-empiricist: a theory is acceptable if it is based inductively on experience.

Conventionalist-simplicist: a theory is acceptable if it explains the greatest number of phenomena in the simplest way.

Pragmatist: a theory is acceptable if it solves the greatest number of problems.

Hypothetico-deductivist: a theory is acceptable if its predictions are confirmed in experiments and observations.

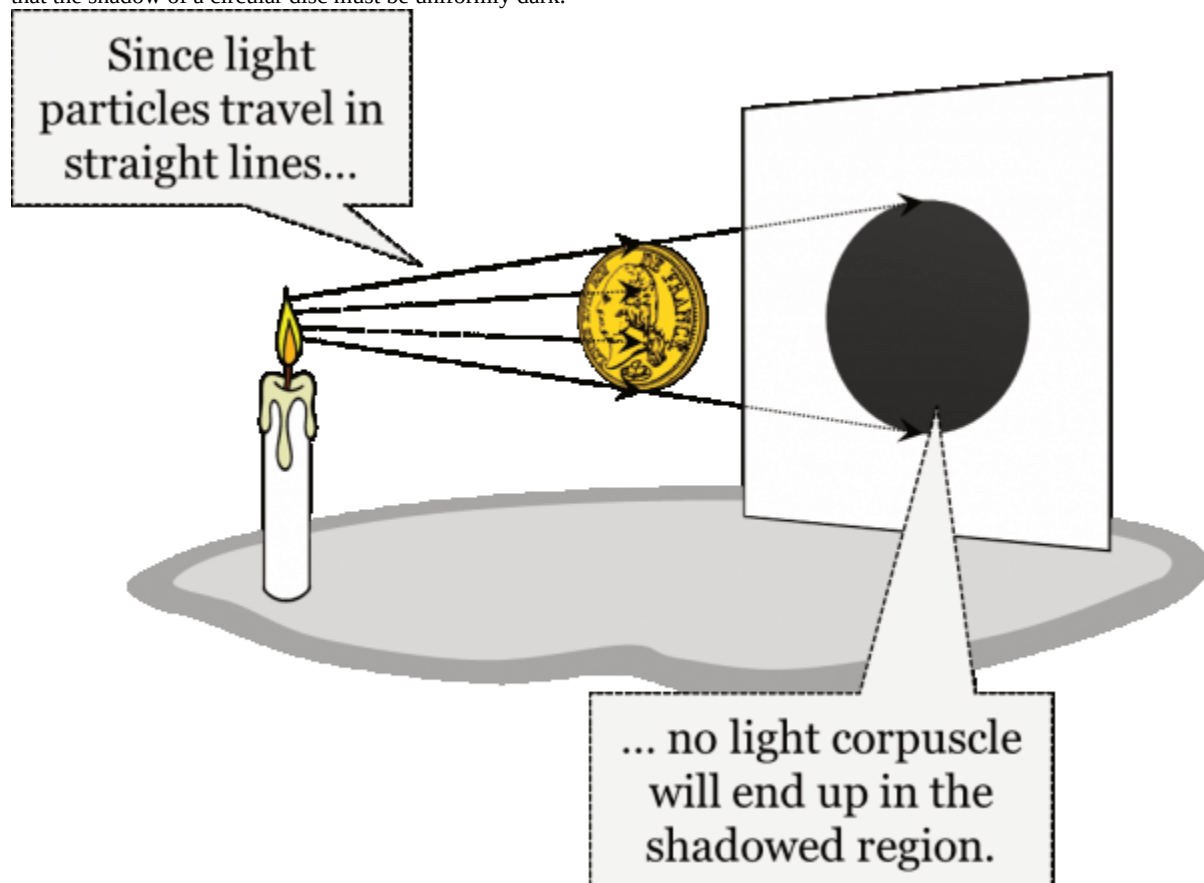
Note that this brief outline doesn't even scratch the surface of the infinitely rich world of discussions on scientific method. It's merely meant as a general sketch.

Now, let us assume for the sake of argument that there is such a thing as a fixed (unchangeable, transhistorical) method of science. How would we spell it out? What are our actual expectations concerning new scientific theories and how can we explicate them?

The best way to explicate our expectations is by studying the transitions in our mosaic over a certain period of time. When focusing on changes in theories that have taken place since the early 1700s, two kinds of communal expectations seem to transpire. While in some cases we seem to expect a new theory to have made successful predictions of *novel* (hitherto unobserved) phenomena, in other cases, we seem to be perfectly happy to accept a theory even if it doesn't make any novel predictions that have been confirmed in experiments and observations.

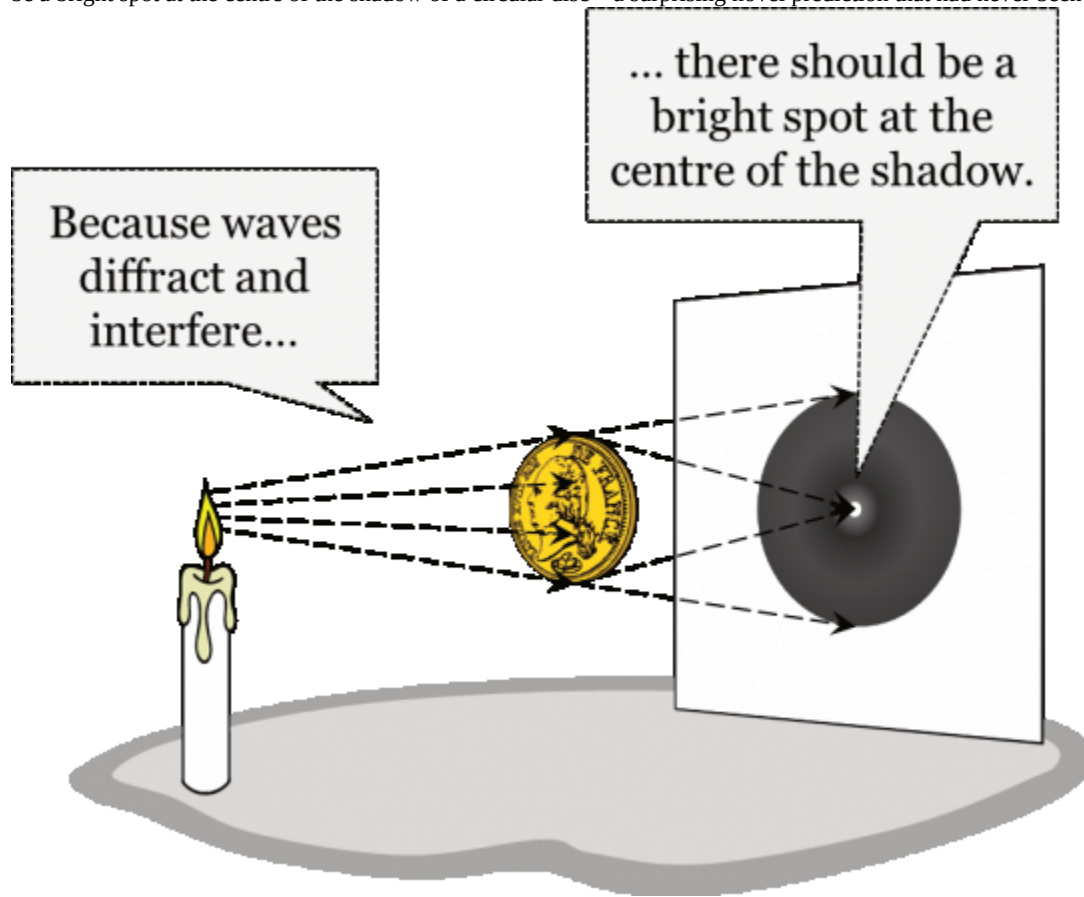
Historically, there have been many cases where a theory has been accepted only after some of its predictions of hitherto unobserved phenomena were confirmed. For instance, the French community of natural philosophers accepted Newtonian physics ca. 1740 only after the confirmation of one of its novel predictions – the prediction that the Earth is slightly flattened towards the poles, i.e. that it is an *oblate* spheroid. The theory's prediction that the Earth is an oblate spheroid was at odds with the then-accepted view of the Earth as a *prolate* spheroid, i.e. as a spheroid that is slightly elongated towards the poles. The prediction of the oblate-spheroid Earth was confirmed by several expeditions organized by the French academy of sciences in the late 1730s. Importantly, Newton's theory became accepted in France only *after* these confirmations.

Similarly, Fresnel's wave theory of light became accepted ca. 1820, right *after* the confirmation of one of its novel predictions – the prediction of a bright spot at the centre of the shadow of a circular disc. According to the previously accepted corpuscular theory of light, light is a chain of tiny particles that travel along straight lines. Like any other particles, light particles were expected to travel along straight lines unless they were affected by an obstacle. In particular, it followed from the corpuscular theory of light that the shadow of a circular disc must be uniformly dark.



The picture suggested by Fresnel's wave theory of light was very different. According to the wave theory, light is a wave that spreads in a universally present medium called *ether* similar to the propagation of water waves as a result of a dropped stone. This

theory also suggested that, just like water waves, light waves should be able to diffract, bend behind obstacles, and interfere with one another. It is because of this ability of light waves to diffract and interfere that the theory made the prediction that there must be a bright spot at the centre of the shadow of a circular disc – a surprising novel prediction that had never been observed before.



The prediction was confirmed in 1819, and the wave theory of light became accepted.

The history of science is full of examples where a theory became accepted only after some of its novel predictions became confirmed by experiments and observations. Other examples of this phenomenon include the acceptance of Einstein's general relativity ca. 1920, the theory of continental drift in the late 1960s, the theory of electroweak unification by Weinberg, Salam, and Glashow in the mid-1970s, as well as the acceptance of the existence of the Higgs boson in 2014.

Yet, a careful study of the history of science reveals many other transitions where a theory became accepted *without* having any confirmed novel predictions whatsoever. There have been many cases where all that the community seemed to care about was whether the theory under evaluation managed to explain the extant observational and experimental data with the desired level of *accuracy* and *precision*.

The acceptance of Mayer's lunar theory in the 1760s provides a nice example of this. It is safe to say that predicting the trajectory of the moon with the same accuracy and precision as the trajectories of other celestial bodies had been a challenging task ever since the inception of astronomy. As a result, a great number of lunar theories had been pursued over the years. The 18th century alone gave birth to several prominent lunar theories, including those of Isaac Newton, Leonhard Euler, Alexis Clairaut, Jean d'Alembert, and Tobias Mayer. When Mayer's theory became accepted, it was precisely because it was found to be accurate enough to determine the moon's position to five seconds of arc. Importantly, the theory became accepted without having any confirmed novel predictions.

The acceptance of Coulomb's inverse square law of electrostatics a few decades later exhibits the same pattern. Charles-Augustin de Coulomb formulated his famous law of electrostatic force in 1785. The law states:

The force of attraction/repulsion between two point charges is proportional to the product of the magnitude of each charge and inversely proportional to the square of the distance between them:

Here is the mathematical formulation of the law:

$$|F| = k_e \frac{|q_1 q_2|}{r^2}$$

According to the law, opposite charges attract, while like charges repel each other. The force of this attraction/repulsion (F) is greater when the charges (q_1 and q_2) are greater and smaller when the distance (r) is greater. Notably, Coulomb's law became accepted in the early 1800s because it fit the available data. The community of the time didn't seem to care about the fact that the theory didn't provide any confirmed predictions of novel, hitherto unobserved phenomena.

There have been many cases in the history of science when a theory became accepted in the absence of any confirmed novel predictions, i.e. cases where the mere precision and accuracy of predictions was sufficient. Other prominent examples of this phenomenon are the acceptance of the three laws of phenomenological thermodynamics in the 1850s, Clark's law of diminishing returns in economics ca. 1900, and quantum mechanics towards 1930. Now, what does this tell us about our expectations? Specifically, do we or do we not require a theory to have confirmed novel predictions to consider it acceptable?

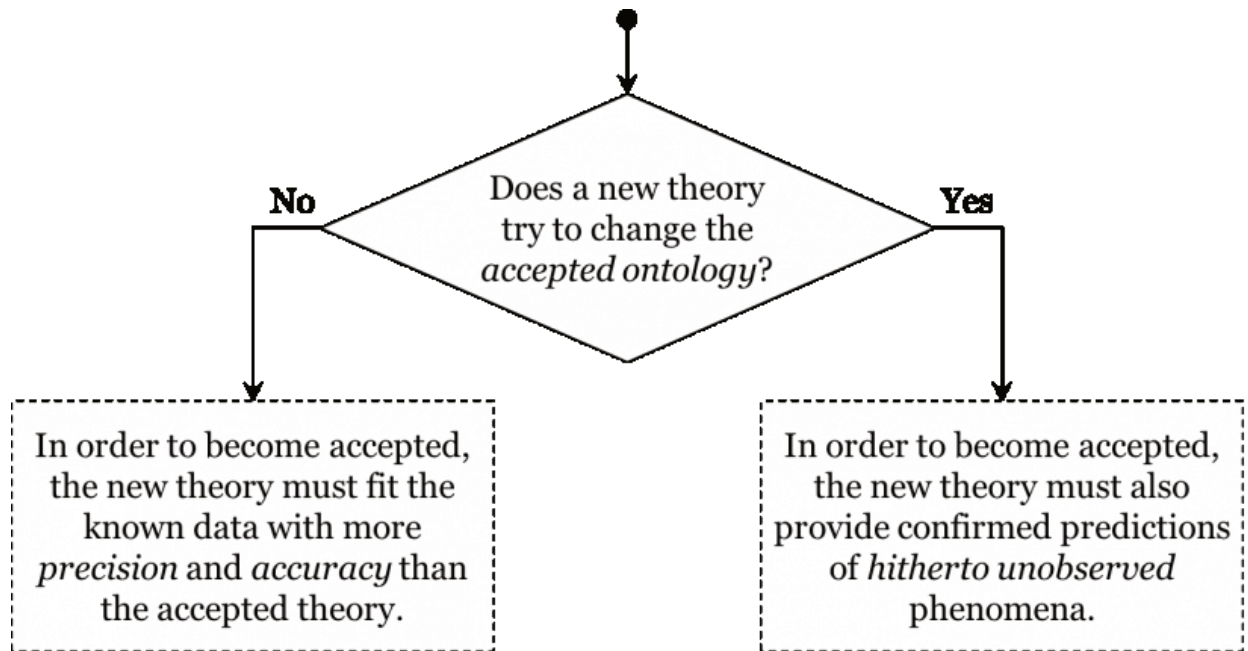
Several generations of philosophers have debated over the role of novel predictions in theory acceptance. Among many others, William Whewell, Karl Popper, and Imre Lakatos all argued that confirmed novel predictions play an indispensable role in theory evaluation. Among their opponents, who argued that confirmed novel predictions don't play any special role in theory evaluation, were John Stuart Mill, Milton Keynes, and Larry Laudan.

A careful examination of historical episodes evinces that scientists do expect confirmed novel predictions, but only in very special circumstances. We expect a theory to be making successful predictions of hitherto unobserved phenomena only when the theory attempts to change our accepted ontology. Now, what is an *ontology*? Roughly speaking, an ontology is a set of views on the types of *entities* and *interactions* that populate the world. For instance, our accepted ontology today includes our belief that there are quarks, leptons, and bosons, that there are more than one hundred chemical elements, that there are millions of species, etc. While most empirical theories presuppose an ontology, not every empirical theory requires the *replacement* of the previously accepted ontology. Very often, a new theory doesn't introduce new entities or new relations, i.e. it doesn't offer an alternative ontology, but shares the ontology of the previously accepted theory. At times, however, a new theory comes with a different ontology – with different assumptions on what entities and relations populate the world. Importantly, our expectations towards new theories depend crucially on whether a theory attempts to introduce a modification to our accepted ontology.

We seem to expect confirmed novel predictions whenever a theory attempts to alter our accepted ontology, i.e. if it tries to convince us to accept the existence of a new entity, new interaction, new process, new particle, new wave, new force, new substance, etc. For instance, Newtonian physics posited the existence of absolute space, absolute time, and the force of gravity acting at a distance – all of which were not part of the then-accepted ontology; thus, it was expected to have some confirmed novel predictions. Similarly, Fresnel's wave theory of light attempted to introduce the idea of light-waves into our ontology; it was also expected to make successful predictions of novel phenomena. The same goes for Einstein's general relativity, which postulated the existence of a curved space-time continuum; such a bold hypothesis could only be accepted if at least some of the novel predictions of the theory were experimentally or observationally confirmed. Note that the confirmation itself need not necessarily be done in a *single* crucial experiment or observation but may rather take a series of experiments and observations that collectively confirm the prediction. For example, in the nineteenth century the Schleiden-Schwann cell theory posited an ontological novelty – the idea that living cells are the fundamental unit of structure in *all* organisms. Thus, the theory was expected to be experimentally confirmed. However, it was accepted because cellular structure was detected under the microscope in a growing catalogue of living species, rather than as a result of one single decisive observation.

By contrast, when it comes to theories that do *not* attempt to change our accepted ontology, we seem to be much more lenient. If a theory doesn't try to alter our views on the constituents of the world, we normally do not require any confirmed novel predictions – we are typically willing to deem the theory acceptable if it provides a more accurate and precise description of already known phenomena. Consider, for instance, Mayer's lunar theory. It wasn't expected to provide any confirmed novel predictions because it didn't try to alter our accepted ontology. It was an attempt to explain and predict the trajectory of the moon by relying exclusively on the elements of the accepted Newtonian ontology, such as *mass*, *force*, *acceleration*, *distance*, etc. Consequently, the community of the time merely expected the theory to be more accurate and precise in its predictions of lunar motion than the previously accepted lunar theory. Likewise, Coulomb's law wasn't attempting to modify the then-accepted ontology of point charges and electric forces. It was merely an attempt to quantify the already known relationship (*attraction/repulsion*) between known entities (*point charges*). Naturally, the community didn't require the theory to predict any previously unseen phenomena. The same goes for any other theory that merely provides a new description of a phenomenon by means of known relations and entities.

To summarize our expectations, if a theory fully relies on the currently accepted ontology, it is not expected to provide any confirmed novel predictions – mere accuracy and precision of its predictions suffice; otherwise, if a theory tries to convince us that there exists some new *type* of particle, substance, interaction, process, force, etc., it *must* provide confirmed novel predictions. As Carl Sagan once said: “extraordinary claims require extraordinary evidence.” After all, if a theory posits the existence of superstrings, the only way to convince the community is by confirming some of the novel predictions that follow from the theory. Here is the summary of our expectations in a flow-chart:



This is essentially the gist of the *hypothetico-deductive method* that we seem to employ nowadays:

Hypothetico-Deductive Method

A hypothesis is allowed to introduce *unobservable entities* (e.g. particles, forces, superstrings, etc.) provided that it predicts something *novel*, hitherto unobserved, and some of these novel predictions are *confirmed*.

Since this is an explication of the implicit expectations of scientific communities, we have to keep in mind that it may or may not be correct. It is our historical *hypothesis* that the requirements of the hypothetico-deductive method have actually been employed in physical sciences since the early 1700s.

Fixed Method?

Let us assume for the sake of argument that this historical hypothesis is correct. A question arises: is this the *fixed* method of science? Is it the case that in all historical periods and all fields of inquiry, new theories have always had to satisfy the requirements of this method in order to become accepted? More generally, can we claim that our expectations towards new theories haven't changed ever since antiquity? In short:

Is there an unchangeable (fixed, transhistorical) method of science?

An analysis of historical episodes reveals that the hypothetico-deductive method has *not* always been employed in theory evaluation. At best, we can show that the method has been employed in most physical sciences since around 1700. If we went back to the times of Aristotelian-Medieval science, we would notice that the expectations of that community had little in common with the requirements of the hypothetico-deductive method. So, what were their expectations?

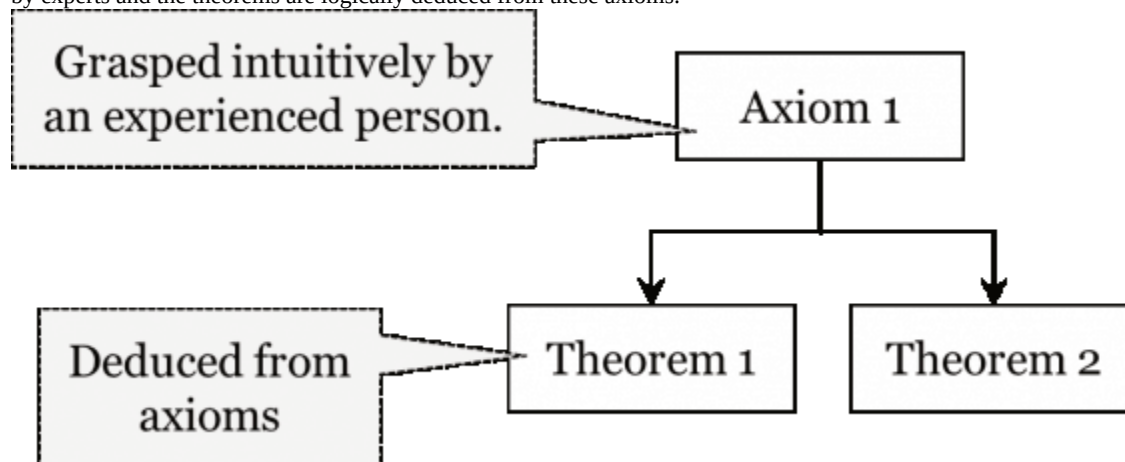
While explicating the exact expectations of a community is never an easy task, we can reasonably claim that in the Aristotelian-Medieval worldview, a theory/proposition was acceptable if it grasped the nature of a thing through intuition schooled by experience, or if it followed deductively from propositions that seemed intuitively true:

Aristotelian-Medieval Method

A proposition is acceptable if it grasps the nature of a thing through intuition schooled by experience or is deduced from the general intuitive propositions.

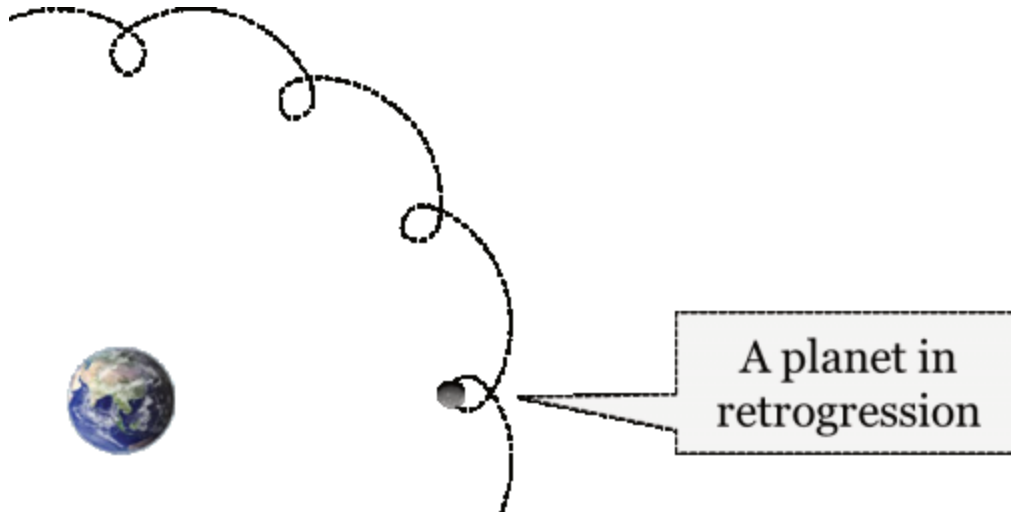
Now, how can someone's intuition be "schooled by experience"? The underlying idea here is straightforward: the more time one spends studying a certain phenomenon, the better one is positioned to reveal the nature of that phenomenon. A beekeeper that has spent a lifetime around bees is assumed to be in a position to tell us what the nature of bees is. Similarly, an experienced arborist will have developed an intuition regarding the nature of different trees – an intuition that has been schooled by her experience. Finally, a natural philosopher who has carefully considered different types of things is best positioned to know what these things have in common; e.g. she can grasp intuitively that everything is made out of four elements (*earth, water, air, fire*) and that these elements tend towards their natural positions – heavy elements tend towards the centre of the universe and light elements tend towards the periphery. In Aristotelian-Medieval science, these expectations applied to any field of study: a theory would be acceptable if it appeared intuitively true to those schooled by experience, i.e. experts.

In the Aristotelian-medieval worldview, the basic intuitions about the nature of things would be considered the *axioms* of their science. Once the nature of a thing under study was grasped intuitively by an expert, we would then proceed to tracing the logical consequences (i.e. theorems) of our basic intuitions. For example, once it is grasped that heavy elements tend towards the centre of the universe, we can deduce a theorem that the Earth, which is predominantly a combination of the elements *earth* and *water*, is located at the centre of the universe. As a result, we have an axiomatic-deductive system, where the axioms are grasped intuitively by experts and the theorems are logically deduced from these axioms.

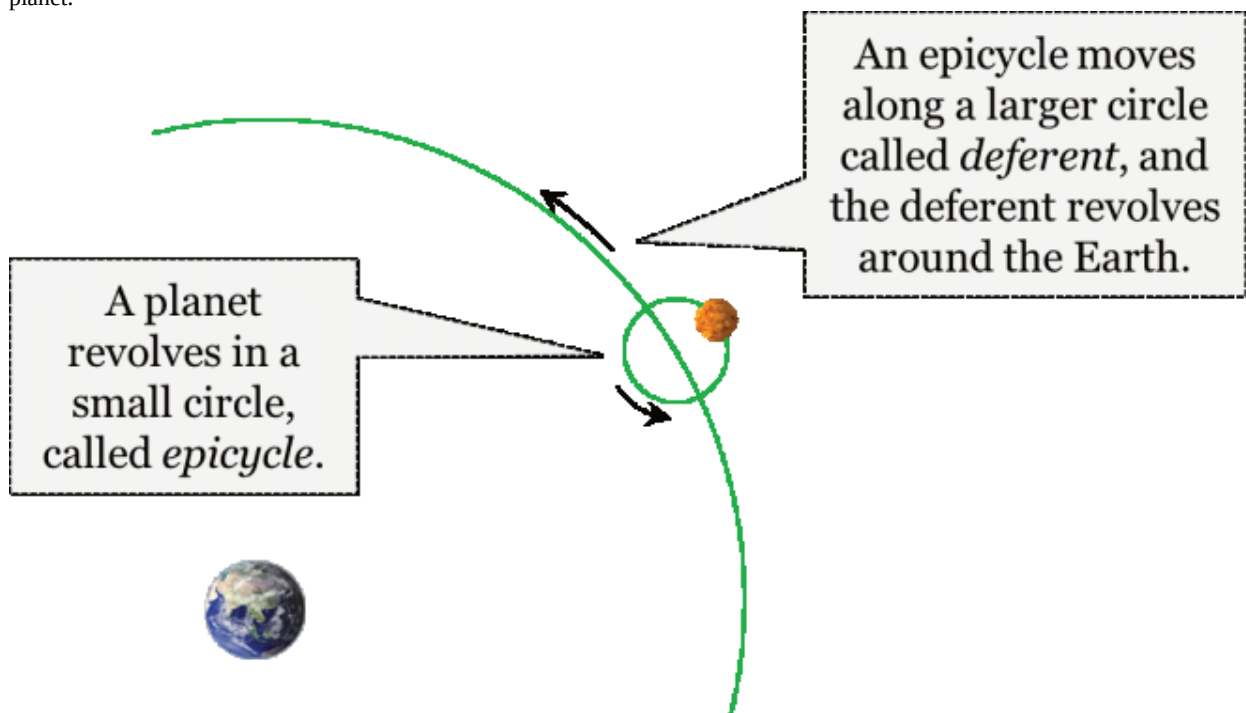


Importantly, the method of the Aristotelian-medieval community had nothing to do with confirmed novel predictions; the community expected basic intuitions schooled by experience and deductions from these intuitions. Thus, trying to convince that community by an appeal to confirmed novel predictions would be a fool's errand.

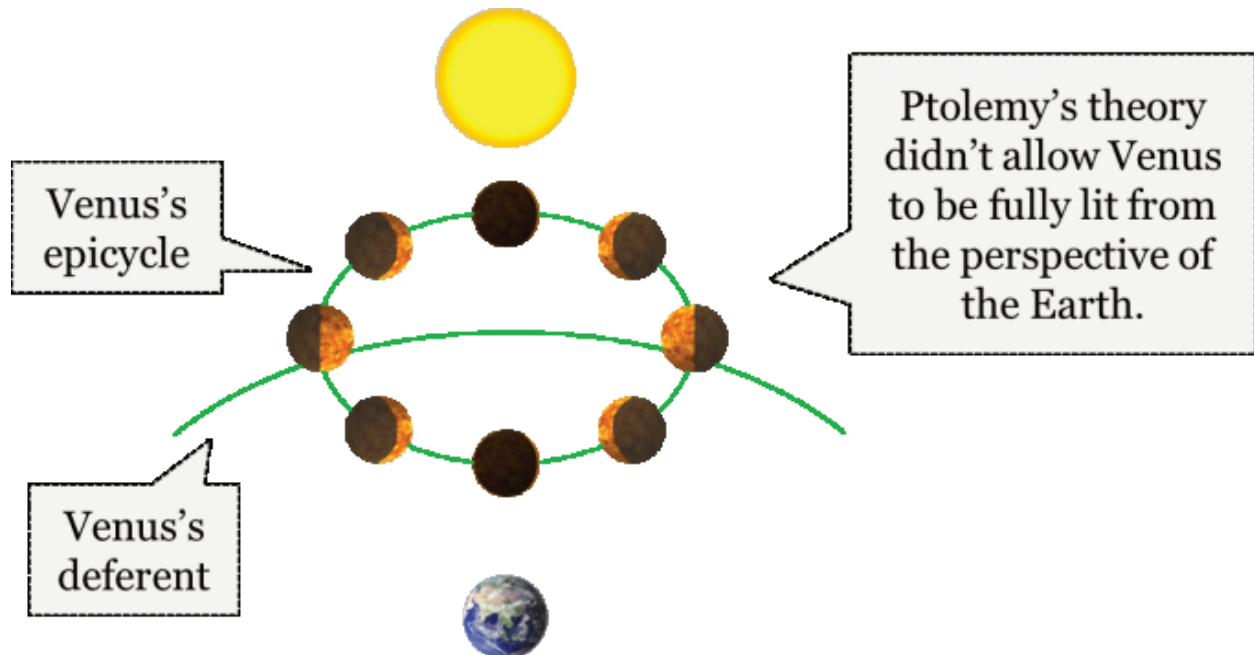
The famous case of Galileo is a great illustration of the Aristotelian-medieval method in action. If we went back to the year 1600, we would discover that the accepted view was that of Ptolemaic geocentrism: the Earth was believed to be at the centre of the universe and all planets, the Moon, and the Sun were believed to revolve around the Earth. From the Earth, the observable motion of planets can seem rather strange against the relatively uniform motion of the stars. While the stars seem to always travel in the same direction and at the same rate, a dedicated observer will notice that planets seem to travel at different speeds, and sometimes move in reverse – what has been called *retrograde motion*.



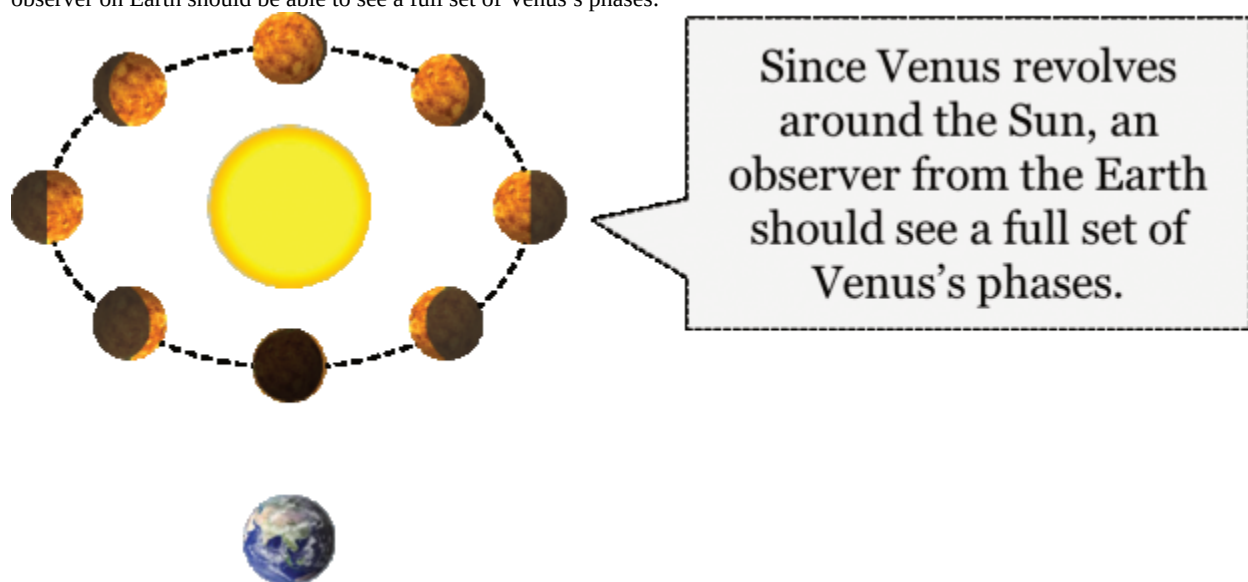
This meandering across the sky is actually what earned them the name planet: recall that the Greek word *planētēs* means *wanderer*. For Ptolemy, all heavenly bodies rotated around the Earth in circular paths that were called *deferents*. To account for the planets' easily observable retrograde motion, Ptolemy suggested that planets moved in an additional, smaller circular orbit along their deferent, called the *epicycle*. Such combinations of epicycles and deferents helped reproduce the exact motions of every planet.



In fact, the theory was extremely successful in its predictions of future positions of planets. Since, in this theory, Venus revolves around the Earth, it was believed that an observer on Earth could never see Venus fully lit; we could, at best, see it half-lit:



A few alternative cosmological theories were pursued at the time, including the heliocentric theory of Copernicus, championed famously by Galileo. In the Copernican heliocentric theory, Venus and the other planets all revolve around the Sun and, thus, an observer on Earth should be able to see a full set of Venus's phases:



This was one of the novel predictions of the Copernican theory. It was not until the invention of the telescope that it became possible to test it. In 1610, Galileo famously tested this prediction and confirmed that we can see a full set of phases of Venus. If only the community of the time cared about confirmed novel predictions, then the Copernican theory would've become accepted. Unfortunately for Galileo, the community of the time didn't care for confirmed novel predictions, as they expected new theories to be commonsensical, i.e. to appear intuitively true to experts. But the idea of the Earth being anywhere but at the centre of the universe was anything but commonsensical. At the time, the idea of the Earth's revolution around the Sun and its diurnal rotation around its own axis appeared equally counterintuitive. Not only was it against our everyday observations that seemed to suggest that the Earth is static, but it also contradicted the laws of the then-accepted Aristotelian natural philosophy which entailed that because the Earth is made of the elements *earth* and *water* it could not possibly revolve or rotate. It is not surprising, therefore, that Galileo failed to convince the community of his time. Importantly, his failure had nothing to do with the alleged dogmatism and obstinacy of the university-educated Aristotelian clergy of the time, but with the fact that the expectations of the community had nothing to do with observational confirmations of novel predictions. It is as though Galileo was trying to beat everyone in chess when others were playing checkers.

It took the Copernican theory almost another century to become accepted. But how did it become accepted? The short answer is

that it was incorporated into a more general system of the world which finally managed to meet the expectations of the community. The author of that system was René Descartes, whose views we will revisit on several occasions, specifically in chapter 8. Descartes devised a system of the world with the goal of meeting the requirements of the Aristotelian-Medieval community. The most comprehensive exposition of this new system was presented in his *Principia Philosophiae* (Principles of Philosophy), first published in 1644. Descartes clearly realized that if his theory was going to succeed it had to strike the experts as intuitively true. But, says Descartes, if it is intuitive truth we are after, wouldn't it be more reasonable to start by erasing everything we think we know and then proceed to accepting only those theories that are established beyond any reasonable doubt? That's exactly what Descartes set off to accomplish.

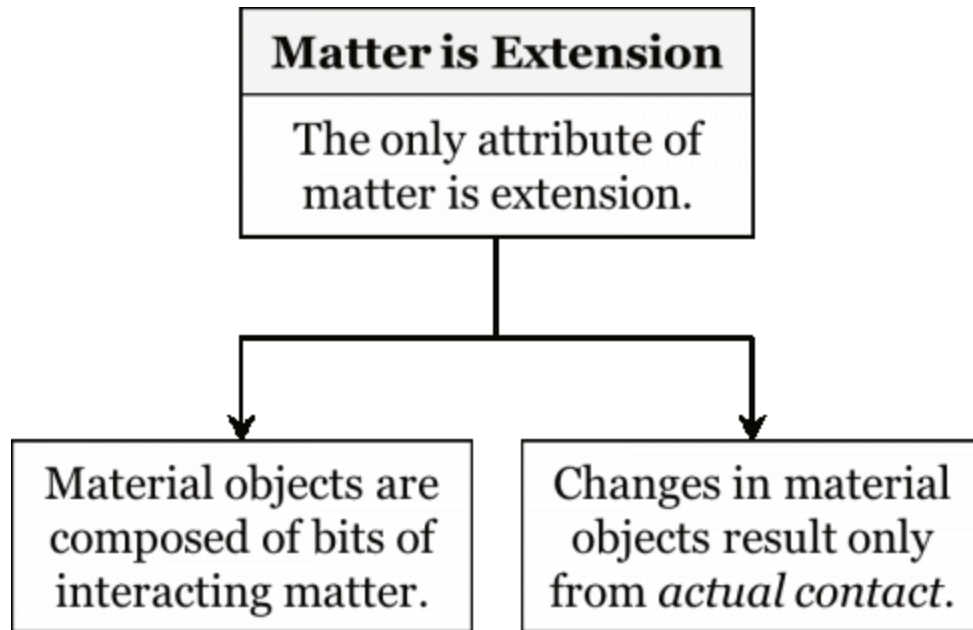
Among many other topics, Descartes also attempted to reveal the attributes of matter, i.e. those qualities of material objects that are indispensable. Take any material object: a rock, a plant, an animal, a human body – anything. Question: what are the qualities that any material object must necessarily have? A material object can have some *colour*, *sound*, *taste*, *smell*, and *shape*. Of these usual suspects, which are truly indispensable? For instance, can we think of a material thing that doesn't have any colour? Yes, says Descartes, it is possible: we can think of something material yet transparent, such as air. Therefore, colour is not an indispensable property of material objects. But how about sound: can we conceive of a material object that doesn't make any sound? Clearly, we can do that. Therefore, according to Descartes, sound is not indispensable. The same goes for taste: there are many objects that don't have any taste. Moreover, there are some foods that are supposed to have taste but don't seem to have any (I am looking at you, American apricots)! Thus, taste is not an indispensable quality of matter. What about smell? Clearly, not everything smells in our world, so smell is yet another dispensable quality. This only leaves us with shape. Descartes argues that we cannot possibly conceive of a material object that does not occupy some space. A material thing can have a well-defined shape, as is the case for most of the solid objects around us, or it could have a pretty fuzzy shape, as is the case with water, fire, or air. But in any case, all material objects necessarily occupy some *space*, i.e. they are all spatially *extended*. Indeed, can we conceive of a material object that doesn't occupy any space whatsoever, even a very minute space? That's impossible, says Descartes; if you happen to come across a material object that doesn't occupy any space whatsoever, then it's not much of a material object, is it? Recall that infamous cheese shop from the Monty Python sketch that didn't have any cheese: it wasn't much of a cheese shop, was it? Similarly, for a thing to be material, according to Descartes, it has to be spatially extended, i.e. it must occupy *some* space. This brings Descartes to the formulation of one of his fundamental principles – the idea that the only attribute of matter is extension:

Matter is Extension

The only attribute of
matter is extension.

Once again, by *attribute*, Descartes means *an indispensable quality*.

Once we accept this principle, a number of notable conclusions logically follow. One such conclusion is the idea that material objects should be composed of bits of interacting matter, each spatially extended. Indeed, says Descartes, if all material objects can do is occupy some space, then everything in the universe is composed of smaller bits of matter that also occupy some space and interact with each other. But how can two bits of matter possibly affect each other if all they can do is occupy space? Consider two billiard balls. How can one billiard ball possibly affect another billiard ball? Their *actual contact* seems to be the only way in which they can interact. According to Descartes, that's exactly the case with *all* interactions between material objects; because they all are spatially extended bits, they can only affect each other by touching and pushing. That requires actual contact; for Descartes, there can be no such thing as *action at a distance*.



While these conclusions may sound trivial, they are the cornerstones of the *mechanistic Cartesian worldview* that we will study in chapter 8.

Importantly, Descartes wanted to persuade his Aristotelian peers that these conclusions are nothing but intuitive truths, just as Aristotle himself would have wanted. Isn't it intuitive, says Descartes, that the only attribute of matter is extension? It is common sense! So, if we are expecting commonsensical, intuitively true axioms, we have them. In essence, Descartes was claiming that his theory satisfies the Aristotelian-Medieval requirements better than Aristotle's theory itself!



The difference between Descartes' strategy and that of Galileo is apparent. Where Galileo was trying to convince the community by quoting the results of his observations, Descartes knew that the only way he could convince the Aristotelians was through meeting *their* expectations. So, it shouldn't come as a surprise that it was Descartes' theory that became accepted on the Continent ca. 1700. The key point is that it was accepted not because it provided confirmed novel predictions, but because it appeared intuitively true to the community of the time. What we need to appreciate is that the method of the time was very different from the hypothetico-deductive method we employ nowadays.

Summary

To answer the central question of this chapter: it is accepted nowadays that there is no fixed (unchangeable, transhistorical) method of science. The historical record shows that, when studying the world, we transition not only from one theory to the next, but also, importantly, from one method of theory evaluation to the next. In other words, we nowadays reject the idea of a *static method* and accept *the dynamic method thesis*, thus:

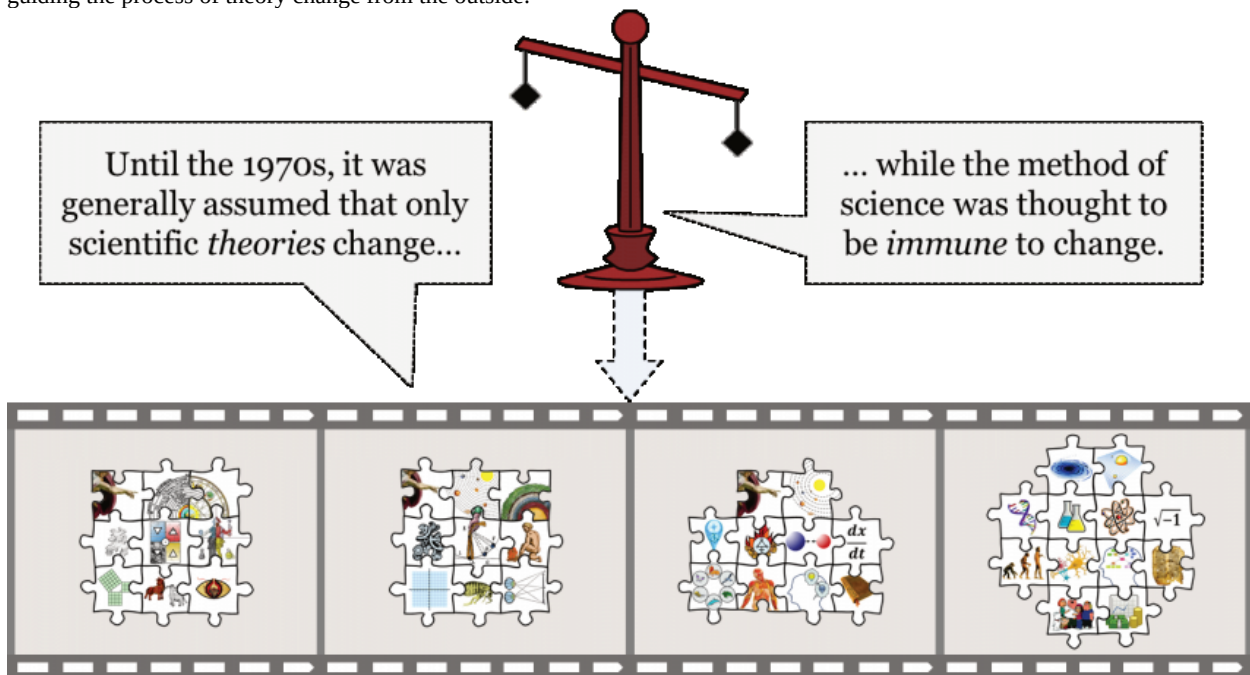


Is there an unchangeable method of science?

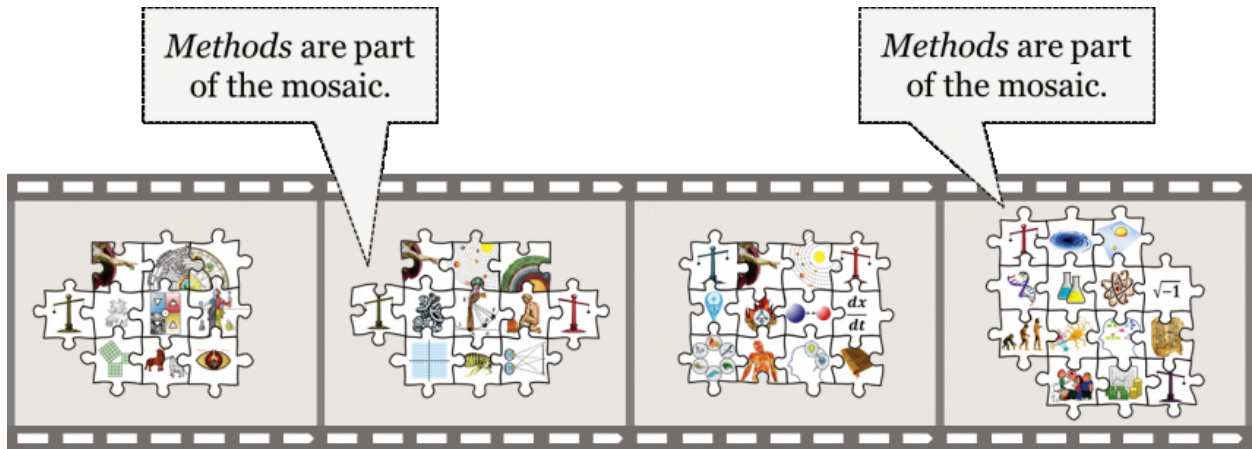
Yes	No
<p>Static Method thesis: The method of science is unchangeable.</p> <p> Historical record (e.g. Aristotelian method vs. HD method)</p>	<p>Dynamic Method thesis: There is no unchangeable method of science.</p> <p> Historical record (e.g. Aristotelian method vs. HD method)</p>

For most of the history of knowledge the static method thesis was taken for granted. While Aristotle, Newton, Kant, and Popper would never agree as to what the requirements of the method are, they would all agree that there is one set of requirements that any acceptable theory should meet. The transition to the dynamic method thesis took place ca. 1980 and was mostly due to the pioneering work of Thomas Kuhn and Paul Feyerabend as well as a number of historians of science, who showed that methods of theory evaluation often change as we learn new things about the world. Nowadays, it is commonly accepted that there is no such thing as a fixed method of science; methods change.

This transition seriously reshaped our views on the process of scientific change. Until about 1980, it was assumed that the process of scientific change concerns only *theories*, while methods were thought to be *external* to the process, as if they were guiding the process of theory change from the outside:



Our current view is different, for we no longer think that methods are external to the process of scientific change. In fact, they are part of the process of scientific change:



But if we understand that methods are part of the process of scientific change, we should redefine our notion of *scientific mosaic*. We originally defined it as a set of all accepted theories. Now the notion of scientific mosaic should also include all methods employed by the community at a certain time:

<p>Scientific Mosaic \equiv</p>
<p>A set of all accepted <i>theories</i> and employed <i>methods</i>.</p>

Note that just as we accept multiple theories at the same time, it is possible for multiple methods to be employed in the same mosaic. For instance, our method of drug testing will likely be different than our method for evaluating the efficacy of surgical techniques. Similarly, the specific method for evaluating hypotheses concerning subatomic particles need not be the same as the method for evaluating hypotheses concerning the existence of different biological species. In short, a community may have different expectations regarding hypotheses that pertain to different domains.

Thus, we arrive at what is arguably the most challenging question in contemporary philosophy of science. If there are no fixed methods of theory evaluation, does it mean that the process of scientific change is *irrational*? In other words, why is it that nowadays we employ the hypothetico-deductive method and not, say, the Aristotelian-Medieval method? Is the choice of methods arbitrary or is there a certain *mechanism* that governs the process of transitions from one employed method to the next? If it turns out that the choice of methods is random – if anyone can freely choose her own method of theory evaluation – then how can we reasonably argue that our contemporary science is better than the science of Aristotle, Descartes, or Newton? If we were to concede that there is no mechanism guiding transitions from one method to the next, then there would be no way of reasonably arguing that one method is better than another and we would end up in what philosophers call *relativism*. So, is there a mechanism that guides the process of changes in theories and methods alike? We will tackle this question in the next chapter.

4.

The Laws of Scientific Change -- Introduction to History and Philosophy of Science

Chapter 4: The Laws of Scientific Change

Intro

So far, we have given a number of examples of how both accepted theories and methods of theory evaluation have changed through time. These changes portray scientific mosaics as dynamic and seemingly always in flux. So, you might be wondering: is there some underlying universal mechanism that governs the changes in theories and methods that we see in scientific mosaics? Or is it the case that theories and methods change in a completely random fashion, without exhibiting any general patterns? To answer that question, we have to first tackle the broader philosophical question of whether a general theory of scientific change is possible at all. Doing so will require making a number of clarifications from the outset.

When we consider any particular science, there are two relatively distinct sets of questions that we can ask: particular questions and general questions. *Particular questions* ask about specific events, data points, phenomena, or objects within a certain science's domain. Particular questions in a field like ornithology (the study of birds) might be:

Is that blue jay a male or a female?

How many eggs did that snowy owl lay this season?

Did the Canada geese begin flying south late this year?

In contrast, *general questions* ask about regular patterns, or about *classes* of events, phenomena, or objects in a certain science's domain. General questions in ornithology (corresponding with the particular questions above) might be:

What are the physical or morphological differences between male and female blue jays?

What environmental factors affect snowy owl egg production?

What are the typical migration patterns of Canada geese?

Often, answers to particular questions are relatively straightforward theories:

The blue jay in question is a male.

The owl laid a clutch of 3 eggs.

The Canada geese left late.

But interestingly, these questions often depend upon answers to the general questions we asked above. In order to determine whether the blue jay is male or female, for instance, the ornithologist would have to consult a more general theory about typical morphological sex differences in blue jays. The most apt general theories seem to emerge as answers to general questions from within a specific scientific domain.

Now, it is possible – and sometimes necessary – to draw from the general theories of other sciences in order to answer particular questions. For instance, if we did not yet have a general theory for determining the sex of a blue jay, we might rely upon general theories about sexual difference from generic vertebrate biology. That said, while helpful, general theories from related disciplines are often vague and more difficult to apply when one is trying to answer a very specific, particular question.

The question “Can there be a general theory of scientific change?” is therefore asking whether we can, in fact, unearth some general patterns about the elements of science and scientific change – the scientific mosaic, theories, methods, etc. – and whether it's possible to come up with a general theory about them, on their own terms. We can clearly ask particular questions about different communities, their mosaics, and changes in them:

What different scientific communities existed in Paris ca. 1720?

What version of Aristotelian natural philosophy was accepted in Cambridge around 1600?

What method did General Relativity satisfy to become accepted?

Yet, we can also ask general questions about the process of scientific change, such as:

What is the mechanism by which theories are accepted?


What is the mechanism of method employment?

But are we in a position to answer these general questions and, thus, produce a general theory of scientific change? Or are particular questions about scientific change answerable by general theories from other fields, like anthropology or sociology?

Generalism vs. Particularism

Let's frame the debate with two opposing conceptions, which we will call generalism and particularism: does the process of changes in theories and methods exhibit certain general patterns? Answering “yes” to this question would make one a supporter of the conception we've called *generalism*. Generalists believe that there is some sort of an underlying mechanism that governs transitions from one theory to the next and one method to the next and, therefore, a general theory of scientific change is possible,

in principle. To answer “no” to this question would make one a supporter of what we’ve called *particularism*. Particularists believe that there can be no general theory of scientific change.

 Can there be a <i>general</i> theory of scientific change?	
Yes	No
Generalism: There can be a general theory of scientific change.	Particularism: There can be no general theory of scientific change.

It should be noted at the outset that there is *no communal* consensus among historians and philosophers of science regarding the issue of particularism vs. generalism. That said, particularism, *not generalism*, has more widespread acceptance among *individual* historians and philosophers of science nowadays. They take this position for a number of reasons, four of which we will unpack below.

Argument from bad track record: There have been a number of attempts to fashion theories that could answer general questions about scientific change. Despite the fact that some of these theories were accepted by individual historians and philosophers of science, none have ever become accepted by the *community* of historians and philosophers of science. While these theories were often ingenious, one of their principal drawbacks has been their failure to account for the most up-to-date knowledge from the history of science. Consider, for instance, Karl Popper’s conception of scientific change, according to which the process of scientific change is essentially a series of bold conjectures and decisive refutations. In this Popperian picture of science, a theory is accepted only provisionally as the best available hypothesis and is immediately rejected once we observe an anomaly, i.e. something that goes against the predictions of the theory. However, historical studies show that scientific communities often exhibit remarkable tolerance towards anomalies. For example, in 1859, it was discovered that the observed orbit of the planet Mercury deviates slightly from the one predicted by the then-accepted Newtonian theory. This was clearly an anomaly for the accepted theory and if science functioned according to Popper’s dicta, the Newtonian theory would have to be rejected. Yet, it is a historical fact that the Newtonian theory remained accepted for another 60 years until it was replaced by general relativity ca. 1920.

These historical failings are explained slightly more thoroughly in points 2 and 3 below. The “argument from bad track record” considers these failed attempts all at once. In short, every general theory of scientific change suggested so far has been left unaccepted. The conclusion of the argument, therefore, is an inductive inference: since every attempt to craft a general theory has ended up unaccepted, it seems likely that no general theory will ever be accepted. This argument isn’t the strongest, since it makes a rather pessimistic inference drawn from very little data, but it is an argument against generalism nonetheless.

Argument from Changeability of Scientific Method: Previous attempts to propose a general theory of scientific change tried to ground those theories in a single set of implicit expectations and requirements which all scientific communities had shared throughout history. That is, philosophers of science like Carnap, Popper, Lakatos and others between the 1920s and the 1970s tried to explicate a single method of science common to all scientific communities, past, present, and future. If we figure out what that method is – they reasoned – we can point to a set of unchanging requirements or expectations to help us answer the general question “how do new scientific theories get accepted?”

But as we know from chapter 3, a more careful historical analysis reveals that there has been no single, transhistorical or transdisciplinary method of science. Scientific methods, in fact, do change through time. If answers to general questions about scientific change depend upon a static method of science, and the method of science is *dynamic*, then when faced with the question “Can there be a general theory of scientific change?” we would seem compelled to respond “no”.

Argument from Nothing Permanent: One can take this line of reasoning from the changeability of the scientific method a step further. The efforts of historians of science have demonstrated that reconstructing the theories and methods held by any scientific community involves also accounting for an incredibly complex set of social, political, and other factors. For instance, to even begin to reconstruct the theories and methods of the scientific community of the Abbasid caliphate, which existed on the Arabian peninsula between 750–1258 C.E., a historian must be familiar with that particular society’s language, authority structure, and what types of texts would hold the accepted theories of that time. All of these would be starkly different from the factors at play in Medieval Europe during the same period! This is all to say that in addition to showing that there are diverse scientific *methods* across geography and time, it seems that almost anything can vary between scientific communities. Across time and space, there might be *nothing permanent* or fixed about scientific communities, their theories, or their methods.

The claim here is that, because neither methods nor nearly *anything else* can ground patterns of change at the level of the mosaics of scientific communities, there’s no way that we could have a general theory of scientific change, nor general answers to the general questions we might ask about it. Those who marshal this argument therefore eschew any general approach to science

and instead focus on answering particular questions about science: questions about the uniqueness and complexity of particular scientific communities or transitions within particular times and places. Often, these historical accounts will be communicated as a focused story or *narrative*, giving an exposition of a particular historical episode without relying on any general theory of scientific change.

Argument from Social Construction: While historical narratives might be able to avoid invoking or relying upon a general theory of scientific change, they cannot avoid relying upon general theories of some sort. Even the most focused historical narratives invoke general theories in order to make sense of the causes, effects, or relationships within those episodes. It is accepted by historians that Queen Elizabeth I cared deeply for Mary, Queen of Scots. The traditional explanation for this is that they were cousins. Clearly this explanation is based on a tacit assumption, such as “cousins normally care for each other”, or “relatives normally care for each other”, or even “cousins in Elizabethan England cared for each other”, etc. Regardless of what the actual assumption is, we can appreciate that it is a *general* proposition. Such general propositions are an integral part of any narrative.

For a large swath of contemporary historians, the constitution and dynamics of the scientific mosaic can be explained by referencing the general theories of sociology and anthropology. In other words, the patterns and regularities needed to have such general theories can be found in the way that communities behave, and how individuals relate to those communities. Since there are good arguments for why scientific change lacks these patterns, and since there is no accepted general theory that identifies and explains them, many philosophers and historians of science see the dynamics of scientific theories and methods as social constructs, and thus best explained by the best general theories of sociology and anthropology. So, the argument from “social construction” is slightly more nuanced than the arguments that have come before it. It claims that particular changes in science can be *best* explained with general theories from sciences like sociology and anthropology, *not* general theories of scientific change which have typically failed to capture the nuances of history. In short, many philosophers and historians are doubtful about the prospects of a general theory of scientific change.

But what if there were a general theory which offered rebuttals to all four of the previously mentioned arguments for particularism? What if there were a general theory of scientific change which could accommodate the nuances and complexities of historical-geographic contexts, and also managed to recognize and articulate constant features in the historical dynamics of scientific mosaics?

The remainder of this chapter is dedicated to introducing a general theory of scientific change which proposes four *laws* which can be said to govern the changes in the scientific mosaic of any community. This general theory has formed the centerpiece of a new empirical science called *scientonomy*, where it is currently accepted as the best available description of the process of scientific change. However, note that from the viewpoint of the larger community of historians, philosophers, and sociologists of science this theory is not currently accepted but merely *pursued*. In the process of explaining this new general theory of scientific change, we will labour to demonstrate how this general theory rebuts all of the major arguments used to support particularism. Hopefully it will become clear by the end of the chapter why the development and refinement of this theory is worth pursuing, and perhaps even why we might be able to answer “yes” to the question “can there be a general theory of scientific change?”

Method Employment

To begin, let’s do a thought experiment based on the history of medical drug testing. Suppose that you and a group of other scientists want to determine whether a certain medicine can help people with insomnia fall asleep more easily. Let’s call the medicine you’d like to test *Falazleep*. What you are ultimately doing is seeing whether the *theory* “Falazleep helps individuals with insomnia fall asleep” is acceptable to the scientific community or not.

One of the most straightforward ways of seeing whether Falazleep can help individuals go to sleep more easily is to simply give the proper dose of Falazleep to people who suffer from *insomnia*. If some of those people report being able to sleep more easily, then it seems reasonable for the scientific community to accept the theory that *Falazleep is therapeutically effective*, i.e. that it works, at least in certain cases. What we’ve articulated here is a very basic version of the hypothetico-deductive (HD) method: in order to accept this causal relationship between Falazleep and sleep, you have to *confirm* your theory’s *prediction* that, when given to insomniacs, Falazleep will indeed help them sleep.

You and your peers give a group of insomniacs Falazleep and, sure enough, many of them do in fact fall asleep more easily. Having confirmed your theory’s prediction, you and the scientific community accept the theory that Falazleep helps individuals with insomnia fall asleep.

A few years later, you and your peers are asked to test a new drug, *Falazleep Z*. However, the scientific community has come to the realization that other factors – such as the recipients’ exercise, diet, stress level, and so forth – could have actually been the cause of the reported easier sleep in your previous trial, *not* Falazleep itself. That is, you and the scientific community now accept the theory of *unaccounted effects*, which says “factors other than the medication may influence the outcome of a medical drug test study.” Knowing what you know now, can you and your peers test Falazleep Z the same way you tested the original Falazleep? No!

Why not? Because now you and the rest of the scientific community accept that your earlier test – simply giving Falazleep to insomniacs – didn’t even begin to distinguish between the effects of Falazleep and the other variations in the test subjects’ lives. In order for your new test results to be accepted, you have to meet the scientific community’s new expectation: that you find a way to distinguish between the effects of Falazleep Z and other variations in the test subjects’ lives.

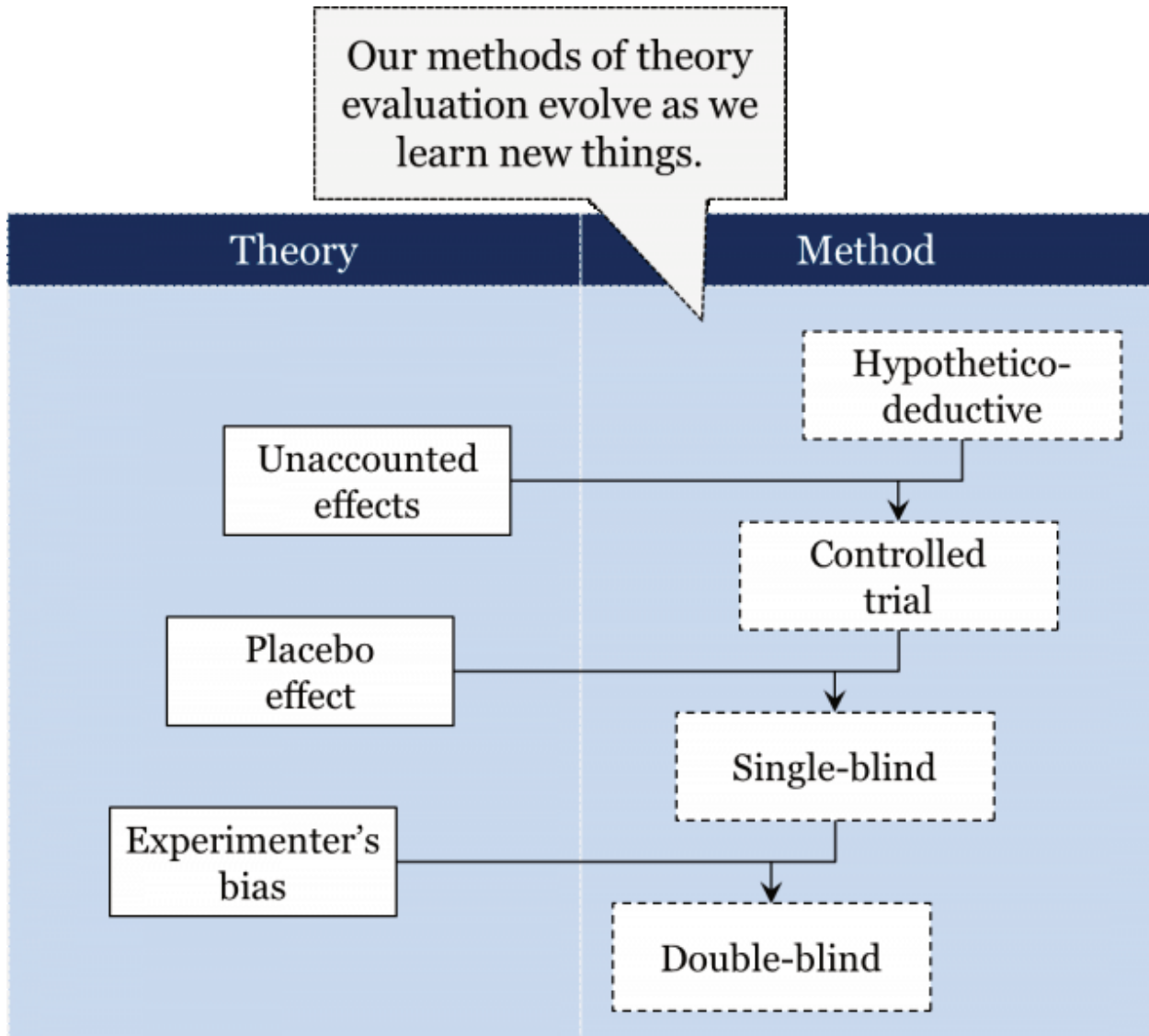
So, you and your colleagues take a large group of insomniacs and split them in two groups, each of which you study for the duration of the trial. However, only one group – the so-called “active group” – actually receives Falazleep Z. The other group, which we will label the “control group”, receives no medication. At the end of the trial, if the active group’s improvement *with* Falazleep Z is significantly greater than the improvement in the control group, then perhaps the scientific community will accept that Falazleep Z is therapeutically effective. This is especially likely because the experimental setup, today called a “controlled trial,” accounts for previously unaccounted effects, such as the body’s natural healing ability, improved weather, diet, and so on.

A few more years go by, and you and your peers are asked to test another drug, *Falazleep Zz*. But in the meantime, it has been found that medical drug trial participants can sometimes show drastic signs of improvement simply because they *believe* they are receiving medication, even if they are only being given a fake or substitute (candy, saline solution, sugar pills, etc.). In other words, the scientific community accepted a theory that *beliefin effective treatment* might actually produce beneficial effects, called the *placebo effect*. Perhaps some of the efficacy of the Falazleep Z trial from a few years back was because some participants in the active group *believed* they were getting effective medicine...whether the medicine was effective or not!

So, in your trial for Falazleep *Zz*, you need to now account for this newly discovered placebo effect. To that end, you and your peers decide to give *both* your active group *and* your control group pills. However, you only give Falazleep *Zz* to the active group, while giving the control group sugar pills, i.e. *placebos*. Importantly, only you (the researchers) know which of the two groups are *actually* receiving Falazleep *Zz*. Because both groups are receiving pills, but neither group knows whether they’re receiving the actual medicine, this experimental setup is called a *blind trial*. Now suppose the improvement in the condition of the patients in the active group is noticeably greater than that of the patients in the control group. As a result, the scientific community accepts that Falazleep *Zz* is therapeutically efficient.

Another few years pass, and you and your group test the brand-new drug *FASTazleep*. Since your last trials with Falazleep *Zz*, the scientific community has discovered that researchers distributing those pills to the control and active groups can give off subtle, unintended cues as to which pills are genuine and which are fake. Unsurprisingly this phenomenon, called *experimenter’s bias*, prompts test subjects in the active group who have picked up on these cues to improve, regardless of the efficacy of the medication. Because you and your peers accept this theory of experimenter’s bias, you know the scientific community expects a new theory to account for this bias. Consequently, you devise a setup in which neither those administering the pills nor those receiving the pills actually know who is receiving *FASTazleep* and who is receiving a placebo. In other words, you perform what is nowadays called a *double-blind trial*.

Hopefully this repetitive example hasn’t put you “*FASTazleep*”! The repetition was purposeful, it was meant to help elucidate the pattern found not only in the actual history of drug trial testing, but in the employment of new methods in science generally. As you may have noticed, when the scientific community accepted a new theory – say, the existence of the placebo effect – their implicit expectations about what would make a new theory acceptable *also* changed. Remember, we’ve called these implicit expectations and requirements of a scientific community their *method*. Thus, we can conclude that changes in our criteria of theory evaluation, i.e. in our methods, were due to changes in our accepted theories:



That new accepted theories cause employed methods to change is the fundamental insight which is the essence of a *law of scientific change*, which we've called the law of method employment. According to the law of method employment, a method becomes employed only when it is deducible from some subset of other employed methods and accepted theories of the time:

3rd Law: Method Employment

A method becomes employed only when it is deducible from some subset of other employed methods and accepted theories of the time.

This law captures a pattern, a regularity that occurs time and again in the history of science when a community comes to employ a new method. Indeed, it seems to successfully describe the mechanism of both minor changes in employed methods, such as any of the examples from medical drug trial testing above, and even more substantial changes in employed method, such as the shift from the Aristotelian-Medieval (AM) to the hypothetico-deductive (HD) method. Importantly, the third law describes not only *that* new methods are shaped by new theories, but *how* those methods are shaped by those theories. As you can see in the formulation

of the law, a new method is employed only if that method is a *deductive consequence* of, i.e. if it logically follows from, some other theories and methods which are part of that scientific community's mosaic.

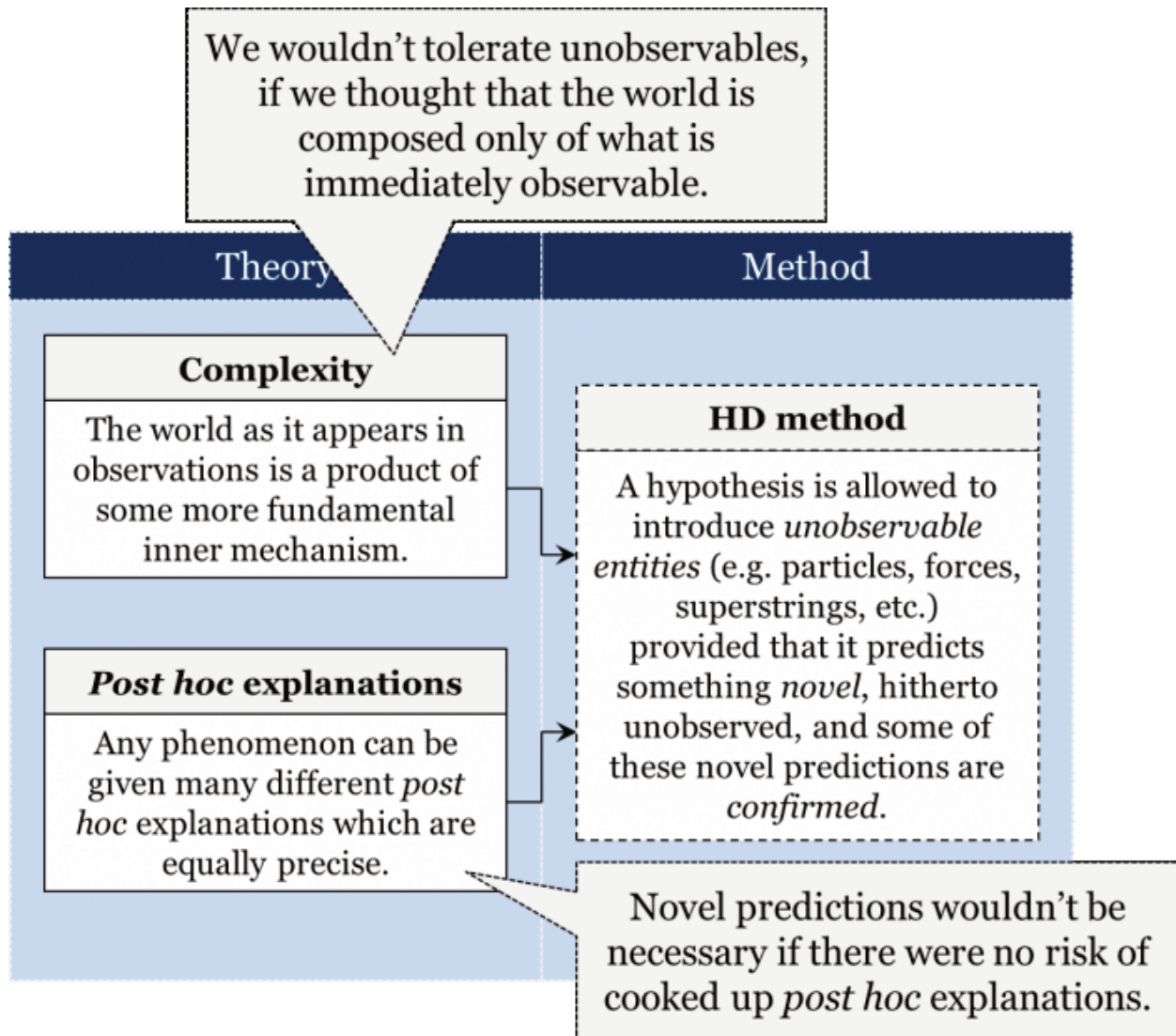
Now that we know the mechanism of transitions from one method to the next, we can appreciate *why* methods are changeable. In chapter 3 we provided a *historical* argument for the dynamic method thesis by showing how methods of theory evaluation were different in different historical periods. Now we have an additional – *theoretical* – argument for the dynamic method thesis. On the one hand, we know from the third law that our employed methods are deductive consequences of our accepted theories. On the other hand, as fallibilists, we realize that our empirical theories can change through time. From these two premises, it follows that all methods that assume at least one empirical theory are, in principle, changeable.

This law also helps us cast light on what features of historical mosaics led to the employment of certain methods at certain times. Consider for instance the assumptions underlying the *hypothetico-deductive method*. It is safe to say that the HD method is a deductive consequence of very specific assumptions about the world.

One such assumption is the idea of *complexity*, according to which the world, as it appears in observations, is a product of some more fundamental inner mechanism. This assumption has been tacitly or explicitly accepted since the early 18th century. It is because of this assumption that we tolerate hypotheses about the existence of unobservable entities, such as elementary particles, waves, forces, etc. The reason why we tolerate such hypotheses is that we accept the possibility that there might be entities and relations that are not directly observable. In other words, we believe that there is more to the world than meets the eye. We wouldn't tolerate any hypotheses about unobservables if we believed that the world is composed only of immediately observable entities, properties, relations, etc.

The second assumption underlying the HD method is the idea that any phenomenon can be given many different *post hoc* explanations which are equally precise. We've already witnessed this phenomenon. Recall, for example, different theories attempting to explain free fall, ranging from the Aristotelian "descend towards the centre of the universe" to the general relativistic "inertial motion in a curved space-time". But since an explanation of a phenomenon can be easily cooked up after the fact, how do we know if it's any good? This is where the requirement of confirmed novel predictions enters the picture. It is the scientists' way of ensuring that they don't end up accepting rubbish explanations, but only those which have managed to predict something hitherto unobserved. A theory's ability to predict something previously unobserved is valued highly precisely due to the risk of post hoc explanations. Indeed, novel predictions wouldn't be necessary if there were no risk of cooked up *post hoc* explanations.

In short, just like the drug testing methods, the HD method too is shaped by our knowledge about the world.

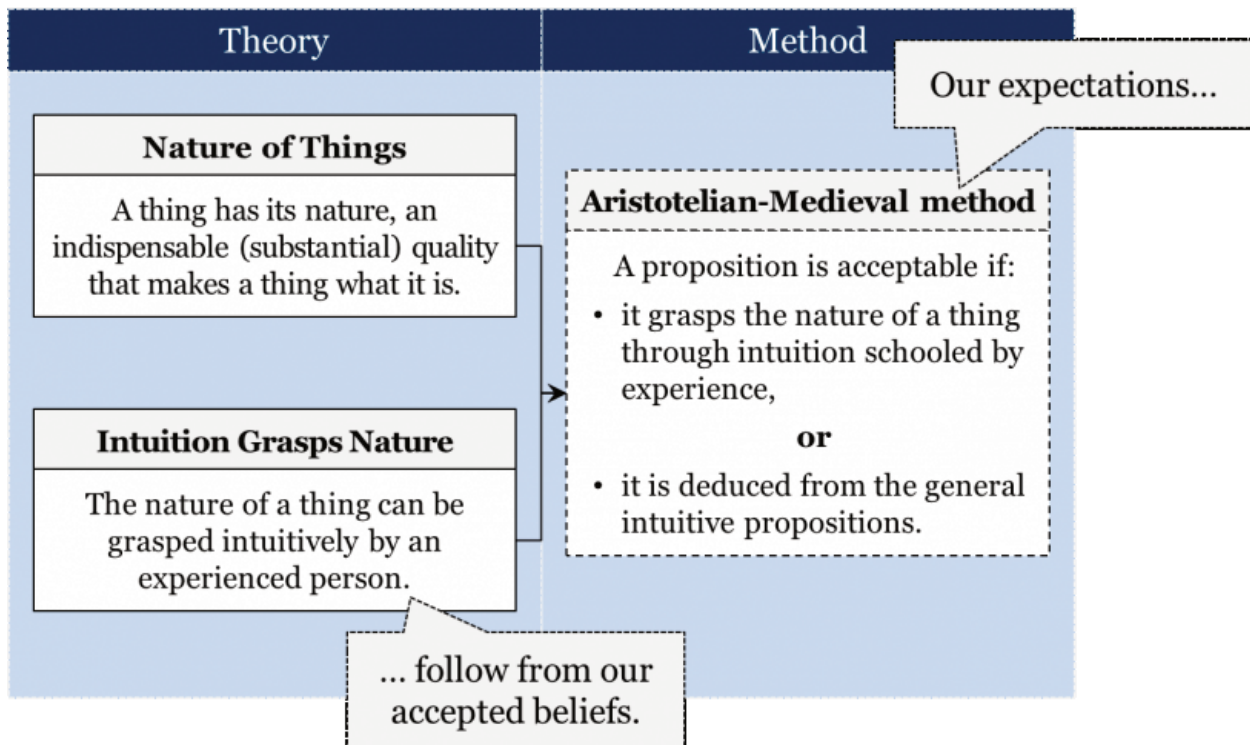


It is important to appreciate that the HD method is also *changeable* due to the fact that the assumptions on which it is based – those of complexity and post hoc explanations – are themselves changeable.

The same goes for the *Aristotelian-medieval method* – it was also shaped by the theories accepted at the time. One of the assumptions underlying the AM method is the idea that every natural thing has its *nature*, an indispensable (substantial) quality that makes a thing what it is. For instance, what makes an acorn an acorn is its capacity to grow into an oak tree. Similarly, the nature of a lion cub is its capacity to grow into a full-fledged lion. Finally, according to this way of thinking, the substantial quality of a human being is the capacity of reason.

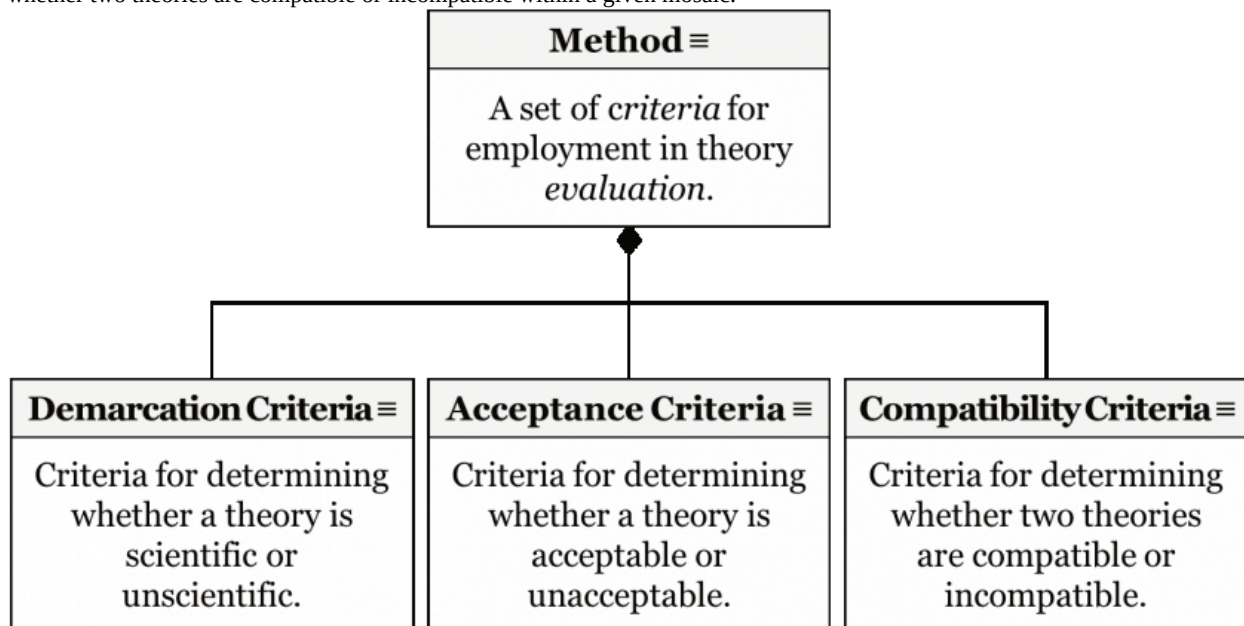
The second assumption underlying the AM method is the idea that the nature of a thing can be grasped intuitively by an experienced person. In a nutshell, the idea was that you get to better understand a thing as you gain more experience with it. Consider a beekeeper who has spent a lifetime among bees. It was assumed that this beekeeper is perfectly positioned to inform us on the nature of bees. The same goes for an experienced teacher who, by virtue of having spent years educating students, would know a thing or two about the nature of the human mind.

The requirements of the AM method deductively follow from these assumptions. The Aristotelian-medieval scholars wouldn't expect intuitive truths grasped by an experienced person if they didn't believe that the nature of a thing can be grasped intuitively by experienced people or if they questioned the very existence of a thing's nature.



Once again, we see how a community’s method of theory evaluation is shaped by the theories accepted by that community.

It is important to understand that a method can be composed of criteria of three different types – *acceptance criteria*, *demarcation criteria*, and *compatibility criteria*. So far, when talking about methods, we have focused on *acceptance criteria*. These are the criteria that determine whether a theory is acceptable or unacceptable. However, there are at least two other types of criteria. *Demarcation criteria* determine whether a theory is scientific or unscientific, while *compatibility criteria* determine whether two theories are compatible or incompatible within a given mosaic.



Since the third law of scientific change applies to all methods, it applies equally to all three types of criteria. While, so far, our historical examples concerned changes in acceptance criteria, the third law also explains how demarcation criteria and compatibility criteria change through time. Below we will discuss some examples of different compatibility criteria and how they are shaped by accepted theories. As for different demarcation criteria, we will discuss them in chapter 6.

To re-emphasize a point made earlier, we claim that this same mechanism of method employment is at work in any historical

period, in any scientific community. In science, methods and any criteria they contain are always employed the same way: our knowledge about the world shapes our expectations concerning new theories. But this is not the only transhistorical feature of scientific mosaics and their dynamics. The law of method employment is only one of *four* laws that together constitute a general theory of scientific change. Let's look at the remaining three laws in turn.

Theory Acceptance

As we've just seen, a mosaic's method is profoundly shaped by its accepted theories, and methods will often change when new theories are accepted. But how does a scientific community accept those new theories? As with method employment, there seems to be a similarly universal pattern throughout the history of science whenever new theories are accepted into a mosaic.

Whether you've noticed it or not we have given you many examples of the mechanism of theory acceptance in action. We saw how Fresnel's wave theory of light was accepted ca. 1820, how Einstein's theory of general relativity was accepted ca. 1920, and how other theories like Mayer's lunar theory and Coulomb's inverse square law were all accepted by their respective scientific communities. In chapter 3 we also considered the case of Galileo, which was a scenario in which the scientific community did *not* accept a proposed new theory. So, what *was not the case* in the Galileo scenario that *was the case* in every other scenario we have just mentioned? In the case of Galileo, his heliocentric theory did not satisfy the method of the time. Every other theory which *became* accepted *did* satisfy the method of their respective communities at the time of assessment.

Indeed, in every situation we've considered where a new contender theory is eventually accepted into a mosaic, the theory is only accepted if it somehow meets the requirements of *that* scientific community's method. To perhaps oversimplify this finding slightly: a theory is only accepted when a scientific community judges that it is acceptable! One of our earlier versions of the law of theory acceptance was actually very similar to this insight, saying that in order to become accepted into the mosaic, a theory is assessed by the method actually employed at the time. This did the trick for a while, but we came to realize that it only tells us that to be accepted a theory must be *assessed*, but that it doesn't tell us the conditions that actually lead to its *acceptance*. For that reason, we have reformulated the law in a way that, while slightly less elegant, clearly links the theory's assessment by the method to the possible assessment outcomes:

2nd Law: Theory Acceptance

If a theory satisfies the acceptance criteria of the method employed at the time, it becomes accepted into the mosaic; if it does not, it remains unaccepted; if assessment is inconclusive, the theory can be accepted or not accepted.

So, if a theory satisfies a mosaic's employed method, it is accepted. If a theory fails to satisfy the employed method – if it fails to meet the implicit expectations of that community – it will simply remain unaccepted and will not become part of the mosaic. If it is unclear whether the method has been satisfied or not (“assessment is inconclusive”), the law simply describes the two logically possible outcomes (accepted or not accepted). Most importantly for our purposes, however, are the two clear-cut scenarios: when a theory is assessed by the method of the time, satisfaction of the method always leads to the theory's acceptance, and failure to satisfy the method always leads to the theory remaining unaccepted.

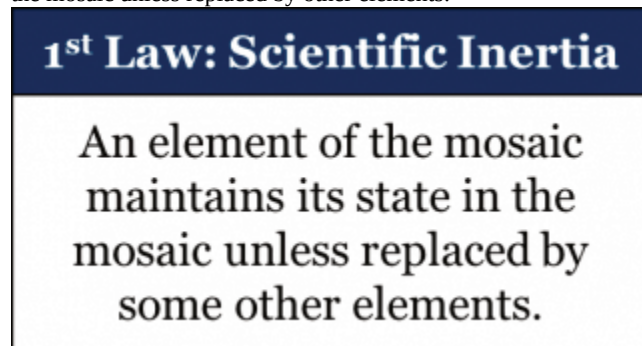
It is important to remember that there is not one universal scientific method that has been employed throughout history, so the requirements that a theory must satisfy in order to be accepted will depend on the method of that particular community at that particular time. Yes, this means that a theory being assessed by the method of one mosaic might be accepted, and a theory assessed by the method of a different mosaic might remain unaccepted. But methods are not employed willy-nilly, and their requirements are not chosen randomly! Despite the fact that *both theories and methods change*, these changes seem to follow a stable pattern, which we've captured in the two laws we've discussed thus far. The exciting diversity and variety in science is not the sign of an underlying anarchism or irrationality, but seems to be the result of stable, predictable laws of scientific change playing out in the rich history of human discovery and creativity.

The laws of method employment and theory acceptance are laws which focus on the way that the elements of mosaics *change*, and how new elements enter into the mosaic. But while change may be inevitable, we also know that science often displays periods of relative stability. So, is there a standard way in which theories and methods behave when there are no new elements being considered?

Scientific Inertia

When Isaac Newton presented his laws of motion in his 1687 opus *Philosophiæ Naturalis Principia Mathematica* (Mathematical Principles of Natural Philosophy), his obvious focus was on the dynamics of moving bodies and their influence on one another. But he realized that even objects which were not noticeably interacting with one another were also part of his physical system and should therefore be described as well. Newton's first law of motion about physical objects therefore answered the question: What happens when something is *not* being acted upon? His answer was that such an object basically *continues doing what it was doing*: if it was moving, it continues moving at the same rate in a straight line. If it was at rest, it remains at rest. He called this law his law of *inertia* – the Latin word *inertia* means “inactivity.”

What happens when the elements in our mosaic are not being “acted upon”? If there's no new theory being accepted, or no new method being employed, how do the existing elements of the mosaic behave? It seems that elements in a scientific mosaic *maintain their place in that mosaic*: that is, if a theory had been accepted, that theory continues to be accepted; if a method had been employed, it continues to be employed. Due to this coincidental similarity between Newton's first law of motion and the historical behaviour of the elements of scientific mosaics in times of relative stability, we have called this law of scientific change the *law of scientific inertia*. As precisely formulated, the law of scientific inertia states that an element of the mosaic remains in the mosaic unless replaced by other elements:



There are two major implications of this law that we need to unpack.

First, recall that if a theory is in a mosaic, it's an *accepted* theory, and if a method is part of a mosaic, it's an *employed* method. That is, if an element is part of the scientific mosaic, it has already been “vetted” by the community. What scientific inertia entails, therefore, is that these same theories are never *re-assessed* by the method, nor are these methods' requirements *re-deduced*. The integration of a particular element into the mosaic happens *once* (for theories, with the mechanism of theory acceptance; for methods, method employment) and it remains part of the mosaic until it's replaced by another element.

Second, *replacement* is essential to understanding this law properly. An element never just “falls out” of a mosaic; it has to be replaced by something. This is most clear in the case of theories, since even the negation of an accepted theory is itself a theory and can only be accepted by satisfying the method of the time. For instance, suppose that a scientific community's accepted theory changes from “the moon is made of cheese,” to “the moon is *not* made of cheese.” Is this an example of a theory simply disappearing from the mosaic? No. The negation of a theory is still a theory! Presumably the theory “the moon is *not* made of cheese” satisfied the method of the time and was accepted as the best available description of the material makeup of the moon, thereby replacing the earlier cheese theory.

What does scientific inertia look like in the real world? In biology today, the accepted theory of evolution is known as the *modern synthesis*; it has been accepted since the middle of the twentieth century. Two central sub-theories of the modern synthesis are “evolution results from changes in gene frequencies in populations of organisms,” and “populations adapt to their environments by natural selection.” Within the last 30 years or so, even strong proponents of the modern synthesis have begun to agree that these two sub-theories are inadequate, and that the modern synthesis struggles to account for well-known biological phenomena (such as epigenetics, niche construction, and phenotypic plasticity). Despite this building dissatisfaction, however, *the modern synthesis remains the accepted theory in contemporary biology*. Why? It is because no contender theory has yet satisfied the method of the biology community, as per the law of theory acceptance. Until a new theory can be proposed that satisfies the method of the biology community – and therefore *replace* the modern synthesis – the modern synthesis will remain the “best available” theory of evolution.

Compatibility

The fourth and final law of scientific change describes the scientific mosaic *at any moment of time* and is known as *the law of compatibility*. The law of compatibility states that at any moment of time, the elements of the scientific mosaic are compatible with each other:

10th Law: Compatibility

At any moment of time, the elements of the scientific mosaic are compatible with each other.

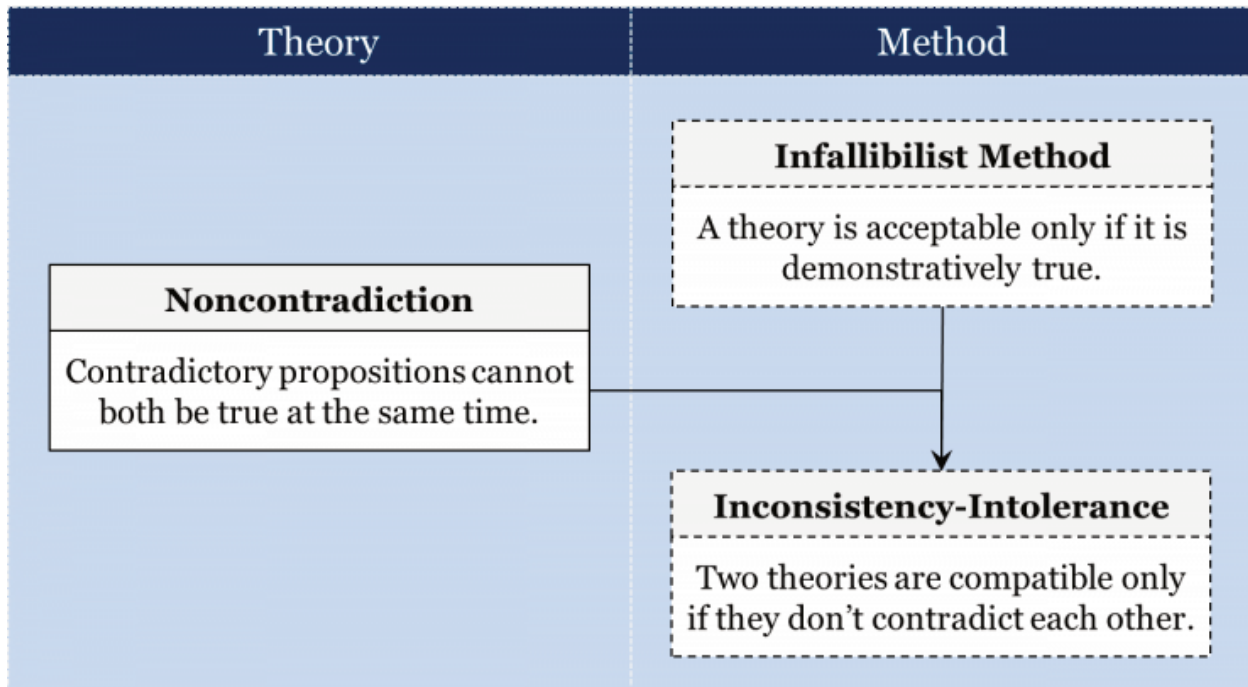
If two or more elements can exist in the same mosaic, we say that they are *compatible* with one another. If two or more elements cannot exist in the same mosaic, we say those elements are *incompatible*. This law simply points out that in any one scientific mosaic, *all elements* – theories and methods – are judged to be compatible.

The notion of compatibility must not be confused with the logical notion of consistency. Two theories are said to be logically *consistent* when there is no contradiction between them. *Compatibility* is the ability of theories to coexist within a community's mosaic regardless of their mutual consistency or inconsistency. Thus, in principle, it is possible for two logically inconsistent theories to be nevertheless compatible within a community's mosaic. In other words, compatibility is a *relative* notion: the same pair of theories may be compatible in one mosaic and incompatible in another.

What, ultimately, arbitrates which theories and methods are compatible with one another? In other words, where can we find the implicit *requirements* (i.e. *expectations, criteria*) which the scientific community would have for what is compatible and what is not? You guessed it: the method of the time. What is compatible in a mosaic, therefore, is determined by (and is part of) the method of the time, specifically its *compatibility criteria*. Just like any other criteria, compatibility criteria can also differ across different communities and time periods. Thus, a pair of theories may be considered compatible in one mosaic but incompatible in another, depending on the compatibility criteria of each mosaic.

But how do these compatibility criteria become employed? They become employed the same way any other method becomes employed: as per the law of method employment, each criterion is a deductive consequence of other accepted theories and employed methods. Let's look at some examples of different compatibility criteria and how they become employed.

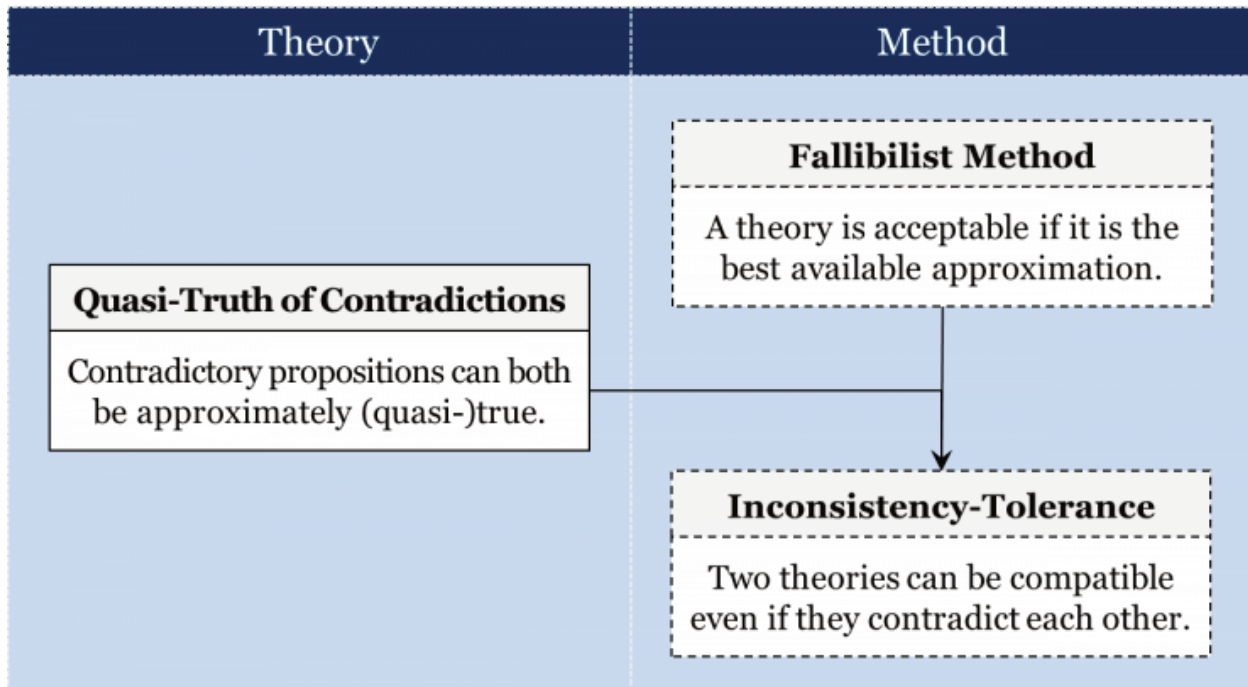
Recall that in the formal sciences we *can* make claims of absolute knowledge, because any formal proposition is simply a definition or follows from definitions. Communities in formal sciences, therefore, would employ an infallibilist method and consider a theory acceptable only if it is demonstratively true. We also know from the logical law of noncontradiction that contradictory propositions cannot both be true at the same time. Since accepted propositions in formal sciences are considered demonstratively true, it deductively follows that two theories which contradict each other cannot possibly be accepted at the same time in the mosaic of the same formal science community, since only one of them can be demonstratively true. This explains why formal science communities are normally intolerant towards inconsistencies. If we assume that the contemporary formal science community accepts that their theories are absolutely true, it follows deductively that in the mosaic of this formal science community, all logically inconsistent theories are also necessarily incompatible. Thus, this community would be *intolerant towards inconsistencies*, i.e. for this community inconsistent theories would be incompatible:



This is therefore a main facet of the compatibility criteria of this community's method. If this scientific community were to accept a new theory that was *inconsistent* with any earlier theories, those earlier theories would be considered *incompatible* and therefore rejected from the mosaic.

But are there communities that are *tolerant* towards inconsistencies? That is, does *inconsistency* always entail *incompatibility*? The short answer is "no". Let's consider the mosaic of an empirical science like contemporary physics. The contemporary physics community accepts both quantum mechanics and general relativity as part of their mosaic. Quantum mechanics deals principally with incredibly small objects, like the elements of the standard model (quarks, leptons, and bosons), while general relativity deals with incredibly large-scale objects and high energy interactions, like the bending of space by massive objects and velocities approximating the speed of light. But, interestingly, these two theories are inconsistent with one another when it comes to their interpretation of an object that belongs to both of their domains, such as the singularity within a black hole. It belongs to both domains because singularities are incredibly small and thus should be covered by quantum physics, yet they also have an incredibly large mass and thus should be covered by general relativity. When the two theories are simultaneously applied to singularities, the inconsistency between the two becomes apparent. And yet these two inconsistent theories are simultaneously accepted nowadays and are therefore considered to be *compatible* despite this inconsistency. Why might that be the case?

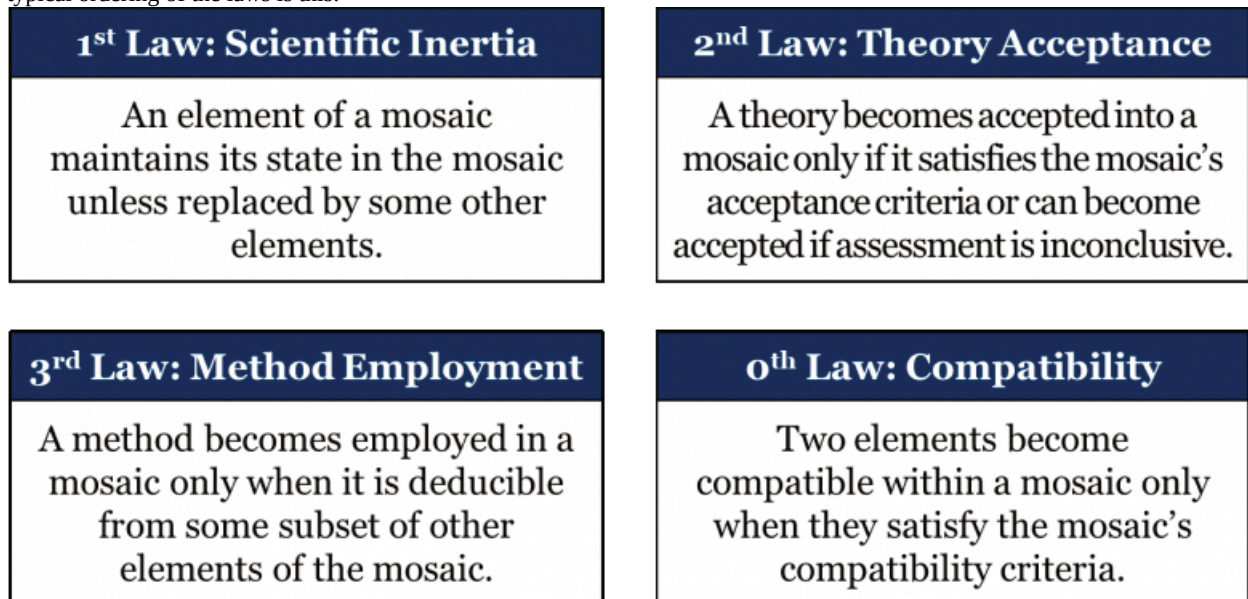
Remember that the contemporary physics community (like all empirical science communities) is fallibilist, so it is prepared to accept theories that are far from being demonstratively true. This community is content to accept the best available approximations. (Look back at chapter 2 if you need a refresher as to why!) In addition, this empirical science community also accepts that contradictory propositions can both be *approximately* (or *quasi-*) true. If theories cannot be absolutely true, then it is at least *possible* that two obviously *inconsistent* theories might still nevertheless be considered – in certain circumstances – compatible. In this case, physicists realize that they have two incredibly successful theories that have independently satisfied the employed HD method, but that these two theories focus on different scales of the physical universe: the tiny, and the massive, respectively. Their fallibilism and the two distinct domains carved out by each branch of physics lead us to conclude that physicists are *tolerant towards inconsistencies*:



Naturally, an empirical science community need not necessarily be tolerant towards *all* inconsistencies. For instance, it may limit its inconsistency toleration only to accepting inconsistent theories that belong to two sufficiently different domains. For example, an empirical science community can be tolerant towards inconsistencies between two accepted theories when each of these theory deals with a very distinct scale. Alternatively, it may only be tolerant towards inconsistencies between an accepted theory and some of the available data obtained in experiments and observations. In any event, it shouldn't be tolerant towards *all* possible inconsistencies to qualify as inconsistency-tolerant.

The Four Laws of Scientific Change

We have now introduced you to all four laws of scientific change. But we didn't introduce you to them in numerical order! The typical ordering of the laws is this:



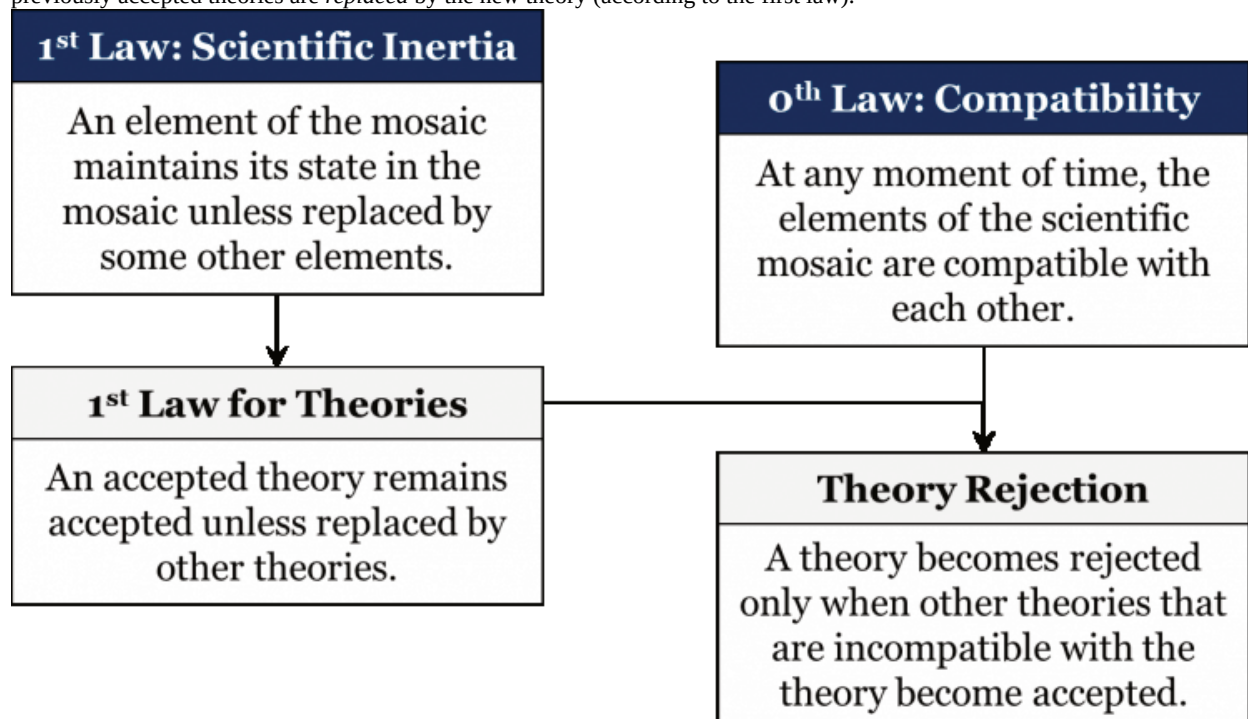
Scientific inertia was given the honour of being the *first law*, as something of an homage to Newton's laws of motion (as explained above). The law of compatibility has been called the *zeroth law* precisely because it is a "static" law, describing the state of the mosaic during and between transitions. For the purposes of this textbook, it is important to learn the numbers of each

respective law, since the numbers will act as a helpful shorthand moving forward. For instance, we can simply reference “the third law” instead of writing about “the law of method employment”.

These four laws of scientific change together constitute a general theory of scientific change that we will use to help explain and understand episodes from the history of science, as well as certain philosophical topics such as scientific progress (chapter 5) and the so-called demarcation problem (chapter 6). The laws serve as axioms in an axiomatic-deductive system, and considering certain laws together helps us shed even more light on the process of scientific change than if we only consider the laws on their own. In an axiomatic-deductive system, propositions deduced from axioms are called *theorems*. Currently, over twenty theorems have been deduced from the laws of scientific change, but we will only briefly consider three.

Some Theorems

We’ve talked about how theories are accepted, but can we be more explicit about how they are rejected? If we look at the first and zeroth laws together we can derive the *theory rejection theorem*. By the first law, an accepted theory will remain accepted until it is replaced by other theories. By the zeroth law, the elements of the scientific mosaic are always compatible with one another. Thus, if a newly accepted theory is incompatible with any previously accepted theories, those older theories leave the mosaic and become *rejected*, and are replaced by the new theory. This way compatibility is maintained (according to the zeroth law) and the previously accepted theories are *replaced* by the new theory (according to the first law).

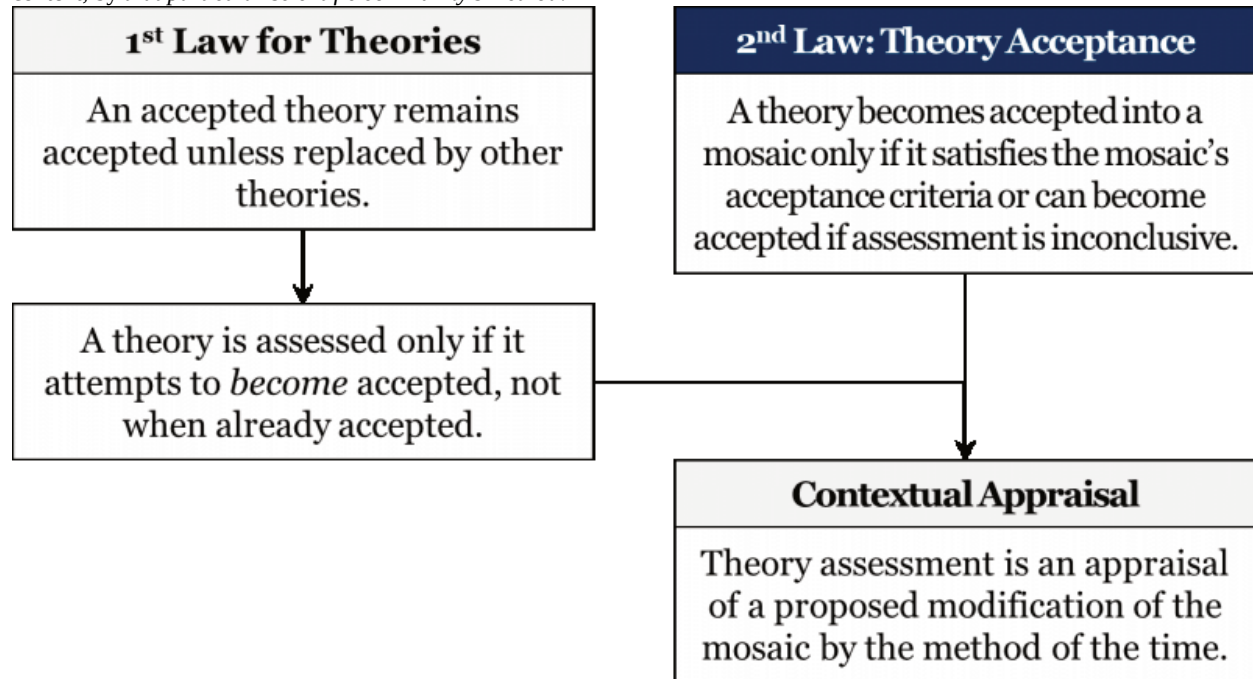


How do new theories become appraised? Specifically, is it at all possible to evaluate a theory in isolation from other theories? For a long time, it was believed by philosophers that the task of theory evaluation is to take the theory and determine whether it is absolutely true or false. If a theory passes the evaluation, then it is considered absolutely true; if not, it is considered false. This is the gist of the conception of *absolute theory appraisal*, according to which it is possible to assess a standalone theory without comparing it to any other theory. However, this conception of absolute appraisal would only make sense if we were to believe that there is such a thing as absolute certainty. What happens to the notion of theory appraisal once we appreciate that all empirical theories are fallible? When philosophers realized that no empirical theory can be shown to be absolutely true, they came to appreciate that the best we can do is to *compare* competing theories with the goal of establishing which among competing theories is *better*. Thus, they arrived at what can be called the conception of *comparative theory appraisal*. On this view of theory appraisal, any instance of theory evaluation is an attempt to compare a theory with its available competitors, and therefore, theory appraisal is never absolute. In other words, according to comparative theory appraisal, there is no such thing as an appraisal of an individual theory independent of other theories.

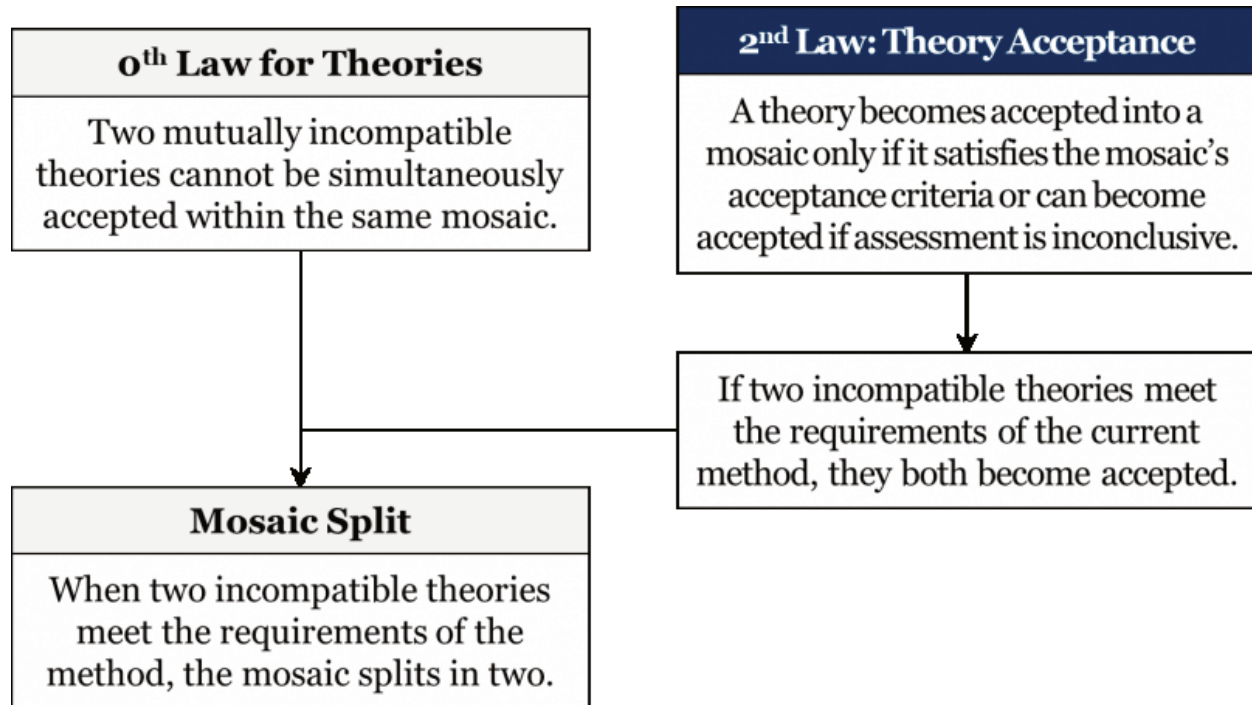
In recent decades, it has gradually transpired that while all theory appraisal is comparative, it is also necessarily *contextual*, as it depends crucially on the content of the mosaic in which the appraisal takes place. Specifically, it is accepted by many philosophers and historians these days that the outcome of a theory’s appraisal will depend on the specific method employed by the community making the appraisal. Thus, it is conceivable that when comparing two competing theories, one community may prefer one theory

while another may prefer the other, depending on their respective employed methods. We can call this view *contextual theory appraisal*. Importantly, it is this view of theory appraisal that follows from the laws of scientific change.

Considering the first and second laws, we can derive the *contextual appraisal theorem*. By the first law, a theory already in the mosaic stays in the mosaic insofar as it is not replaced by other theories. This means that once a theory is accepted into the mosaic, it is no longer expected to satisfy any requirements to stay in the mosaic. By the second law, a theory is assessed by the method employed at the time of the assessment. In other words, the assessment takes place when members of a community attempt to bring a theory into the mosaic. As such, theory assessment is an assessment of a proposed *modification* of the mosaic by the method employed at the time. In other words, theory assessment is contextual as it takes place within a particular historical/geographical context, by that particular scientific community's method.



What happens when two incompatible theories satisfy the community's employed method at the same time? By the second law, we know that if a theory satisfies the method of the time, it will be accepted. From the zeroth law, however, we know that all elements in a mosaic are compatible with one another. So, if two incompatible theories simultaneously satisfy the same employed method, to maintain compatibility *the mosaic itself splits*: you would have one entirely compatible mosaic containing one theory as accepted, and a second entirely compatible mosaic containing the other. This is known as the *mosaic split theorem*.



There are many examples of mosaic split in the history of science, the most famous of which occurred between the Newtonian and Cartesian mosaics at the turn of the 18th century. Up until this point the scientific community throughout Europe employed the Aristotelian-Medieval method. You'll recall that the acceptance criteria for the AM method required that the theory be intuitive to an experienced person. Isaac Newton's vast axiomatic-deductive system of physical theorems was derived from his three simple laws of motion and his law of universal gravitation. René Descartes' equally impressive axiomatic-deductive system (incompatible with that of Newton) was similarly deduced from a set of intuitively obvious axioms, such as "Matter is extension." Faced with these two impressively comprehensive sets of theories, the scientific community found that *both* satisfied their (vague) AM acceptance criterion of being "intuitive": both systems began from easily-graspable axioms and validly deduced a plethora of theorems concerning different aspects of the physical world from them. The result was that *both* theories became accepted, but that the mosaic itself split in two: one Newtonian, one Cartesian.

But what historical factors decide who becomes Newtonian and who becomes Cartesian? The mosaic split theorem tells us that, given the circumstances, a split is going to take place. Yet, it doesn't say which part of a previously unified community is going to accept Newtonianism and which part is going to accept Cartesianism. From the perspective of the laws of scientific change, there is nothing determining which of the two new communities will accept what. It's reasonable to suspect that a variety of sociocultural factors play a role in this process. Some sociological explanations point to the fact that France, home to Descartes, and Britain, home to Newton, were rival nations celebrating and endorsing the discoveries and advancements of their own people first. Other explanations pointing to the theological differences between the mosaics of France and Britain from before 1700, with Catholicism accepted in the former and Anglicanism in the latter. Keep in mind that in early modern Europe, theological propositions and natural philosophical propositions were typically part of the same mosaic and had to be compatible with one another. In such a case, the acceptance of Cartesianism and Newtonianism in France and the Continent respectively represented the widening rather than beginning of a mosaic split. We will have an opportunity to look at these two mosaics (and explore exactly why they were incompatible) in the history chapters of this textbook.

Summary

Can there be a general theory of scientific change? To begin answering the question we first made a distinction between general and particular questions, and between general and particular theories. We went on to clarify that *generalists* answer "yes" to this question, and that *particularists* answer "no." Afterwards we introduced four arguments used to support particularism. However, we went on to argue for generalism, and presented the four *laws of scientific change* as a promising candidate for a general theory of scientific change. By way of review, let's see how we might respond to the four main arguments for particularism via the laws of scientific change.





Bad Track Record: While poor past performance of general theories of scientific change is indeed an argument *against* generalism, it doesn't rule out the possibility of the creation of a new theory that has learned from past mistakes. The laws of

scientific change take seriously history, context, and change, and seem to succeed where other theories have failed. It's a contender, at least!

Nothing Permanent (Including the Scientific Method): Combining the second and third arguments, particularists note that stability and patterns are the cornerstone for answering general questions with general theories, and “science” doesn’t seem to have them. With no fixed method and with highly unique historical contexts, the prospect for grounding a general theory seems nearly impossible. The laws of scientific change, however, are not grounded in some transhistorical element of a mosaic. Rather, they recognize that *the way science changes*, the *patterns* of scientific change, are what remain permanent throughout history. The four laws are descriptions of the stable process by which theories and methods change through time. Besides this, the laws, in their current form, don’t ignore the uniqueness and peculiarities of certain historical periods. The laws rather incorporate them into science’s transhistorical dynamic (see the *contextual appraisal & mosaic split theorems*).

Social Construction: In the absence of a general theory of scientific change, some historians have attempted to explain individual scientific changes by means of the theories of sociology and anthropology. But hopefully our exposition of the laws has, at least, helped to demonstrate that the process of scientific change has its own unique *patterns* that are worth studying and describing on their own terms. Indeed, the laws show that not only can general questions be asked about science, but general answers seem to be discoverable as well. Using general theories from a related discipline can be helpful, but why stick with more vague theories that are harder to apply when more specific theories are at your fingertips?

While particularism remains the currently *accepted* position in this debate, it is clearly unjustified, as all of its main arguments are flawed:

 Can there be a <i>general</i> theory of scientific change?	
Yes	No
<p>Generalism:</p> <p>There can be a general theory of scientific change.</p> <p> All cases of scientific change seem to obey certain general laws</p>	<p>Particularism:</p> <p>There can be no general theory of scientific change.</p> <p> There is no universal and fixed method of science</p> <p> There is nothing universal in science: each historical episode is unique</p>

We will *pursue* and *use* the laws of scientific change throughout the remainder of this textbook. As fallibilists, we realize that the laws of scientific change are imperfect and will likely be improved. But we are confident that the laws of scientific change as they stand now are at least a valuable tool for parsing the sometimes overwhelmingly complex history of science.

5.

Scientific Progress -- Introduction to History and Philosophy of Science

Chapter 5: Scientific Progress

Intro

In previous chapters we've established that the theories within mosaics change through time. We have also made the case that these changes occur in a regular, law-like fashion. Recognizing that a mosaic's theories change through time, and understanding how they change, is important. But what makes us think that our empirical theories succeed in describing the world? Yes, we can make a claim that they change in a law-governed fashion, i.e. that theories only become accepted when they meet the expectations of the respective community. But does that mean that our accepted physical, chemical, biological, psychological, sociological, or economic theories actually manage to tell us anything about the processes, entities, and relations they attempt to describe? In other words, we know that the process of scientific change exhibits certain general patterns of change, but does that necessarily mean that this law-governed process of changes in theories and methods actually takes us closer to the true description of the world as it really is?

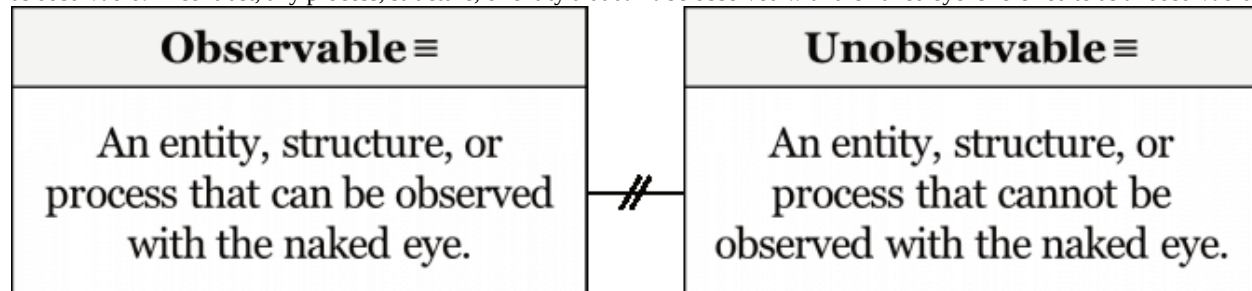
The position of fallibilism that we have established in chapter 2 suggests that all empirical theories are – *at best* – approximations. Yet, it doesn't necessarily follow from this that our empirical theories actually succeed in approximating the world they attempt to describe. To begin this chapter, we will tackle the following question:

Do our best theories correctly describe the mind-independent external world?

Before we go any further, it's important to keep in mind that we are committed to the philosophical viewpoint of *fallibilism*. As such, whenever we talk about an empirical theory's *correctness*, *success*, or *truth*, those theories are always subject to the problems of sensations, induction, and theory-ladenness, and are therefore not absolutely certain, not absolutely true. So, the question is not whether our empirical theories provide *absolutely* correct descriptions of the world; it is nowadays accepted that they do *not*. The question is whether we can claim that our best theories get at least *something* correct, i.e. whether they succeed as approximate descriptions of the world as it really is.

Scientific Realism vs. Instrumentalism

Many theories of empirical science attempt to explain easily observable facts, events, processes, relations, etc. Let's take the free-fall of a coffee cup as an example. We can easily observe a coffee cup fall from a table to the ground, i.e. the fall itself is an *observable* process. For instance, we can measure the time it took the coffee cup to hit the ground and formulate a proposition that describes the results of our observation: "it takes 0.5 seconds for the cup to hit the ground". But what about explaining *why* the coffee cup falls from the table to the ground? To describe the underlying mechanism which produces the fall, we nowadays cite the theory of general relativity. The explanation provided by general relativity is along these lines: the coffee cup falls to the ground because the Earth's mass bends the space-time around it in such a way that the inertial motion of the cup is directed towards the ground. While a cup falling through the air is easily observable, the bent space-time which general relativity invokes is not. Similarly, many scientific theories postulate the existence of entities, structures, or processes that are too small (like quarks, atoms, or microbes), too large (like clusters of galaxies), too fast (like the motion of photons), or too slow (like the process of biological evolution by natural selection) to be directly detected with the human senses. In addition, some of the entities, structures, or processes invoked by our scientific theories are such that cannot be directly observed even in principle (like space-time itself). Any process, structure, or entity that can be perceived without any technological help ("with the naked eye") is typically referred to as *observable*. In contrast, any process, structure, or entity that *can't* be observed with the naked eye is referred to as *unobservable*.



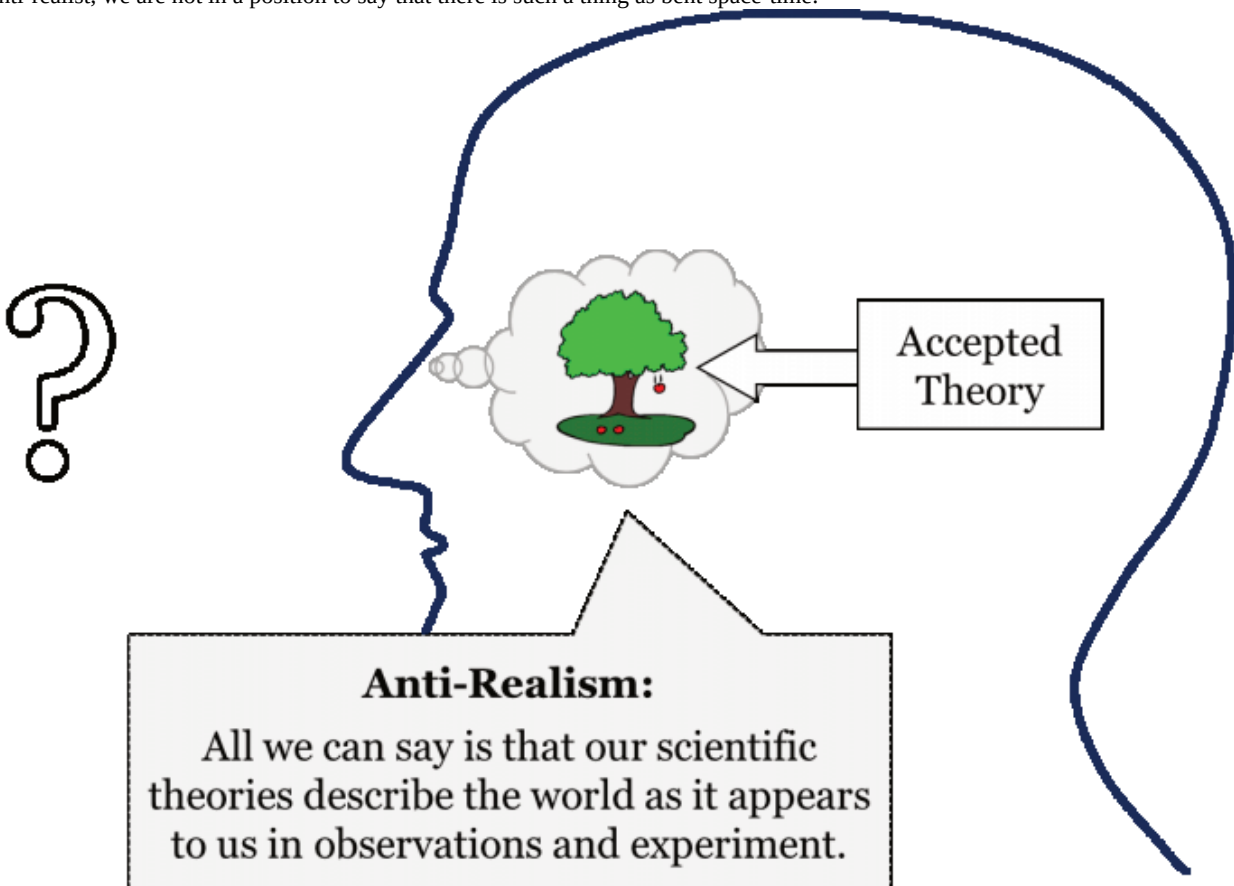
Since many – if not most – empirical scientific theories invoke *unobservables* to help explain their objects, this makes our initial question a little bit more interesting. When we ask whether our theories correctly describe the external world, we are actually asking the more specific question:

Do our scientific theories correctly describe both *observables* and *unobservables*?

Generally speaking, scientists and philosophers of science today accept that our theories succeed in correctly describing observables, even as fallibilists. When we've dropped that coffee cup for the 400th time and still clock its airtime at 0.5 seconds, we'll consider the claim "it takes 0.5 seconds for the cup to hit the ground" confirmed and correct. But how should we understand claims concerning unobservables, like invoking the notion of bent space-time to explain the fall of the coffee cup? Can we legitimately make any claims about unobservable processes, entities, or relations? Thus, the question of interest here is that concerning *unobservables*:

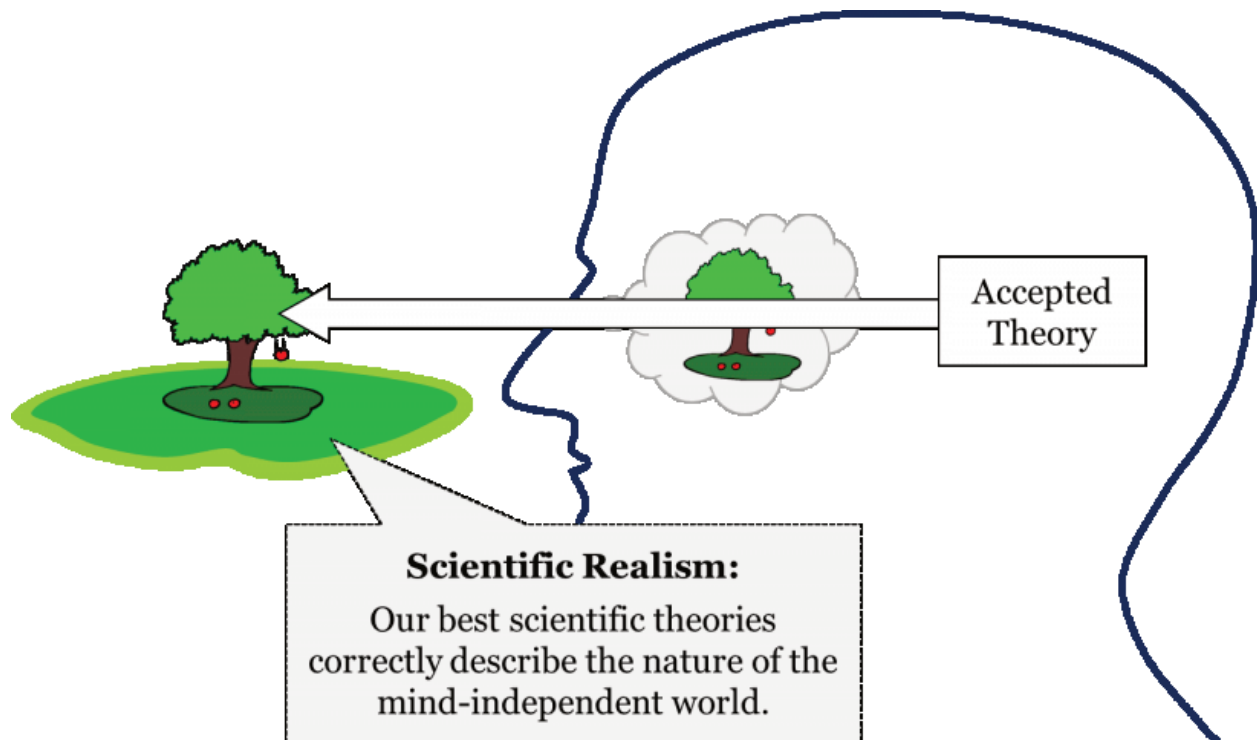
Do our scientific theories correctly describe *unobservables*?

There are many philosophers who are relatively *pessimistic* when it comes to our ability to make legitimate claims about the unobservable entities, structures, or processes posited by our scientific theories. Those who hold this position would answer "no" to the question of whether scientific theories correctly describe unobservables. This is the position of *anti-realism*. While anti-realists don't deny that our theories often correctly describe *observables*, they do deny that we can make any legitimate claims about the reality of unobservable entities, processes, or relations invoked by our scientific theories. For instance, according to an anti-realist, we are not in a position to say that there is such a thing as bent space-time:



The position of anti-realism is also often called *instrumentalism* because it treats scientific theories generally – and the parts of those theories that make claims about unobservables specifically – merely as tools or *instruments* for calculating, predicting, or intervening in the world of observables. Thus, according to instrumentalists, the notion of bent space-time is merely a useful mathematical *tool* that allows us to calculate and predict how the locations of observable objects change through time; we can't make any legitimate claims concerning the reality of that bent space-time. In other words, while instrumentalism holds that theories invoking unobservables often yield practical results, those same theories might not actually be succeeding in describing genuine features of the world.

But there are also those who are *optimistic* when it comes to our ability to know or describe unobservables posited by our scientific theories. This position is called *scientificrealism*. Those who hold this position would answer "yes" to the question of whether scientific theories correctly describe unobservables. For them, unobservables like quarks, bosons, natural selection, and space-time are not merely useful instruments for making predictions of observable phenomena, but denote entities, processes, and relations that likely exist in the world.



It is important to note that the question separating scientific realists from instrumentalists doesn't concern the *existence* of the external mind-independent world. The question of whether our world of observable phenomena is all that there is or whether there is an external world beyond what is immediately observable is an important question. It is within the domain of metaphysics, a branch of philosophy concerned with the most general features of the world. However, that metaphysical question doesn't concern us here. Our question is not about the *existence* of the external mind-independent world, but about the ability or inability of our best scientific theories to tell us anything about the features of that mind-independent external world. Thus, an instrumentalist doesn't necessarily deny the existence of a reality beyond the world of observable phenomena. Her claim is different: that even our best scientific theories fail to reveal anything about the world as it really is.

To help differentiate scientific realism from instrumentalism more clearly, let's use the distinction between *acceptance* and *use* we introduced in chapter 3. Recall that it is one thing to accept a theory as the best available description of its respective domain, and it's another thing to use it in practical applications. While a community can *use* any number of competing theories in practice (like different tools in a toolbox), it normally *accepts* only one of the competing theories as the best available description of its object.

Now, instrumentalists and scientific realists don't deny that communities do *in fact* often accept theories and use them in practice; the existence of these epistemic stances in the actual practice of science is beyond question. What separates instrumentalists and scientific realists is the question of the *legitimacy* of those stances. The question is not whether scientists have historically accepted or used their theories – it is clear that they have – but whether it is *justifiable* to do so. Since we are dealing with two different stances (*acceptance* and *use*) concerning two different types of claims (about *observables* and *unobservables*), there are four distinct questions at play here:

- Can we legitimately *use* a theory about *observables*?
- Can we legitimately *accept* a theory as describing *observables*?
- Can we legitimately *use* a theory about *unobservables*?
- Can we legitimately *accept* a theory as describing *unobservables*?

As far as the practical *use* of theories is concerned, there is no disagreement between instrumentalists and realists: both parties agree that any type of theory can legitimately become useful. This goes both for theories about observable phenomena and theories about unobservables. Instrumentalists and realists also agree that we can legitimately *accept* theories about observables. Where the two parties differ, however, is in their attitude concerning the legitimacy of *accepting* theories about *unobservables*. The question that separates the two parties is whether we can legitimately *accept* any theories about *unobservables*. In other words, of the above four questions, realists and instrumentalists only differ in their answer to one of them. This can be summarized in the following table:

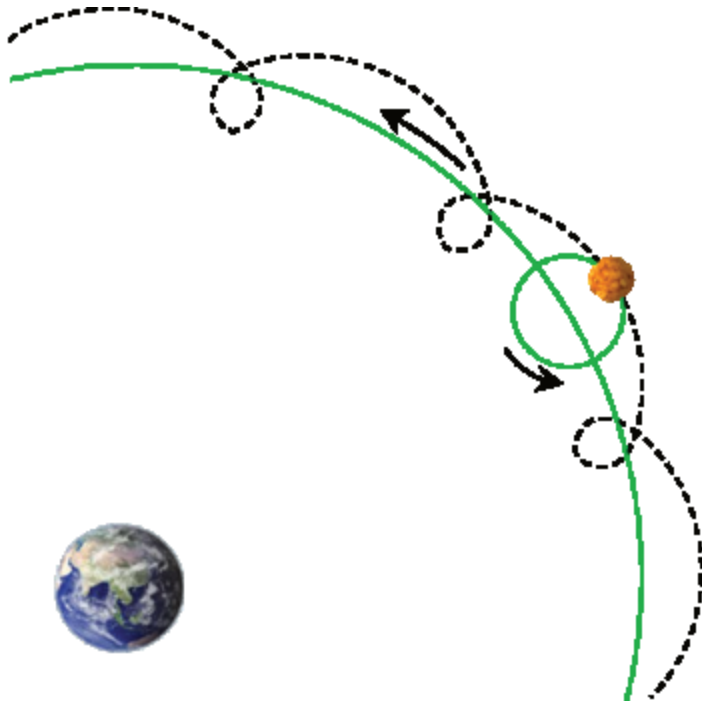
	? Can our descriptions of <i>observables</i> be legitimately...		? Can our descriptions of <i>unobservables</i> be legitimately...	
	... used?	... accepted?	... used?	... accepted?
Instrumentalism:	Yes	Yes	Yes	No
Scientific Realism:	Yes	Yes	Yes	Yes

For an instrumentalist, theories concerning unobservables can sometimes be legitimately used in practice (e.g. bridge building, telescope construction, policy making) but *never* legitimately accepted. For a realist, theories concerning unobservables can sometimes be legitimately used, sometimes legitimately accepted, and sometimes both legitimately used and accepted. The following table summarizes the difference between the two conceptions:

? Do our best scientific theories correctly describe the nature of the external (mind-independent) world?	
Yes	No
<p>Scientific Realism:</p> <p>Our best scientific theories correctly describe the nature of the mind-independent world.</p> <p>Thus, we can both <i>use</i> theories in practical applications and <i>accept</i> them as best available descriptions of the external world.</p>	<p>Instrumentalism:</p> <p>We do not know whether our theories correctly describe the nature of the mind-independent world.</p> <p>Thus, we shouldn't <i>accept</i> theories, but should only <i>use</i> them in practical applications, for theories are nothing more than useful <i>tools</i>.</p>

Now that we've established the basic philosophical difference between instrumentalism and realism, let's consider some historical examples to help illustrate these contemporary distinctions.

For centuries, astronomers accepted Ptolemy's model of a geocentric universe, including the Ptolemaic theory of planetary motion. First, recall that for Ptolemy the observable paths of the planets, including their retrograde motion, were produced by a combination of epicycles and deferents. Here is the diagram we considered in chapter 3:

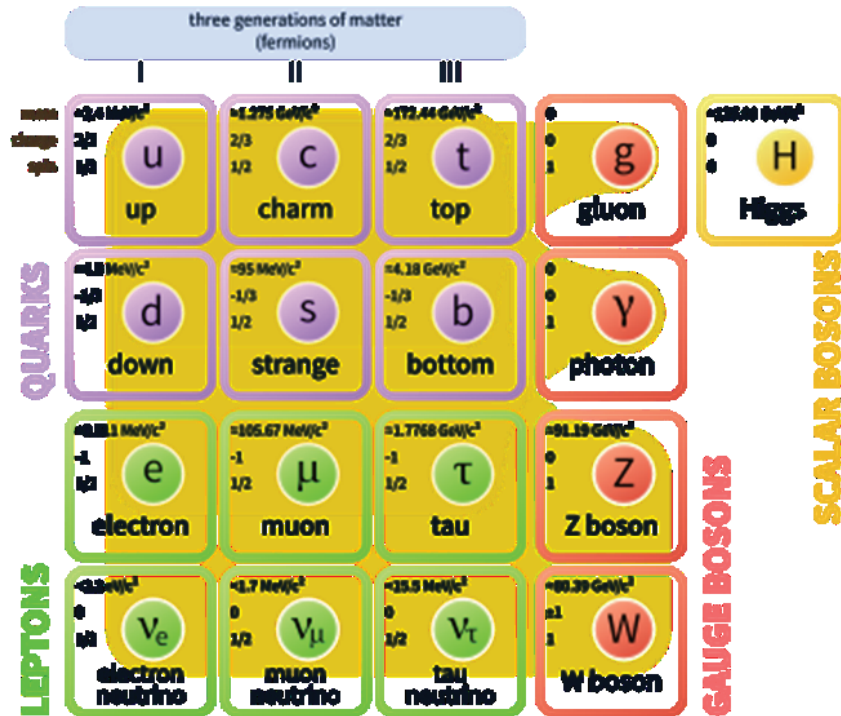


In addition to being accepted, this epicycle-on-deferent theory and its corresponding mathematics was used for centuries to predict the position of the planets with remarkable accuracy. The tables of planetary positions composed by means of that theory (the so-called *ephemerides*) would then be used in astrology, medicine, navigation, agriculture, etc. (See chapter 7 for more detail.)

Now let's break this example down a little bit. On the one hand, the planets and their retrograde motion are *observables*: we don't need telescopes to observe most of the planets as points of light, and over time we can easily track their paths through the night sky with the naked eye. On the other hand, from the perspective of an observer on the Earth, deferents and epicycles are *unobservables*, as they cannot be observed with the naked eye. Instead, they are the purported mechanism of planetary motion – the true shape of the orbs which underlie the wandering of the planets across the night sky.

So how would we understand Ptolemy's epicycle-on-deferent theory of planetary motion from the perspectives of instrumentalism and realism? For both instrumentalists and realists, the Ptolemaic account of the meandering paths of the planets through the night sky would be *acceptable* since these paths are observable. That is, both conceptions agree that, at the time, the Ptolemaic predictions and projections for planets' locations in the night sky could be legitimately considered the *best* description of those phenomena. But instrumentalism and realism disagree over whether it was legitimate for medieval astronomers to also accept those parts of Ptolemy's theory which referred to *unobservable* deferents and epicycles. According to realism, at the time, Ptolemy's theory about epicycles and deferents could be legitimately accepted as the best available description of the actual mechanism of planetary motion. In contrast, the instrumentalist would insist that astronomers had to refrain from committing themselves to the reality of epicycles and deferents, and instead had to focus on whether the notions of epicycle and deferent were *useful* in calculating planetary positions.

Consider a slightly more recent historical example: the standard model of quantum physics. According to the standard model, there are six *quarks*, six *leptons*, four *gauge bosons* and one *scalar boson*, the recently discovered Higgs boson. Here is a standard depiction of the standard model:



For our purposes, we can skip the specific roles each type of elementary particle plays in this model. For the purposes of our discussion it is important to note that all of these elementary particles are unobservables; due to their minute size, they cannot be observed with the naked eye. In any event, it is a historical fact that this standard model is currently *accepted* by physicists as the best available description of its domain.

The philosophical question separating instrumentalists and scientific realists is whether it is legitimate to believe that these particles are more than just a useful calculating tool. Both scientific realists and instrumentalists hold that we can legitimately *use* the standard model to predict what will be *observed* in a certain experimental setting given such-and-such initial conditions. However, according to instrumentalists, scientists should *not accept* the reality of these particles, but should consider them *only* as useful instruments that allow them to calculate and predict the results of observations and experiments. In contrast, scientific realists would claim that we can legitimately accept the standard model as the best available description of the world of elementary particles, i.e. that the standard model is *not* a mere predicting tool.


Species of Scientific Realism

A historical note is in order here. In the good-old days of infallibilism, most philosophers believed that, in one way or another, we manage to obtain absolute knowledge about the mind-independent external world. Naturally, the list of theories that were considered strictly true changed through time. For instance, in the second half of 18th century, philosophers would cite Newtonian physics as an exemplar of infallible knowledge, while the theories of the Aristotelian-medieval mosaic would be considered strictly true in the 15th century. But regardless of which scientific theories were considered absolutely certain, it was generally accepted that such absolute knowledge does exist. In other words, most philosophers accepted that our best scientific theories provide us with a picture of the world just as it really is. This species of scientific realism is known as *naïve realism*.

However, as we have seen in chapter 2, philosophers have gradually come to appreciate that all empirical theories are, in principle, fallible. Once the transition from infallibilism to fallibilism was completed, the position of naïve realism could no longer be considered viable, i.e. it was no longer possible to argue that our best theories provide us with the exact image of the world as it really is. Instead, the question became:

Do our scientific theories *approximate* the mind-independent external world (i.e. the world of unobservables)?

Thus, the question that fallibilists ask is not whether our theories succeed in providing an exact picture of the external world – which, as we know, is impossible – but whether our theories at least succeed in approximating the external world. It is this question that nowadays separates scientific realists from instrumentalists. Thus, the version of scientific realism that is available to fallibilists is not naïve realism, but *critical realism*, which holds that empirical theories can succeed in approximating the world. While critical realists believe that at least some of our best theories manage to provide us with some, albeit fallible, knowledge of the world of unobservables, nowadays there is little agreement among critical realists as to how this approximation is to be understood.

 Can our empirical theories provide us with a picture of the world as it really is?	
Yes	No
Naïve Realism: Our theories provide us with a picture of the world just as it really is.	Critical Realism: All empirical theories are at best approximations.

Once the fallibility of our empirical theories became accepted, many critical realists adopted the so-called *selective* approach. *Selective scientific realists* attempt to identify those *aspects* of our fallible theories which can be legitimately accepted as approximately true. On this selective approach, while any empirical theory taken as a whole is strictly speaking false, it may nevertheless have *some parts* which scientists can legitimately accept as approximately true. Thus, the task of a selective scientific realist is to identify those aspects of our theories which warrant such an epistemic stance. While selective scientific realism has many sub-species, here we will focus on two of the most common varieties – *entity realism* and *structural realism*. These two varieties of selective realism differ in their answers to the question which aspects of our theories scientists can legitimately consider approximating reality.

According to *entity realism*, scientists can be justified in accepting the reality of unobservable *entities* such as subatomic particles or genes, provided that they are able to manipulate these unobservable entities in such a way as to accurately bring about observable phenomena. For instance, since scientists can accurately predict what exactly will be observed when the putative features of an electron are being manipulated, then they have good reason to accept the reality of that electron. Entity realists hold that it is the scientists' ability to causally manipulate unobservable entities and produce very precise observable outcomes that justifies their belief that these unobservable entities are real. The key reason why some philosophers find the position of entity realism appealing is that it allows one to continue legitimately accepting the reality of an entity despite of any changes in our knowledge concerning the behaviour of that entity. For example, according to entity realists, we can continue accepting the reality of an electron regardless of any additional knowledge concerning specific features of the electron and its behaviour that we may acquire in the future. Importantly, entity realism takes a *selective* approach as to which parts of our theories can be legitimately accepted and which parts can only be considered useful.

A different version of the selective approach is offered by *structural realism*. While structural realists agree with entity realists that scientists can legitimately accept only certain *parts* of the best scientific theories, they differ drastically in their take on which parts scientists are justified in accepting. According to structural realists, scientists can justifiably accept not the descriptions of unobservable *entities* provided by our best theories, but only the claims these theories make about unobservable *structures*, i.e. about certain *relations* that exist in the world. One key motivation for this view is a historical observation that often our knowledge of certain structural relations is being preserved despite fundamental changes in our views concerning the types of entities that populate the world. Consider, for instance, the law of refraction from optics:

The ratio of the sines of the angles of incidence θ_1 and angle of refraction θ_2 is equivalent to the ratio of phase velocities (v_1 / v_2) in the two media:

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{v_1}{v_2}$$

Setting aside the question of who is to be rightfully credited with the discovery of this law – René Descartes, or the Dutch astronomer Willebrord Snellius, or even the Persian scientist Ibn Sahl – we can safely say that all theories of light accepted since the 17th century contained a version of this law. This goes for Descartes' mechanistic optics, Newton's optics, Fresnel's wave theory of light, Maxwell's electrodynamics, as well as contemporary quantum optics. These different optical theories had drastically opposing views on the nature of light: some of these theories understood light as a string of corpuscles (particles), while other theories treated light as a wave-like entity, while yet others considered light as both corpuscular and wave-like. But despite these drastic changes in our understanding of the nature of light, the law of refraction has maintained its state in our mosaic. Thus, according to structural realists, scientists are justified in accepting those claims of our theories which reveal certain underlying structures (relations), while the claims about unobservable entities are to be taken sceptically. In this view, scientists can still find the claims regarding unobservable entities *useful*, but they are not justified in *accepting* those claims. All they can legitimately do is to believe that their theories provide acceptable approximations of the underlying *structures* that produce observable phenomena. This is another example of *selective scientific realism*.

Upon closer scrutiny, however, both entity realism and structural realism fail to square with the history of science. Let us begin with entity realism, according to which, we are justified in accepting the existence of those unobservable entities which we

manage to manipulate. This view assumes that there are claims about certain unobservable entities that maintain their positions in the mosaic despite all the changes in the claims concerning their behaviour and relations with other entities. Unfortunately, a quick glance at the history of science reveals many once-accepted unobservable entities that are no longer accepted. Consider, for instance, the theory of *phlogiston* accepted by chemists until the late 18th century. According to this theory, what makes something combustible is the presence of a certain substance, called phlogiston. Thus, firewood burns because it contains phlogiston. When burning, so the story goes, the firewood is being de-phlogisticated while the air around it becomes phlogisticated. In other words, according to this theory, phlogiston moves from the burning substance to the air. What's important for our discussion is that the existence of phlogiston, an unobservable entity, was accepted by the community of the time. Needless to say, we no longer accept the existence of phlogiston. In fact, the description of combustion we accept nowadays is drastically different. Presently, chemists accept that, in combustion, a substance composed primarily of carbon, hydrogen, and oxygen (the firewood) combines with the oxygen in the air, producing carbon dioxide and water vapor. In short, the chemical entities that we accept nowadays are very different from those that we used to accept in the 18th century. When we study the history of science, we easily find many cases when a once-accepted entity is no longer accepted. Consider, for example, the four Aristotelian elements, the electric fluids of the 18th century, or the luminiferous ether of the 19th century. How then can entity realists argue that scientists are justified in accepting the existence of unobservable entities if our knowledge of these entities is as changeable as our knowledge of structures?

Of course, an entity realist can respond by saying that those entities that we no longer accept were accepted by mistake, i.e. that scientists didn't really manage to causally manipulate them. But such a response is historically-insensitive, as it assumes that only our contemporary science succeeds in properly manipulating the unobservable entities, while the scientists of the past were mistaken in their beliefs that they managed to manipulate their unobservable entities. In reality, however, such "mistakes" only become apparent when a theory becomes rejected and replaced by a new theory that posits new unobservable entities. Chances are, one day our current unobservable entities will also be replaced by some new unobservables, as has happened repeatedly throughout history. Would entity realists be prepared to admit that our contemporary scientists weren't really justified in accepting our current unobservable entities, such as quarks, leptons, or bosons, once these entities become replaced by other unobservable entities? Clearly, that would defeat the whole purpose of entity realism, which was to select those parts of our theories that successfully approximate the world.

Structural realism faces a similar objection. Yes, it is true that sometimes our claims concerning structural relations withstand major transitions from one set of unobservable entities to another. However, this is by no means a universal feature of science. We can think of many instances from the history of science where a long-accepted proposition describing a certain relation eventually becomes rejected. Consider the Aristotelian law of violent motion that was accepted throughout the medieval and early modern periods:

If the force (F) is greater than the resistance (R) then the object will move with a velocity (V) proportional to F/R. Otherwise the object won't move.

Among other things, the law was accepted as the correct explanation of the motion of projectiles, such as that of an arrow shot by an archer. The velocity of the moving arrow was believed to depend on two factors: the force applied by the archer and the resistance of the medium, i.e. the air. The greater the applied force, the greater the velocity; the greater the resistance of the medium, the smaller the velocity. It was accepted that the initial force was due to the mover, i.e. the archer. But what type of force keeps the object, i.e. the arrow, moving *after* it has lost contact with the initial mover? Generations of medieval and early modern scholars have attempted to answer this question. Yet, it was accepted that any motion – including continued motion – necessarily requires a certain force, be it some external force or internal force stemming from the moving object itself.

Now compare this with the second law of Newtonian physics:

The acceleration (a) of a body is proportional to the net force (F) acting on the body and is inversely proportional to the mass (m) of the body:

$$a = \frac{F}{m}$$

How would we parse out the same archer-arrow case by means of this law? First, we notice that the *mass* of the arrow suddenly becomes important. We also notice that the *force* is now understood as the *net force* and it is proportional not to the velocity but to the *acceleration*, i.e. the *change* of velocity per unit time. According to the law, the greater the net force, the greater the acceleration, and the greater the mass, the smaller the acceleration. In short, the second law expresses relations that are quite different from those expressed by the Aristotelian law of violent motion. So how can a structural realist claim that our knowledge of relations is normally being preserved in one form or another?

There have been other historical cases where we accepted the existence of new relations and rejected the existence of previously accepted relations. Consider, for example, the transition from Newton's law of gravity to Einstein's field equations which took place as a result of the acceptance of general relativity ca. 1920. Here is a typical formulation of the law of gravity:

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

And here is a typical formulation of Einstein's field equation:

$$G_{\mu\nu} + g_{\mu\nu}\Lambda = \frac{8\pi G}{c^4} T_{\mu\nu}$$

The two equations are not only very different visually, but they also capture very different relations. The Newtonian law of gravity posits a certain relation between the *masses* of two objects, the *distance* between them, and the *force of gravity* with which they attract each other. In contrast, Einstein's field equation posits a relation between the *space-time curvature* expressed by the Einstein tensor ($G_{\mu\nu}$) and a specific distribution of matter and energy expressed by the stress-energy tensor ($T_{\mu\nu}$). While Newton's law tells us what the force of gravity will be given objects with certain masses at a certain distance, Einstein's equation tells us how a certain region of space-time will be curved given a certain arrangement of mass and energy in that region. Saying that the two equations somehow capture the same relation would be an unacceptable stretch.

In short, there are strong *historical* reasons against both entity realism and structural realism. Both of these versions of selective scientific realism fail to square with the history of science, which shows clearly that both our knowledge of entities and our knowledge of structures have repeatedly changed through time. This is not surprising, since as fallibilists we know that no synthetic proposition is immune to change. Clearly, there is no reason why some of these synthetic propositions – either the ones describing entities or the ones describing relations – should be any different.

In addition, there is a strong *theoretical* reason against selective scientific realism. It is safe to say that both entity realism and structural realism fail in their attempts due to a fatal flaw implicit in *any* selective approach. The goal of any selective approach is to provide some *criteria* – i.e. some *method* – that would help us separate those parts of our theories that approximate the world from those parts that are at best mere useful instruments. We've seen how both entity realism and structural realism attempted to provide their distinct methods for distinguishing acceptable parts of theories from those that are merely useful. The entity realist method of selecting acceptable parts would go along these lines: "the existence of an unobservable entity is acceptable if that entity has been successfully manipulated". Conversely, the structural realist method can be formulated as "the claim about unobservables is acceptable if it concerns a relation (structure)". Importantly, these methods were meant to be both universal and transhistorical, i.e. applicable to all fields of science in all historical periods. After our discussions in chapters 3 and 4, it should be clear why any such attempt at identifying transhistorical and universal methods is doomed. We know that methods of science are changeable: what is acceptable to one community at one historical period need not necessarily be acceptable to another community at another historical period. Even the same community can, with time, change its attitude towards a certain relation or entity.

Take, for instance, the idea of *quantum entanglement*. According to quantum mechanics, sometimes several particles interact in such a way that the state of an individual particle is dependent on the state of the other particles. In such cases, the particles are said to be *entangled*: the current state of an entangled particle cannot be characterized independently of the other entangled particles. Instead, the state of the whole entangled system is to be characterized collectively. If, for instance, we have a pair of entangled electrons, then the measurement of one electron's quantum state (e.g. spin, polarization, momentum, position) has an impact on the quantum state of the other electron. Importantly, according to quantum mechanics, entanglement can be *nonlocal*: particles can be entangled even when they are separated by great distances.

Needless to say, the existence of nonlocal entanglement didn't become immediately *accepted*, for it seemingly violated one of the key principles of Einstein's relativity theory, according to which nothing can move and no information can be transmitted faster than the speed of light. Specifically, it wasn't clear how a manipulation on one of the entangled particles can possibly affect the state of the other particle far away without transmitting this information faster than the speed of light. This was one of the reasons why the likes of Albert Einstein and Erwin Schrödinger were against accepting the notion of nonlocal entanglement. Thus, for a long time, the idea of entanglement was considered a *useful* calculating tool. But the reality of entanglement was challenged, i.e. it wasn't *accepted* as a real physical phenomenon. It was not until Alain Aspect's experiments of 1982 that the existence of nonlocal entanglement became accepted. Nowadays, it is *accepted* by the physics community that subatomic particles can be entangled over large distances.

What this example shows is that the stance towards a certain entity or a relation can change even within a single community. A community may at first be instrumentalist towards an entity or a structure, like quantum entanglement, and then may later become realist about the same entity or structure. By the second law of scientific change, the acceptance or unacceptance of a claim about an entity or a relation by a certain community depends on the respective *method* employed by that community. Thus, to assume that we as philosophers are in a position to provide transhistorical and universal criteria for evaluating what's acceptable and what's merely useful would be not only anachronistic and presumptuous but would also go against the laws of scientific change. Therefore, we have to refrain from drawing any such transhistorical and universal lines between what is acceptable and what is merely useful. In other words, the approach of selective scientific realism is untenable.

This brings us to the position that can be called *nonselective scientific realism*. According to this view, our best scientific theories do somehow approximate the world and we are justified to accept them as the best available descriptions of their respective domains, but we can never tell which specific parts of our theories are acceptable approximations and which parts are not. Nonselective scientific realism holds that any attempt at differentiating acceptable parts of our theories from merely useful parts is doomed, since all synthetic propositions – both those describing entities and those describing relations/structures – are inevitably fallible and can be replaced in the future.

The following table summarizes the major varieties of scientific realism:

<p>? Can our empirical theories provide us with a picture of the world as it really is?</p>		
Yes	No	
<p>Naïve Realism: Our theories provide us with a picture of the world just as it really is.</p>	<p>Critical Realism: All empirical theories are at best approximations.</p>	
	<p>? Is it possible to separate those parts of our theories that approximate the world from those that do not?</p>	
	Yes	No
	<p>Selective Scientific Realism: It is possible to pin down those parts of our theories that approximate.</p>	<p>Non-Selective Realism: It is impossible to pin down the approximating parts of theories.</p>
	<p>? Which parts are those?</p>	
	Structures	
<p>Structural Realism</p>	<p>Entity Realism</p>	

From Realism to Progress

Once we appreciate that the selective approach is not viable, we also understand that we can no longer attempt to draw transhistorical and universal lines between legitimate approximations and mere useful instruments. This is decided by an individual community’s employed method which can change through time. Thus, the question that a realist wants to address is not whether our theories approximate the world – they somehow do – but whether new accepted theories provide *better* approximations than the previously accepted theories. Thus, the question is that of progress:

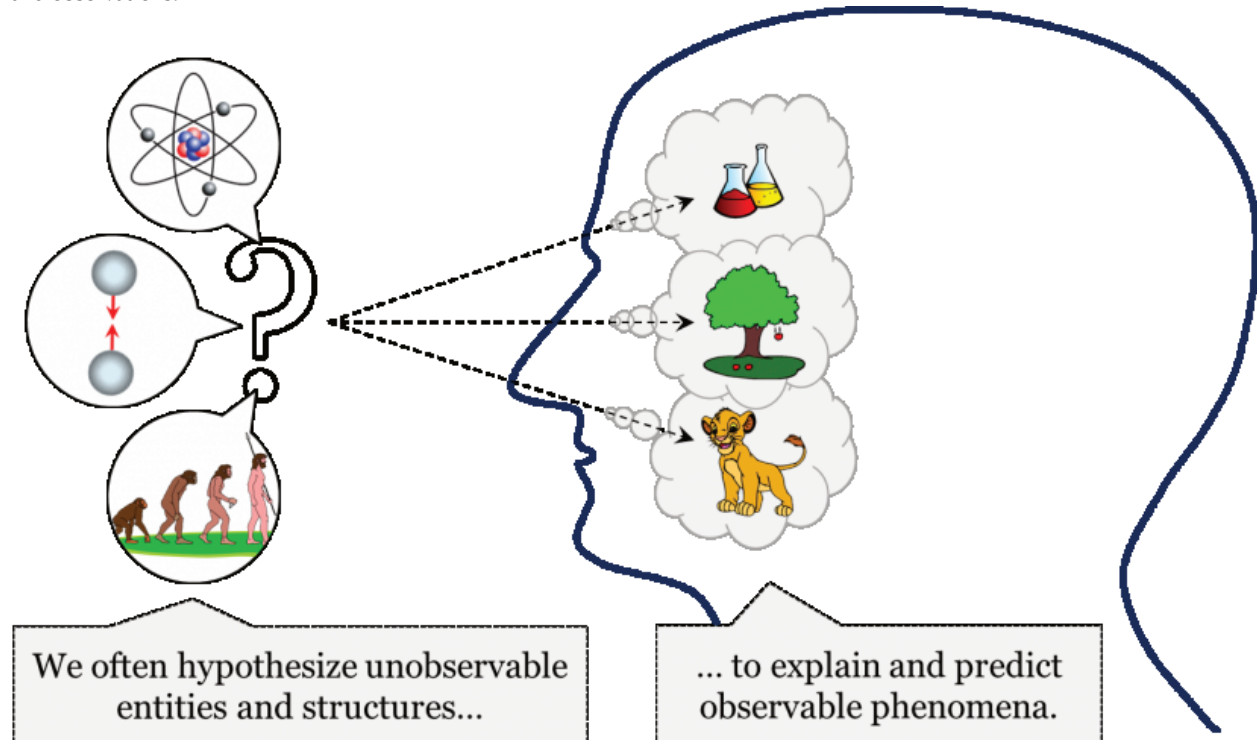
Does science actually *progress* towards truth?

Note that, while there have been many different notions of *progress*, here we are interested exclusively in progress towards the correct description of the mind-independent external world. Indeed, nobody really questions the existence of technological, i.e. instrumental progress. It is a historical fact that our ability to manipulate different phenomena has constantly increased. After all, we can now produce self-driving cars, smartphones, and fidget spinners – something we weren’t able to do not long ago. In other words, the existence of technological (instrumental, empirical) progress is beyond question. What’s at stake here is whether we also progress in our descriptions of the world as it really is. This is worth explaining.


The question of whether scientific theories progress towards truth is important in many respects. Science, after all, seems to be one of very few fields of human endeavour where we can legitimately speak of progress. The very reason why we have grant agencies funding scientific research is that we assume that our new theories can improve our understanding of the world. The belief that contemporary scientific theories are much better approximations of the world than the theories of the past seems to be implicit in the whole scientific enterprise. In short, the belief that currently accepted theories are better than those of the past is integral to our culture. This is more than can be said about many other fields of human endeavour. For example, can anyone really show that contemporary rock music is better than that of the 1970s and the 1980s? How could we even begin to compare the two? Are contemporary authors better than Jane Austen, Leo Tolstoy, or Marcel Proust? Is contemporary visual art better than that of Da Vinci or Rembrandt? It is nowadays accepted that art manages to produce different forms, i.e. different ways of approaching the subject, but the very notion that one of these forms can be better than others is considered problematic. Yet, when it comes to science, our common attitude these days is that it steadily progresses towards an increasingly better approximation of the world. Our question here is to determine whether that is really the case. What we want to establish is whether it is the case that science

gradually approximates the true description of the world, or whether we should concede that science can at best give us different ways of looking at the world, none of which is better or worse.

In science, we often postulate different hypotheses about *unobservable* entities or relations/structures so that we could explain *observable* phenomena. For instance, we hypothesized the existence of a certain attractive force that was meant to explain why apples fall down and why planets revolve in ellipses. Similarly, we hypothesized the existence of a certain atomic structure which allowed us to explain many observable chemical phenomena. We hypothesized the existence of evolution through mutation and natural selection to explain the observable features of biological species. In short, we customarily hypothesize and accept different unobservable entities and structures to explain the world of phenomena, i.e. to explain the world as it appears to us in experiments and observations.



In addition, as we have already seen, sometimes scientists change their views on what types of entities and structures populate the world: the entities and relations that were accepted a hundred years ago may or may not still be accepted nowadays. But if our views on unobservable entities and structures that purportedly populate the world change through time, can we really say that we are getting better in our knowledge of the world as it really is? There are two opposing views on this question – *the progress thesis* and *the no-progress thesis*. While the champions of the progress thesis believe that science gradually advances in its approximations of the world, the proponents of the no-progress thesis hold that we are not in a position to know whether our scientific theories progress towards truth.

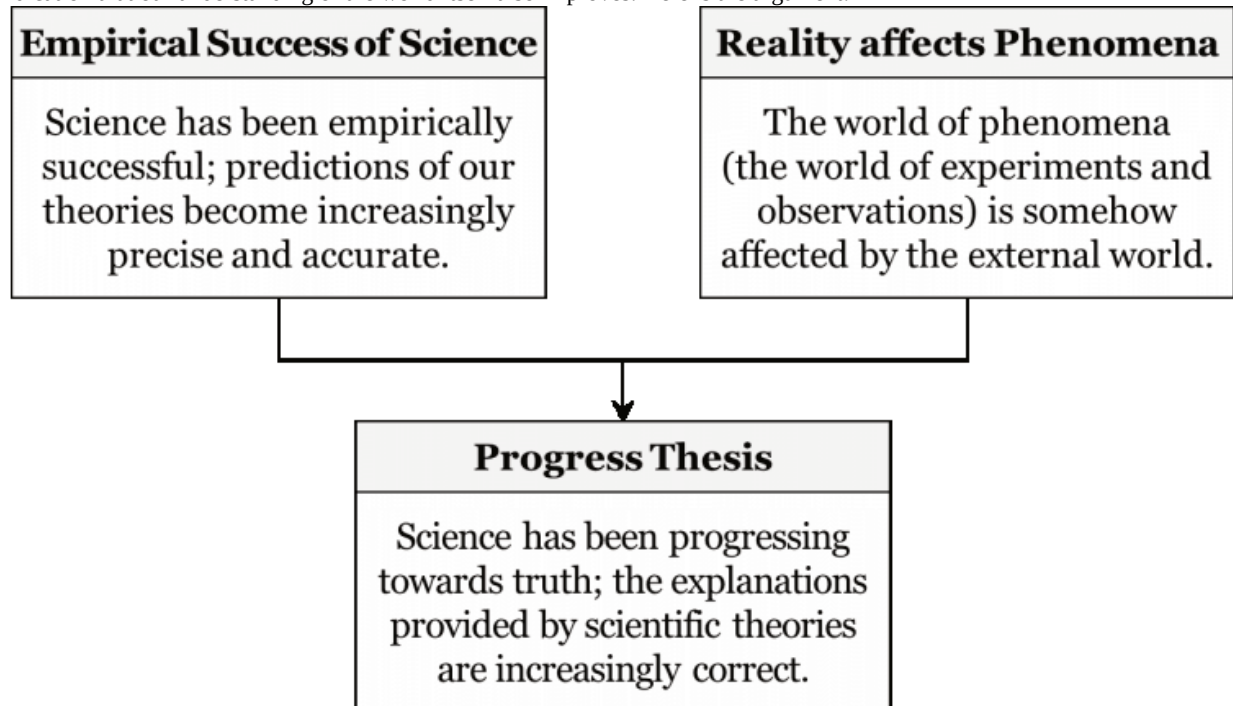
 Does science <i>actually</i> progress towards truth?	
Yes	No
<p>Progress thesis: Science progresses towards truth, i.e. scientific theories provide increasingly correct descriptions of the external world.</p>	<p>No-Progress thesis: We can't know whether science progresses towards truth, i.e. whether some descriptions are closer to truth than others.</p>

How can we find out which of these opposing parties is right? Naturally, we might be inclined to refer to the laws of scientific change and see what they have to tell us on the subject. Yet, a quick analysis reveals that the laws of scientific change as they currently stand do not really shed light on the issue of scientific progress. Indeed, consider the second law, according to which theories normally become accepted by a community when they somehow meet the acceptance criteria of the community's employed method. Now recall the definition of *acceptance*: to accept a theory means to consider it the best available description of whatever it is the theory attempts to describe. Thus, if we ask any community, they will say that their current theories are better approximation of their objects than the theories they accepted in the past. After all, we wouldn't accept general relativity if we didn't think that it is better than the Newtonian theory. Any community whatsoever will always believe that they have been getting better in their knowledge of the world. While some communities may also believe that they have already achieved the absolutely true description of a certain object, even these communities will accept that their current theories are better than the theories of the past. In other words, when we look at the process of scientific change from the perspective of any community, the historical sequence of their accepted theories will always appear progressive to them.

Clearly, this approach doesn't take us too far, since the question wasn't whether the process of scientific change *appears* progressive from the perspective of the scientific community, but whether the process is *actually* progressive. The laws of scientific change tell us that if we were to ask any scientific community, they would subscribe to the notion of progress. Yet, to find out whether we do, in fact, progress towards truth, we need a different approach. In the remainder of this chapter, we will consider the two most famous arguments for and against the progress thesis – the no-miracles argument and the pessimistic meta-induction argument.

As we have already established, nobody denies that science becomes increasingly successful in its ability to manipulate and predict observable phenomena. It is this *empirical success* of our theories that allows us to predict future events, construct all sorts of useful instruments, and drastically change the world around us. In its ability to accurately predict and manipulate observable phenomena, our 21st-century science is undeniably head and shoulders above the science of the past. But does this increasing *empirical* success of our theories mean that we are also getting closer to the true picture of the world as it really is? This is where the progress thesis and no-progress thesis drastically differ.

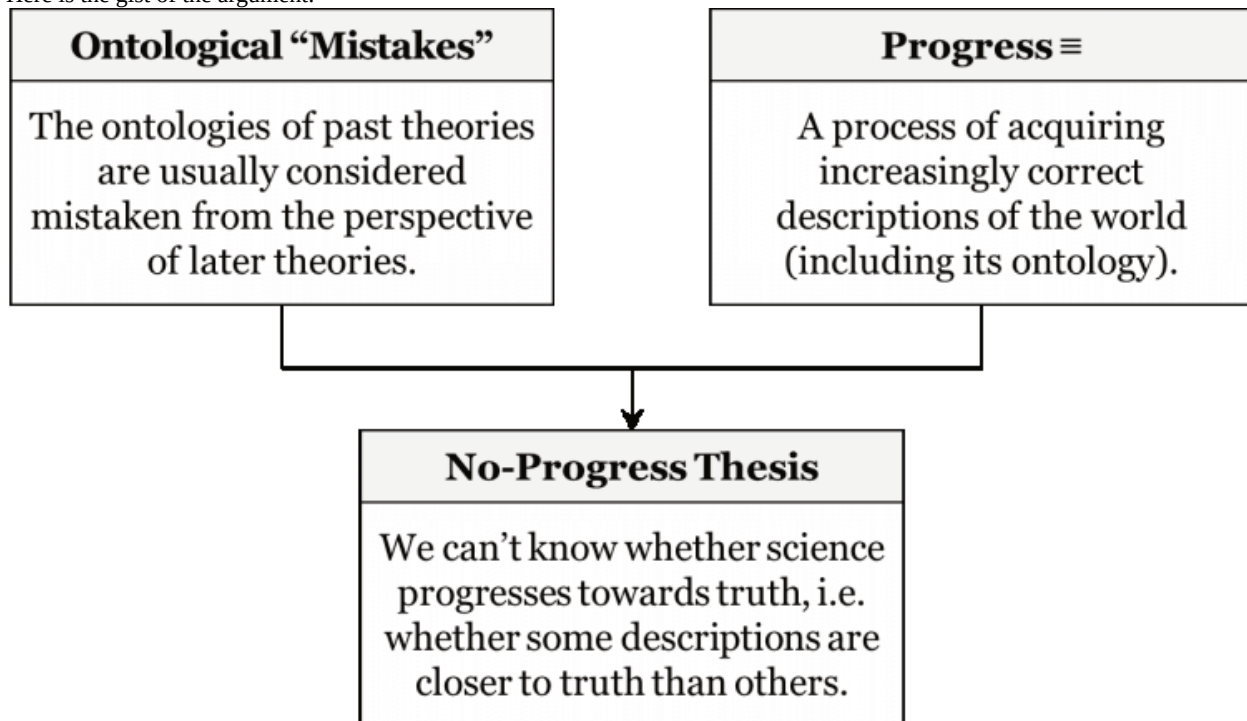
According to the champions of the progress-thesis, the empirical success of our theories is a result of our ever-improving understanding of the world as it really is. This is because the world of phenomena cannot be altogether divorced from the external world. After all, so the argument goes, the world of phenomena is an effect of the external world upon our senses: what we see, hear, smell, taste, and touch should be somehow connected to how the world really is. Of course, nobody will claim that the world as it is in reality is exactly the way we perceive it – i.e. nobody will champion the view of naïve realism these days – but isn't it reasonable to suggest that what we perceive depends on the nature of the external world *at least to some degree*? Thus, the results of experiments and observations are at least partially affected by things as they really are. But this means that by getting better in our ability to deal with the world of observable phenomena, we are also gradually improving our knowledge of the world of unobservables. In other words, as the overall predictive power of our theories increases, this is, generally speaking, a good indication that our understanding of the world itself also improves. Here is the argument:



The underlying idea is quite simple: if our theories didn't manage to get at least something right about the world as it really is,

then the empirical success of our science would simply be a miracle. Indeed, how else could we explain the unparalleled empirical success of our science if not by the fact that it becomes better and better in its approximations of the external world? Surely, if we could manage to have this much empirical success without ever getting anything right about the external world, that would be a miracle. The only reasonable explanation, say the champions of the progress thesis, is that our approximations of the world also improve, i.e. that we gradually progress towards truth. This is the gist of the famous *no-miracles argument* for scientific progress.

How would a champion of the no-progress thesis reply to this? One common reply is the so-called *pessimistic meta-induction argument*. Let us first appreciate, says a champion of the no-progress thesis, a simple historical fact: we have been quite often mistaken in our hypotheses concerning unobservable entities or structures. When we try to hypothesize what unobservable entities or structures populate the world, i.e. when we try to guess the *ontology* of the world, we often end up accepting entities and structures which we eventually come to reject as wrong. The history of science, so the argument goes, provides several great examples of this. Consider the idea of the four terrestrial elements of *earth, water, air, and fire* which were an integral part of any Aristotelian-medieval mosaic all the way into the 17th century. Similarly, recall the idea of *phlogiston* accepted in the 18th century. Also recall the idea of *the force of gravity* acting at a distance between any two objects in the universe, which was accepted until ca. 1920. It is a historical fact, say the proponents of the no-progress thesis, that our knowledge about entities and structures that populate the world has changed through time. We are often wrong in our hypotheses concerning the *ontology* of the world. As a result, we often reject old ontologies and accept new ones. But if the ontologies of our past theories are normally considered mistaken from later perspectives, how can we ever claim that our theories gradually get better in approximating the world? Shouldn't we rather be more modest and say that all we know is that we are often mistaken in our hypotheses concerning the unobservable entities and structures, that we often come to reject long-accepted ontologies, and that the ontologies of our current theories could be found, one day, to be equally mistaken? In other words, we should accept that there is no guarantee that our theories improve as approximations of the world as it really is; we are not in a position to claim that science provides increasingly correct descriptions of the world. But that is precisely what the idea of progress is all about! Thus, there is no progress in science. Here is the gist of the argument:



If the long-accepted ontology of four terrestrial elements was eventually considered mistaken, then why should our ontology of quarks, leptons, and bosons be any different? The unobservable entities and structures that our theories postulate come and go, which means that in the future the ontologies of our current theories will most likely be considered mistaken from the perspective of the theories that will come to replace them. This is the substance of what is known as the *pessimistic meta-induction argument*.

Now, why such a strange label – “pessimistic meta-induction”? In literature, the argument is often portrayed as inductive: because the ontologies of our past theories have been repeatedly rejected, we generalize and predict that the ontologies of the currently accepted theories will also one day be rejected. This is a *meta-inductive* step, as it concerns not our descriptions of the world, but our descriptions concerning our descriptions, i.e. our meta-theories (e.g. the claim that “ontologies of past theories have been rejected”). As the whole argument questions the ability of our current ontologies to withstand future challenges, it is also clearly *pessimistic*. It is important to note that while the inductive form of the argument is very popular in the philosophical literature, it can also be formulated as a *deductive* argument, as we have done above. The main reason why we nowadays believe

that our claims about unobservable entities and structures can be rejected in the future is our *fallibilism*, for which we have several theoretical reasons, such as the problem of sensations, the problem of induction, and the problem of theory-ladenness. Since we accept that all our empirical theories are fallible, we don't make any exceptions for our claims concerning the ontology of the world – and why should we? So, even though the argument is traditionally labelled as “pessimistic meta-induction”, it can be formulated in such a way as to avoid any direct use of induction. We don't have to even mention the failure of our *past* ontologies; our fallibilism alone is sufficient to claim that our current ontologies are also likely doomed.

Regardless of how the argument is formulated, its main message remains the same: we are not in a position to say there is actual progress towards truth. Since the ontologies of past theories are usually considered mistaken from the perspective of future theories, the process of scientific change produces one faulty ontology of unobservable entities and structures after another. What we end up with is essentially a graveyard of rejected ontologies – a series of transitions where one false ontology replaces another false ontology and so on. All that science gives us is different perspectives, different ways of approaching the world, which can be more or less *empirically* successful, yet none of these can be said to be approximating the world better than others. Thus, all that we can legitimately claim, according to the no-progress thesis, is that science increases its overall predictive power but doesn't take us closer to the truth.

Does this argument hold water? As opposed to the no-miracles argument, the pessimistic meta-induction argument divorces the empirical success of a theory from its ability to successfully approximate the world. Indeed, the fact that a theory is predictively accurate and can be used in practical applications doesn't necessarily make its ontology any more truthlike. Consider, for instance, the Ptolemaic geocentric astronomy which was extremely successful in its predictions of planetary positions, but postulated an ontology of eccentrics, equants, epicycles, and deferents, which was considered dubious even in the Middle Ages. In addition, the history of science provides many examples in which several theories, with completely different ontologies, were equally successful in their predictions of observable phenomena. Thus, in the early 17th century, the Copernican heliocentric theory and the Tychonic geo-heliocentric theory posited distinct ontologies but were almost indistinguishable in their predictions of observable planetary positions. Nowadays, we have a number of different quantum theories – the so-called “interpretations” of quantum mechanics – which make exactly the same predictions but postulate very different ontologies. If theories with completely different ontologies manage to be equally successful in their predictions, then how can we even choose which of these distinct ontologies to accept, let alone argue that one of them is a better approximation of the world? The champions of the no-progress thesis do a great job highlighting this discrepancy between empirical successes and approximating the world. In other words, they point out that, from the mere fact that the world of phenomena is affected by reality, it does not follow that by improving our knowledge of phenomena, we simultaneously improve our knowledge of the world as it really is. Thus, they question the validity of the no-miracles argument.




However, the pessimistic meta-induction argument has a fatal flaw, for it is based on the premise that our past ontologies are considered *mistaken* from the perspective of future theories. If we are truly fallibilists, then we should be very careful when deeming ontologies of the past as false in the *absolute* sense. Instead, we should accept that they are not absolutely false, but contain at least some *grains* of truth, i.e. that they somehow approximate the world, albeit imperfectly. For instance, we eventually came to reject the ontology of four elements, but we don't think it was absolutely false. Instead, we think it contained some grains of truth, as it clearly resembles the contemporary idea of the four states of matter: *solid*, *liquid*, *gas*, and *plasma*. Similarly, we no longer accept the theory of phlogiston, but saying that its ontology was absolutely wrong would be a stretch; it was an approximation – a pretty bad one to be sure, but an approximation nevertheless. The fallibilist position is not that the old ontologies are strictly false, but that the ontology of general relativity and quantum physics is slightly better than the ontology of classical physics, just as the ontology of classical physics was slightly better than the ontology of Descartes' natural philosophy, which itself was slightly better than that of Aristotle. Thus, we can't say we commit ontological mistakes in the absolute sense. The old ontologies are rejected not because they are considered absolutely false, but because we think we have something better. A useful way of thinking of it is as a series of photographs of the same person – from the blurriest to the sharpest. Compared to the sharp photographs, the blurry photographs would be worse approximations of the person's appearance; yet, importantly, they are all approximations. In short, the premise of ontological “mistakes” doesn't hold water, and thus the whole argument is unsound.

Summary

We commenced this chapter by posing the question of scientific realism: do our best theories correctly describe the mind-independent external world? We have learned that the central point of contention between realists and instrumentalists doesn't concern our ability to describe what is immediately *observable*, but our ability to provide trustworthy descriptions of *unobservable* entities and structures. While both parties agree that theories can be legitimately used in practical applications, scientific realists also believe that we can legitimately *accept* the claims of our theories about unobservables. We have discussed a number of sub-species of scientific realism. We've also seen how selective approaches fail in their attempts to differentiate acceptable parts of our theories from those that are merely useful. Our knowledge about both unobservable entities and unobservable structures changes through time and there is no transhistorical and universal method that would indicate which parts of our theories are to be accepted and which only used. Only the actual methods employed by a given community at a given time can answer that question.

We then suggested that there is a more interesting question to discuss – that of scientific progress. While it is generally agreed that our scientific theories have been enormously successful in dealing with observable phenomena, there is a heated debate on whether scientific theories gradually progress towards ever-improving approximations of the world. We've seen that the main

argument for the no-progress thesis – the pessimistic meta-induction argument – has a serious flaw. The debate can be summed up in the following table:

 Does science <i>actually progress</i> towards truth?	
Yes	No
<p>Progress thesis: Science progresses towards truth, i.e. scientific theories provide increasingly correct descriptions of the external world.</p> <p> No-Miracles Argument</p>	<p>No-Progress thesis: We can't know whether science progresses towards truth, i.e. whether some descriptions are closer to truth than others.</p> <p> Pessimistic Induction Argument</p>

While we can tentatively say that the progress thesis seems slightly better supported, we have to be cautious not to jump to the conclusion that the debate is over. Far from it: the question of progress is central to the contemporary philosophy of science and is a subject of continuous discussions.

6.

Science and Non-Science -- Introduction to History and Philosophy of Science

Chapter 6: Science and Non-Science

Intro

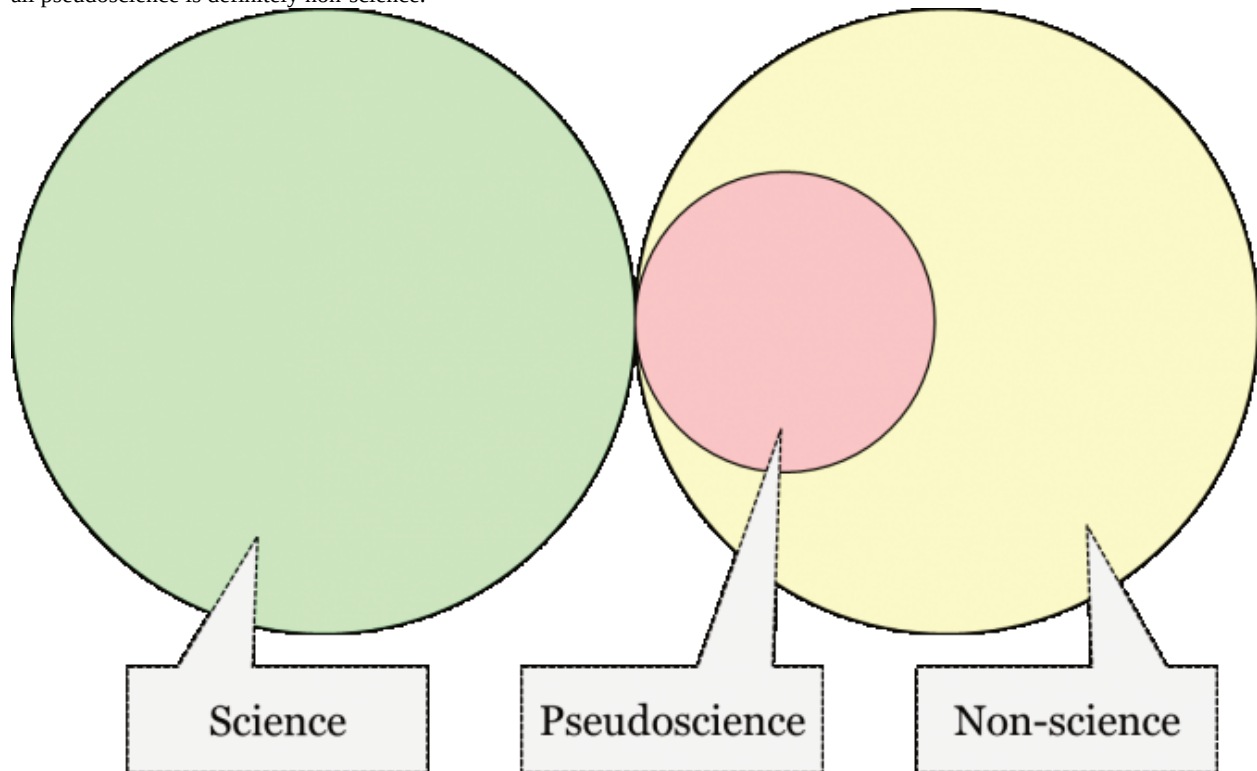
So far, we have discussed scientific knowledge, scientific methods, scientific change, and scientific progress. Despite all these philosophical investigations of science, we haven't yet had a focused discussion on what makes *science* what it is, or what differentiates it from other human endeavours. We have taken for granted that science is something different – something unique. But what actually makes science different? What makes it unique?

In philosophy this question has been called *the demarcation problem*. To demarcate something is to set its boundaries or limits, to draw a line between one thing and another. For instance, a white wooden fence *demarcates* my backyard from the backyard of my neighbour, and a border demarcates the end of one country's territory and the beginning of another's. The demarcation problem in the philosophy of science asks:

What is the difference between *science* and *non-science*?

In other words, what line or "fence" – if any – separates science from non-science, and where exactly does science begin and non-science end?

Historically, many philosophers have sought to demarcate science from non-science. However, often, their specific focus has been on the demarcation between science and *pseudoscience*. Now, what is pseudoscience and how is it different from non-science in general? Pseudoscience is a very specific subspecies of non-science which *masks* itself as science. Consider, for instance, the champions of intelligent design, who essentially present their argument for the existence of God as a properly scientific theory which is purportedly based on scientific studies in fields such as molecular biology and evolutionary biology but incorporates both blatant and subtle misconceptions about evolutionary biology. Not only is the theory of intelligent design unscientific, but it is pseudoscientific, as it camouflages and presents itself as a legitimate science. In short, while not all non-science is pseudoscience, all pseudoscience is definitely non-science.



While pseudoscience is the most dangerous subspecies of non-science, philosophical discussions of the problem of demarcation aim to extract those features that make science what it is. Thus, they concern the distinction between science and non-science in general, not only that between science and pseudoscience. So, our focus in this chapter is not on pseudoscience exclusively, but on the general demarcation between science and non-science.

Practical Implications

As with most philosophical questions concerning science, this question too has far reaching practical implications. The question of demarcation is of great importance to *policy-making*, *courts*, *healthcare*, *education*, and *journalism*, as well as for the proper functioning of *grant agencies*. To appreciate the practical importance of the problem of demarcation, let's imagine what would happen if there was no way of telling science from non-science. Let's consider some of these practical implications in turn.

Suppose a certain epistemic community argues that we are facing a potential environmental disaster: say, an upcoming massive earthquake, an approaching asteroid, or slow but steady global warming. How seriously should we take such a claim? Naturally our reaction would depend on how trustworthy we think the position of this community is. We would probably not be very concerned, if this was a claim championed exclusively by an unscientific – or worse, pseudoscientific – community. However, if the claim about looming disaster was accepted by a *scientific* community, it would likely have serious effect on our *environmental policy* and our decisions going forward. But this means that we need to have a way of telling what's science and what's not.

The ability to demarcate science from non-science and pseudoscience is equally important in *courts*, which customarily rely on the testimony of experts from different fields of science. Since litigating sides have a vested interest in the outcome of the litigation, they might be inclined towards using any available “evidence” in their favour, including “evidence” that has no scientific foundation whatsoever. Thus, knowing what's science and what's not is very important for the proper function of courts. Consider, for example, the ability to distinguish between claimed evidence obtained by psychic channelling, and evidence obtained by the analysis of DNA found in blood at the scene of the crime.

The demarcation of science from non-science is also crucial for *healthcare*. It is an unfortunate fact that, in medicine, the promise of an easy profit often attracts those who are quick to offer “treatments” whose therapeutic efficacy hasn't been properly established. Such “treatments” can have health- and even life-threatening effects. Thus, any proper health care system should use only those treatments whose therapeutic efficacies have been scientifically established. But this assumes a clear understanding as to what's science and what merely masks itself as such.

A solid *educational system* is one of the hallmarks of a contemporary civilized society. It is commonly understood that we shouldn't teach our children any pseudoscience but should build our curricula around knowledge accepted by our scientific community. For that reason, we don't think astrology, divination, or creation science have any place in school or university curricula. Of course, sometimes we discuss these subjects in history and philosophy of science courses, where they are studied as examples of non-science or as examples of what was once considered scientific but is currently deemed unscientific. Importantly, however, we don't present them as accepted science. Therefore, as teachers, we must be able to tell pseudoscience from science proper.

In recent years, there have been several organized campaigns to portray pseudoscientific theories as bearing the same level of authority as the theories accepted by proper science. With the advent of social media, such as YouTube or Facebook, this becomes increasingly easy to orchestrate. Consider, for instance, the deniers of climate change or deniers of the efficacy of vaccination who have managed – through orchestrated *journalism* – to portray their claims as a legitimate stance in a scientific debate. Journalists should be properly educated to know the difference between science and pseudoscience, for otherwise they risk hampering public opinion and dangerously influencing policy-makers. Once again, this requires a philosophical understanding on how to demarcate science from non-science.

Finally, scientific *grant agencies* heavily rely on certain demarcation criteria when determining what types of research to fund and what types of research not to fund. For instance, these days we clearly wouldn't fund an astrological project on the specific effect of, say, Jupiter's moons on a person's emotional makeup, while we would consider funding a psychological project on the effect of school-related stress on the emotional makeup of a student. Such decisions assume an ability to demarcate a scientific project from unscientific projects.

In brief, the philosophical problem of demarcation between science and non-science is of great practical importance for a contemporary civilized society and its solution is a task of utmost urgency. While hopefully science's general boundaries have started to come into view as we've surveyed it over the last five chapters, in this final philosophical chapter we will attempt to bring them into sharper focus.

What are the Characteristics of a Scientific Theory?

Traditionally, the problem of demarcation has dealt mainly with determining whether certain *theories* are scientific or not. That is, in order to answer the more general question of distinguishing science and non-science, philosophers have focused on answering the more specific question of identifying features that distinguish scientific *theories* from unscientific *theories*. Thus, they have been concerned with the question:

What are the characteristics of a scientific *theory*?

This more specific question treats the main distinction between science and non-science as a distinction between two different kinds of theories. Philosophers have therefore been trying to determine what features *scientific theories* have which *unscientific theories* lack. Consider for instance the following questions:

Why is the theory of evolution scientific and creationism unscientific?

Is the multiverse theory scientific?

Are homeopathic theories pseudoscience?

Our contemporary scientific community answers questions like these on a regular basis, assessing theories and determining whether those theories fall within the limits of science or sit outside those limits. That is, the scientific community seems to have an implicit set of *demarcation criteria* that it employs to make these decisions.

You may recall that we mentioned *demarcation criteria* back in chapter 4 as one of the three components of a scientific method, along with *acceptance criteria* and *compatibility criteria*. A scientific method consists of *all* criteria actually employed in theory assessment. Demarcation criteria are a specific subset of those criteria which are employed to assess whether a theory is scientific or not.

So, what are the demarcation criteria that scientists employ to evaluate whether a theory is scientific? First, let's look at our implicit expectations for what counts as a science and what doesn't. What are our current demarcation criteria? What criteria does the contemporary scientific community employ to determine which theories are scientific and which are not? Can we discover what they are and make them explicit? We can, but it will take a little bit of work. We'll discuss a number of different characteristics of scientific theories and see whether those characteristics meet our contemporary implicit demarcation criteria. By considering each of these characteristics individually, one step at a time, hopefully we can refine our initially proposed criteria and build a clearer picture of what our implicit demarcation criteria actually are.

Note that, for the purposes of this exercise, we will focus on attempting to explicate our contemporary demarcation criteria for *empirical science* (as opposed to formal science). As we have learned in chapter 2, empirical theories consist, not merely of analytic propositions (i.e. definitions of terms and everything that follows from them), but also of *synthetic propositions* (i.e. claims about the world). This is true by definition: a theory is said to be empirical if it contains at least one synthetic proposition. Therefore, empirical theories are not true by definition; they can either be confirmed by our experiences or contradicted by them. So, propositions like "the Moon orbits the earth at an average distance of 384,400 km", or "a woodchuck could chuck 500 kg of wood per day", or "aliens created humanity and manufactured the fossil record to deceive us" are all empirical theories because they could be confirmed or contradicted by experiments and observations. We will therefore aim to find out what our criteria are for determining whether an *empirical theory* is scientific or not.

First, let us appreciate that not all empirical theories are scientific. Consider the following example:

Theory A: You are currently in Horseheads, New York, USA.

That's right: we, the authors, are making a claim about *you*, the reader. Right now. *Theory A* has all the hallmarks of an empirical theory: It's not an analytic proposition because it's not true by definition; depending on your personal circumstances, it might be correct or incorrect. But it's not *based on experience* because we, the authors, have no reason to think that you are, in fact, in Horseheads, NY: we've never seen you near Hanover Square, and there is no way you'd choose reading your textbook over a day at the Arnot Mall. *Theory A* is a genuine claim about the world, but it is a claim that is in a sense "cooked-up" and based on no experience whatsoever. Here are two other examples of empirical theories *not* based on experience:

Theory B: The ancient Romans moved their civilization to an underground location on the far side of the moon.

Theory C: A planet 3 billion light years from Earth also has a company called Netflix.

Therefore, we can safely conclude that not every empirical theory can be said to be scientific. If that is so, then what makes a particular empirical theory *scientific*? Let's start out by suggesting something simple.

Suggestion 1: An empirical theory is scientific if it is based on experience.

This seems obvious, or maybe not even worth mentioning. After all, don't all empirical theories have to be based on experience? Suggestion 1 is based on the fact that we don't want to consider empirical theories like *A*, *B*, and *C* to be *scientific* theories. Theories that we can come up with on a whim, grounded in no experience whatsoever, do not strike us as scientific. Rather, we expect that even simple scientific theories must be somehow grounded in our experience of the world.

This basic contemporary criterion that empirical theories be grounded in our experience has deep historical roots but was perhaps most famously attributed to British philosopher John Locke (1632–1704) in his text *An Essay Concerning Human Understanding*. In this work, Locke attempted to lay out the limits of human understanding, ultimately espousing a philosophical position known today as empiricism. *Empiricism* is the belief that all synthetic propositions (and consequently, all empirical theories) are justified by our sensory experiences of the world, i.e. by our experiments and observations. Empiricism stands against the position of *apriorism* (also often referred to as "rationalism") – another classical conception that was advocated by the likes of René Descartes and Gottfried Wilhelm Leibniz. According to apriorists, there are at least some fundamental synthetic propositions which are knowable independently of experiments and observations, i.e. *a priori* (in philosophical discussions, "a priori" means "knowable independently of experience"). It is this idea of a priori synthetic propositions that apriorists accept and empiricists deny. Thus, the criterion that all physical, chemical, biological, sociological, and economical theories must be justified by experiments and observations *only*, can be traced back to empiricism.

But is this basic criterion sufficient to properly demarcate scientific empirical theories from unscientific ones? If an empirical theory is based on experience, does that automatically make it scientific?

Perhaps the main problem with *Suggestion 1* can best be illustrated with an example. Consider the contemporary opponents of the use of vaccination, called "anti-vaxxers". Many anti-vaxxers today accept the following theory:

Theory D: Vaccinations are a major contributing cause of autism.

This theory results from sorting through incredible amounts of medical literature and gathering patient testimonials. Theory D is clearly an *empirical* theory, and – interestingly – it's also in some sense *based on experience*. As such, it seems to satisfy the criterion we came up with in *Suggestion 1*: Theory D is both *empirical* and is *based on experience*.

However, while being based on experience, Theory D also results from *willingly ignoring some of the known data on that topic*. A small study by Andrew Wakefield, published in *The Lancet* in 1998, became infamous around 2000-2002 when the UK media caught hold of it. In that article, the author hypothesized an alleged link between the measles vaccine and autism despite a small sample size of only 12 children. Theory D fails to take into account the sea of evidence suggesting both that no such link (between vaccines and autism) exists, and that vaccines are essential to societal health.

In short, Suggestion 1 allows for theories that have “cherry-picked” their data to be considered scientific, since it allows scientific theories to be based on any arbitrarily selected experiences whatsoever. This doesn’t seem to jibe with our implicit demarcation criteria. Theories like Theory D, while based on experience, aren’t generally considered to be scientific. As such, we need to refine the criterion from Suggestion 1 to see if we can avoid the problems illustrated by the anti-vaxxer Theory D. Consider the following alternative:

Suggestion 2: An empirical theory is considered scientific if it explains all the known facts of its domain.

This new suggestion has several interesting features that are worth highlighting.

First, note that it requires a theory to explain *all* the known facts of its domain, and not only a selected – “cherry-picked” subset of the known facts. By ensuring that a scientific empirical theory explains the “known facts,” Suggestion 2 is clearly committed to being “based on experience”. In this it is similar to Suggestion 1. However, Suggestion 2 also stipulates that a theory must be able to explain *all* of the known facts of its domain precisely to avoid the cherry-picking exemplified by the anti-vaxxer Theory D. As such, Suggestion 2 excludes theories that clearly cherry-pick their evidence and disqualifies such fabricated theories as unscientific. Therefore, theories which choose to ignore great swathes of relevant data, such as decades of research on the causes of autism, can be deemed unscientific by Suggestion 2.

Also, Suggestion 2 explicitly talks about the facts within a certain *domain*. A domain is an area (field) of scientific study. For instance, life and living processes are the *domain* of biology, whereas the Earth’s crust, processes in it, and its history are the *domain* of geology. By specifying that an empirical theory has to explain the known facts of *its* domain, Suggestion 2 simply imposes more realistic expectations: it doesn’t expect theories to explain all the known facts from all fields of inquiry. In other words, it doesn’t stipulate that, in order to be scientific, a theory should explain *everything*. For instance, if an empirical theory is about the causes of autism (like Theory D), then the theory should merely account for the known facts regarding the causes of autism, *not* the known facts concerning black holes, evolution of species, or inflation.

Does Suggestion 2 hold water? Can we say that it correctly explicates the criteria of demarcation currently employed in empirical science? The short answer is: not quite.

When we look at general empirical theories that we unproblematically consider scientific, like the theory of general relativity or the theory of evolution by natural selection, we notice that even they may fail to meet the stringent requirements of Suggestion 2. Indeed, do our best scientific theories explain *all* the known facts of their respective domains? Can we reasonably claim that the current biological theories explain *all* the known biological facts? Similarly, can we say that our accepted physical theories explain *all* the known physical phenomena?

It is easy to see that even our best accepted scientific theories today cannot account for absolutely every known piece of data in their respective domains. It is a known historical fact that scientific theories rarely succeed in explaining all the known phenomena of their domain. In that sense, our currently accepted theories are no exception.

Take, for example, the theory of evolution by natural selection. Our contemporary scientific community clearly considers evolutionary theory to be a proper scientific theory. However, it is generally accepted that evolution itself is a very slow process. For the first 3.5 billion years of life’s history on Earth, organisms evolved from simple single-celled bacterium-like organisms to simple multicellular organisms like sponges. About 500 million years ago, however, there was a major – relatively sudden (on a geological timescale of millions of years) – diversification of life on Earth which scientists call the *Cambrian explosion*, wherein we see the beginnings of many of the forms of animal life we are familiar with today, such as arthropods, molluscs, chordates, etc. Nowadays, biologists accept both the existence of the Cambrian explosion and the theory of evolution by natural selection. Nevertheless, the theory doesn’t currently explain the phenomenon of the Cambrian explosion. In short, what we are dealing with here is a well-known fact in the domain of biology, which our accepted biological theory doesn’t explain.

What this example demonstrates is that scientific theories do not always explain *all* the known facts of its domain. Thus, if we were to apply Suggestion 2 in actual scientific practice, we would have to exclude virtually all of the currently accepted scientific empirical theories. This means that Suggestion 2 cannot possibly be the correct explication of our current implicit demarcation criterion. So, let’s make a minor adjustment:

Suggestion 3: An empirical theory is scientific if it explains, by and large, the known facts of its domain.

Like Suggestion 2, this formulation of our contemporary demarcation criterion also ensures that scientific theories are based on experience *and* can’t merely cherry-pick their data. However, it introduces an important clause – “by and large” – and thus clarifies that an empirical theory simply has to account for the great majority of the known facts of its domain. Note that this new clause is not quantitative: it doesn’t stipulate what percentage of the known facts must be explained. The clause is qualitative, as it requires a theory to explain virtually all but not necessarily all the known facts of its domain. This simple adjustment accomplishes an important task of imposing more realistic requirements. Specifically, unlike Suggestion 2, Suggestion 3 avoids excluding theories like the theory of evolution from science.

How well does Suggestion 3 square with the actual practice of science today? Does it come close to correctly explicating the actual demarcation criteria employed nowadays in empirical science?

To the best of our knowledge, Suggestion 3 seems to be a *necessary condition* for an empirical theory to be considered scientific

today. That is, any theory that the contemporary scientific community deems scientific *must* explain, by and large, the known facts of its domain. Yet, while the requirement to explain most facts of a certain domain seems to be a *necessary* condition for being considered a scientific theory today, it is not a *sufficient* condition. Indeed, while every scientific theory today seems to satisfy the criterion outlined in Suggestion 3, that criterion on its own doesn't seem to be sufficient to demarcate scientific theories from unscientific theories. The key problem here is that some famous unscientific theories also manage to meet this criterion.

Take, for example, the theory of astrology. Among many other things, astrology claims that processes on the Earth are least partially due to the influence of the stars and planets. Specifically, astrology suggests that a person's personality and traits depend crucially on the specific arrangement of planets at the moment of their birth. As the existence of such a connection is far from trivial, astrology can be said to contain synthetic propositions, by virtue of which it can be considered an empirical theory. Now, it is clear that astrology is notoriously successful at explaining the known facts of its domain. If the Sun was in the constellation of Taurus at the moment of a person's birth, and this person happened to be even somewhat persistent, then this would be in perfect accord with what astrology says about Taurus. But even if this person was not persistent at all, a trained astrologer could still explain the person's personality traits by referring to the subtle influences of other celestial bodies. No matter how much a person's personality diverges from the description of their "sign", astrology somehow always finds a way to explain it. As such, astrology can be considered an empirical theory that explains, by and large, the known facts of its domain, as per Suggestion 3.

This means that while Suggestion 3 seems to faithfully explicate a necessary part of our contemporary demarcation criteria, there must be at least *another necessary condition*. It should be a condition that the theory of evolution and other scientific theories satisfy, while astrology and other unscientific theories do not. What can this additional condition be?

One idea that comes to mind is that of *testability*. Indeed, it seems customary in contemporary empirical science to expect a theory to be testable. Thus, it seems that in addition to explaining, by and large, the known facts of their domains, scientific empirical theories are also expected to be *testable, at least in principle*.

It is important to appreciate that what's important here is not whether we as a scientific community currently have the technical means and financial resources to test the theory. No, what's important is whether a theory is testable in principle. In other words, we seem to require that there be a *conceivable* way of testing a theory, regardless of whether it is or isn't possible to conduct that testing in practice. Suppose there is a theory that makes some bold claims about the structure and mechanism of a certain subatomic process. Suppose also that the only way of testing the theory that we could think of is by constructing a gigantic particle accelerator the size of the solar system. Clearly, we are not in a position to actually construct such an enormous accelerator for obvious technological and financial reasons. Such issues actually arise in string theory, a pursued attempt to combine quantum mechanics with general relativity theory into a single consistent theory. They are a matter of strong controversy among physicists and philosophers. What seems to matter to scientists is merely the ability of a theory to be tested *in principle*. In other words, even if we have no way of testing a theory currently, we should at least be able to conceive of a means of testing it. If there is no conceivable way of comparing the predictions of a theory to the results of experiments or observations, then it would be considered *untestable*, and therefore unscientific.

But what exactly does the requirement of *testability* imply? How should testability itself be understood? In the philosophy of science, there have been many attempts to clarify the notion of testability. Two opposing notions of *testability* are particularly notable – *verifiability* and *falsifiability*. Let's consider these in turn.

Among others, Rudolph Carnap suggested that an empirical theory is scientific if it has the possibility of being *verified* in experiments and observations. For Carnap, a theory was considered *verified* if predictions of the theory could be *confirmed* through experience. Take the simple empirical theory:

Theory E: The light in my refrigerator turns off when I close the door.

If I set up a video camera inside the fridge so that I could see that the light does, indeed, turn off whenever I close the door, Carnap would consider the theory to be *verified* by my experiment. According to Carnap, every scientific theory is like this: we can, in principle, find a way to test and confirm its predictions. This position is called *verificationism*. According to *verificationism*, an empirical theory is scientific if it is possible to confirm (verify) the theory through experiments and observations.

Alternatively, Karl Popper suggested that an empirical theory is scientific if it has the possibility of being *falsified* by experiments and observations. Whereas Carnap focused on the ability of theories to become verified by experience, Popper held that what truly makes a theory scientific is its potential ability to be *disconfirmed* by experiments and observation. Science, according to Popper, is all about bold conjectures which are tested and tentatively accepted until they are falsified by counterexamples. The ability to withstand any conceivable test, for Popper, is not a virtue but a vice that characterizes all unscientific theories. What makes *Theory E* scientific, for Popper, is the fact that we can imagine the possibility that what I see on the video camera when I close my fridge might *not* match my theory. If the light, in fact, *does not* turn off when I close the door a few times, then Theory E would be considered *falsified*. What matters here is not whether a theory has or has not *actually* been falsified, or even if we have the technical means to falsify the theory, but whether its falsification is *conceivable*, i.e. whether there can, in principle, be an observational outcome that would falsify the theory. According to *falsificationism*, a theory is scientific if it can conceivably be shown to conflict with the results of experiments and observations.

Falsifiability and verifiability are two distinct interpretations of what it means to be testable. While both verifiability and falsifiability have their issues, the requirement of falsifiability seems to be closer to the current expectations of empirical scientists. Let's look briefly at the theory of young-Earth creationism to illustrate why.

Young-Earth creationists hold that the Earth, and all life on it, was directly created by God less than 10,000 years ago. While

fossils and the layers of the Earth's crust appear to be millions or billions of years old according to today's accepted scientific theories, young-Earth creationists believe that they are not. In particular, young-Earth creationists subscribe to:

Theory F: Fossils and rocks were created by God within the last 10,000 years but were made by God to appear like they are over 10,000 years old.

Now, is this theory testable? The answer depends on whether we understand *testability* as verifiability or falsifiability. Let's first see if *Theory F* is verifiable.

By the standards of verificationism, *Theory F* is verifiable, since it can be tested and confirmed by the data from experiments and/or observations. This is so because any fossil, rock, or core sample that is measured to be older than 10,000 years will actually confirm the theory, since *Theory F* states that such objects were *created to seem that way*. Every ancient object further confirms *Theory F*, and – from the perspective of verificationism – these confirmations would be evidence of the theory's testability. Therefore, if we were to apply the requirement of verifiability, young-Earth creationism would likely turn out scientific.

In contrast, by the standards of falsificationism, *Theory F* is *not* falsifiable; we can try and test it as much as we please, but we can never show that *Theory F* contradicts the results of experiments and observations. This is so because even if we were to find a trillion-year-old rock, it would not in any way contradict *Theory F*, since proponents of *Theory F* would simply respond that God made the rock to seem one trillion years old. *Theory F* is formulated in such a way that no new data, no new evidence, could ever possibly contradict it. As such, from the perspective of falsificationism, *Theory F* is untestable and, thus, unscientific.

Understanding a theory's *testability* as its *falsifiability* seems to be the best way to explicate this second condition of our contemporary demarcation criteria: for the contemporary scientific community, to say that a theory is *testable* is to say that it's *falsifiable*, i.e. that it can, in principle, contradict the results of experiments and observations. With this understanding of *testability* clarified, it seems we have our second necessary condition for a theory to be considered scientific. To sum up, in our contemporary empirical science, we seem to consider a theory scientific if it explains, by and large, the known facts of its domain, and it is testable (falsifiable), at least in principle:

Contemporary Demarcation Criteria

An empirical theory is scientific if it explains, by and large, the known facts of its domain and it is testable (falsifiable), at least in principle.

We began this exercise as an attempt to explicate our contemporary criteria for demarcation, and we've done a lot of work to distill the contemporary demarcation criteria above. It is important to note that this is merely *our attempt* at explicating the contemporary demarcation criteria. Just as with any other attempt to explicate a community's method, our attempt may or may not be successful. Since we were trying to make explicit those *implicit* criteria employed to demarcate science from non-science, even this two-part criterion might still need to be refined further. It is quite possible that the actual demarcation criteria employed by empirical scientists are much more nuanced and contain many additional clauses and sub-clauses. That being said, we can take our explication as an acceptable first approximation of the contemporary demarcation criteria employed in empirical science.

This brings us to one of the central questions of this chapter. Suppose, for the sake of argument, that the contemporary demarcation criteria are along the lines of our explication above, i.e. that our contemporary empirical science indeed expects scientific theories to explain by and large, the known facts of its domain, and be in principle falsifiable. Now, can we legitimately claim that these same demarcation criteria have been employed in all time periods? That is, could these criteria be the *universal* and *transhistorical* criteria of demarcation between scientific and unscientific theories? More generally:

Are there universal and transhistorical criteria for demarcating scientific theories from unscientific theories?

The short answer to this question is *no*. There are both theoretical and historical reasons to believe that the criteria that scientists employ to demarcate scientific from unscientific theories are neither fixed nor universal. Both the history of science and the laws of scientific change suggest that the criteria of demarcation can differ drastically across time periods and fields of inquiry. Let's consider the historical and theoretical reasons in turn.

For our theoretical reason, let's look at the laws of scientific change. Recall the third law of scientific change, *the laws of method employment*, which states that newly employed methods are the deductive consequences of some subset of other accepted theories and employed methods. As such, when theories change, methods change with them. This holds equally for the criteria of acceptance, criteria of compatibility, and criteria of demarcation. Indeed, since the demarcation criteria are *part of the method*, demarcation criteria change in the same way that all other criteria do: they become employed when they follow deductively from accepted theories and other employed methods. As such, the demarcation criteria are not immune to change, and therefore our *contemporary* demarcation criteria – whatever they are – cannot be universal or unchangeable.

This is also confirmed by historical examples. The historical reason to believe that our contemporary demarcation criteria are neither universal nor unchangeable is that *there have been other demarcation criteria employed in the past*. Consider, for instance, the criteria of demarcation that were employed by many Aristotelian-Medieval communities. As we already known, one

of the essential elements of the Aristotelian-Medieval mosaic was the idea that all things not crafted by humans have a *nature*, an indispensable quality that makes a thing what it is. It was also accepted that an experienced person can grasp this nature through *intuition* schooled by experience. We've already seen in chapter 4 how the Aristotelian-Medieval method of intuition was a deductive consequence of these two accepted ideas. According to their *acceptance criteria*, a theory was expected to successfully grasp the nature of a thing in order to become accepted. In their *demarcation criteria*, they stipulated that a theory should at least *attempt* to grasp the nature of a thing under study, regardless of whether it actually succeeded in doing so. Thus, we can explicate the Aristotelian-Medieval demarcation criterion as:

Aristotelian Demarcation Criteria

An empirical theory is scientific if it attempts to uncover the nature of a thing.

Thus, both natural philosophy and natural history were thought to be scientific: while *natural philosophy* was considered scientific because it attempted to uncover the nature of physical reality, *natural history* was scientific for attempting to uncover the nature of each creature in the world. *Mechanics*, however, was *not* considered scientific precisely because it dealt with things crafted by humans. As opposed to natural things, *artificial things* were thought to have no intrinsic nature, but were created by a craftsman for the sake of something else. Clocks, for instance, don't exist for their own sake, but for the sake of timekeeping. Similarly, ships don't have any nature, but are built to navigate people from place to place. Thus, according to the Aristotelians, the study of these artefacts, mechanics, is not scientific, since there is no nature for it to grasp. We will revisit this distinction between artificial and natural in chapter 7.

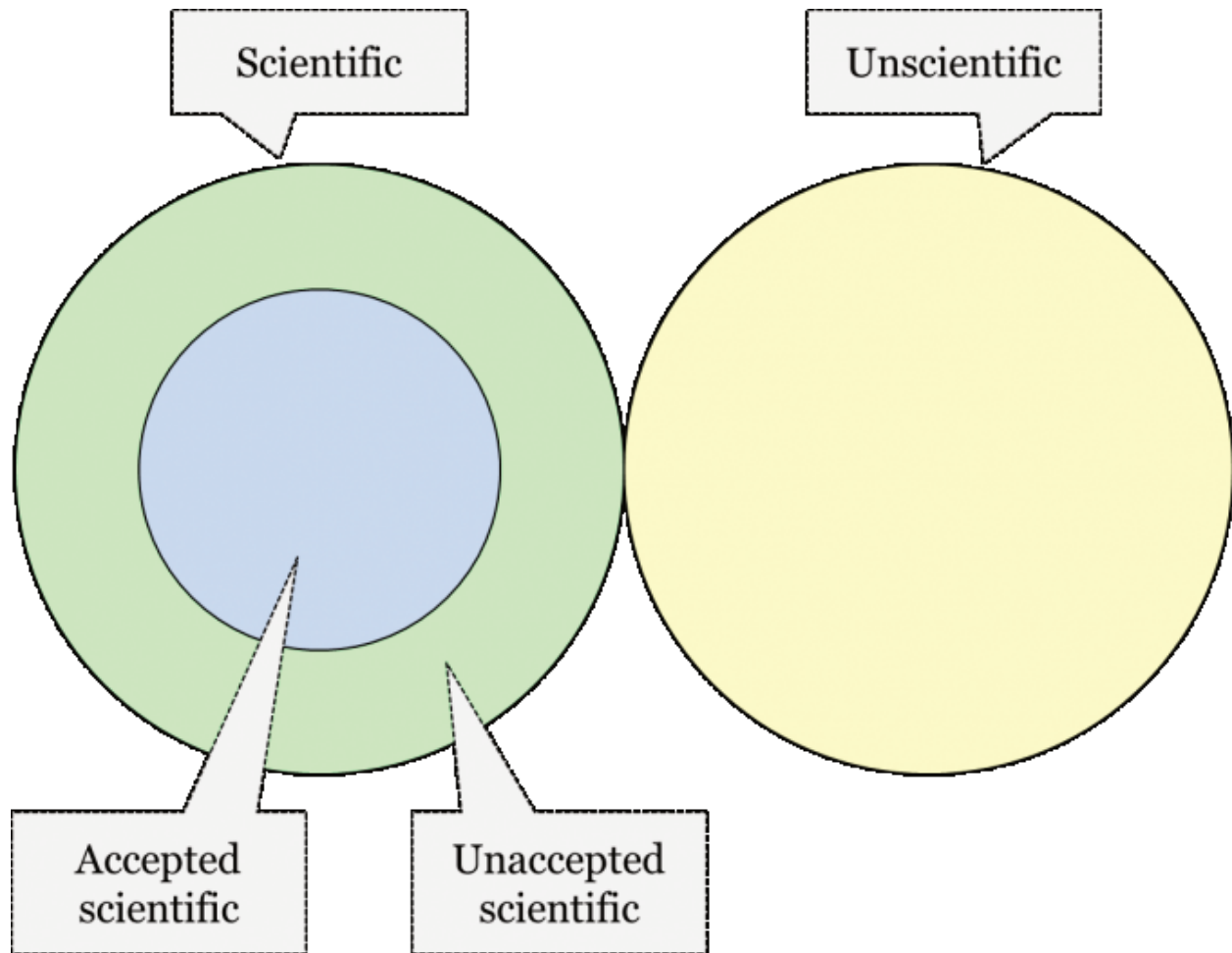
It should be clear by now that the same theory could be considered scientific in one mosaic and unscientific in a different mosaic depending on the respective demarcation criteria employed in the two mosaics. For instance, astrology satisfied the Aristotelian-Medieval demarcation criteria, as it clearly attempted to grasp the nature of celestial bodies by studying their effects on the terrestrial realm. It was therefore considered scientific. As we know, astrology is not currently considered scientific since it does not satisfy our current demarcation criteria. What this tells us is that demarcation criteria change through time.

Not only do they change through time, but they can also differ from one field of inquiry to another. For instance, while some fields seem to take the requirement of falsifiability seriously, there are other fields where the very notion of empirical falsification is problematic. This applies not only to formal sciences, such as logic and mathematics, but also to some fields of the social sciences and humanities.

In short, we have both theoretical and historical reasons to believe that there can be no universal demarcation criteria. Because demarcation criteria are part of the method of the time, theories are appraised by different scientific communities at different periods of history using different criteria, and it follows that their appraisals of whether theories are scientific or not may differ.

Scientific vs. Unscientific and Accepted vs. Unaccepted

Before we proceed, it is important to restate that demarcation criteria and acceptance criteria are not the same thing, as they play different roles. While demarcation criteria are employed to determine whether a theory is *scientific* or not, acceptance criteria are employed to determine whether a theory ought to be accepted as the *best available description of its object*. Importantly, therefore, it is possible for a community to consider a theory to be scientific and to nevertheless leave the theory unaccepted. Here is a Venn diagram illustrating the relations between the categories of *unscientific*, *scientific*, *accepted*, and *unaccepted*:



A few examples will help to clarify the distinction. General relativity is considered to be both scientific and accepted, because it passed the strong test of predicting the degree to which starlight would be bent by the gravitational field of the sun, and other subsequent tests. In contrast, *string theory* is considered by the contemporary scientific community as a *scientific* theory, but it is not yet *accepted* as the best available physical theory. Alchemy had a status similar to that of string theory in the Aristotelian-Medieval mosaic. The Aristotelian-Medieval community never accepted alchemy but considered it to be a legitimate science.

Changing the Question: From Theories to Changes

So, to recap, how do we tell if an empirical theory is a scientific theory or not? We do so by consulting the demarcation criteria of that particular community, at that particular time in history. If a theory satisfies the demarcation criteria employed by the community, it is considered scientific; if it doesn't satisfy the demarcation criteria, it is considered unscientific. What about pseudoscience? A theory is considered pseudoscientific if it is assessed to be unscientific by the scientific community's employed demarcation criteria but is nevertheless presented as though it were scientific. Just as with any other method, all of these demarcation criteria are changeable; they change in accord with the third law of scientific change.

Now, it is clear that if we look for transhistorical and universal criteria for demarcating *theories*, we most likely won't find any, since these criteria – just as any other method – change through time. Thus, if we are to draw a meaningful demarcation line between science and non-science, we need to change our focus from individual theories, to the process of scientific *change* itself.

As we mentioned earlier, most of the discussion of the demarcation problem in the philosophical literature has centred on *theories*, and we have just summarized some aspects of that grand debate. However, there have been some philosophers who suggested that the real distinction between science and non-science is to be located in the way science *modifies* its theories (and methods). Imre Lakatos, a Hungarian-born British philosopher of science, famously argued that it is the *transitions* from one scientific theory to another that can be qualified as scientific or unscientific, not individual theories themselves. According to Lakatos, what we evaluate as scientific are the *steps* in the process of scientific change, such as the transition from Descartes' theory to that of Newton in France ca. 1740, or from that of Newton to that of Einstein ca. 1920. This Lakatosian idea can be extended to apply also to the methods of theory evaluation. How exactly can this be accomplished?

Once we switch our focus from evaluating the scientific status of individual scientific theories to evaluating the transitions in a mosaic, we can pose a number of interesting questions: Can the acceptance of a new theory, or the employment of a new method, happen in a way that strikes us as *unscientific*? What are the features of scientific *changes* in a mosaic? What are examples of unscientific changes? In the most general form, our question becomes:

What are the characteristics of the *process* of scientific change?

So how can we evaluate whether a certain transition in a mosaic was indeed scientific? Let us appreciate that while the criteria of demarcation of scientific *theories* change through time, the *mechanism* underlying those changes is universal and transhistorical. As we learned in chapter 4, the laws of scientific change attempt to describe and explain that universal and transhistorical mechanism, to capture the general patterns exhibited by the process of scientific change. Thus, we can look at *any* episode of scientific change and, using the laws and our knowledge of the intricacies of the respective mosaic at that historical period, determine whether that change was scientific or not. How exactly can we do this?

Violations of the Laws of Scientific Change

We suggest that a modification of the mosaic is scientific if that modification takes place in accord with the laws of scientific change. It is unscientific, if it violates at least one of the laws. If, for instance, a new theory is accepted in accord with the second law, that change would be considered a scientific change. If, however, a theory's acceptance happened in violation of the second law, then chances are that the acceptance itself was an unscientific step. We suggest that an actual or potential change in the mosaic is considered unscientific if it violates any of the laws of scientific change.

But how can the laws of scientific change be violated? Indeed, the laws of scientific change are meant to be descriptions of general patterns at the organizational level of the scientific mosaic, in the same way that any chemical law is a description of general patterns that hold at the atomic and molecular levels. What, then, do we mean by a "violation" of the laws of scientific change?

When scientists talk about "the laws of nature," they are generally referring to the regular, law-like patterns that emerge at any specific level of organization, be it the physical level, the chemical, the biological, the psychological, or the social. As you might recall, we made the claim that the laws of scientific change describe the general, law-like patterns that emerge at the level of the scientific mosaic, one level "above" that of the social. The scientific community generally accepts that at the most basic, fundamental level (i.e. that of physical processes) the laws of nature are *inviolable*. That is, all fundamental physical processes always obey, without exception, the same set of physical laws, and nothing can ever prevent them from doing so. However, for any "higher" level law – such as those of biology, psychology, economics, or sociology – there's always a chance that they can be violated, because the regularities they describe can have exceptions, usually due to the interference of another (usually "lower") level process. Laws at levels above the fundamental physical level are therefore considered *local* and are said to therefore hold only in very specific conditions.

Consider an example. The dinosaurs dominated the global biosphere for hundreds of millions of years, all the while slowly adapting, developing, and evolving according to the general mechanism of evolution... that is, until one large rock (an asteroid) collided with another, larger rock (planet Earth) about 65 million years ago. Suddenly, the very conditions for dinosaur evolution were wiped away. We could say that the slow gradual process of evolutionary change through natural selection was interrupted and replaced by sudden and extraordinary catastrophic change. Alternatively, we could say that the "laws" of evolution were "violated". This violation only demonstrates that the general patterns of biological evolution emerge and hold only under very specific physical and chemical conditions. This makes the laws of biology as well as any other "higher" level laws *local*.

A similar hypothetical example can be used to illustrate what a violation of *the first law* of scientific change, *scientific inertia*, might consist of. As we know, the law states that any element in the mosaic remains in the mosaic unless it is replaced by other elements. Let's imagine that, sometime before the creation of the Internet, every palaeontologist in the world had gathered together at DinoCon to discuss the extinction of the dinosaurs. In their enthusiasm, this scientific community brought *all* of their books and *all* of their data with them. Then, in an ironic twist of fate, a much smaller asteroid struck the DinoCon convention centre, wiping all of the world's palaeontologists, and all their data and books, off the face of the Earth. What would happen to their theories? It would seem that paleontological theories would simply be erased from the mosaic and would be replaced by nothing at all. But this removal of paleontological theories is a clear violation of *the first law* of scientific change. This sudden loss of theories would intuitively strike us as an unscientific change in the mosaic even if we didn't know anything about the first law. However, with the first law at hand, we can now appreciate exactly *why* it would strike us as an unscientific change.

How might the other laws of scientific change be violated? Let us consider some hypothetical scenarios illustrating potential violations of the remaining three laws.

The *second law* can be violated if a new theory is accepted *without* satisfying the acceptance criteria employed at the time. Let's consider a theory like that of *sound healing*, which proposes that the sound vibrations of certain percussive instruments, like gongs and bells, can bring health benefits by rearranging the ions of one's cell membranes. Today, these claims would have to be assessed by our current acceptance criterion, namely the HD method, and would thus require a confirmed novel prediction, since the relation of sound vibrations to cell ion arrangement is a new causal relation. Suppose that a brutal dictator rose to power, who had a penchant for torturing and executing anyone who disagreed with his beliefs, and that among his most treasured beliefs was a belief in the efficacy of sound healing. If the scientific community were to suddenly accept sound healing with no confirmed novel

predictions whatsoever, due to such an influence, we would say that this change had violated the second law. Such a transition – one that did not satisfy the requirements of the employed method due to the action of an unscientific influence, such as fear of the dictator – would clearly strike us as an unscientific step.

The third law can be violated by employing a method that simply does not deductively follow from our other accepted theories and employed methods. As we've seen in chapter 4, the method currently employed in drug testing is the so-called double-blind trial method. The method is currently employed, since it is a deductive consequence of our accepted theories of unaccounted effects, the placebo effect, and experimenter's bias (as well as a number of other relevant theories). Scientists employ this method on a daily basis to determine whether a certain drug is therapeutically effective. Now, imagine a hypothetical situation in which a biomedical community suddenly begin relying solely on *patient testimonials* instead of performing double-blind trials. Also imagine that they decided to rely on patient testimonials while still accepting the existence of the placebo effect. Such a drastic transition to a new employed method strikes us as an unscientific step and rightly so: if scientists were to employ a new method which is not a deductive consequence of their accepted theories, it would violate the third law of scientific change and would, thus, be an unscientific step.

Finally, *the zeroth law* can be violated if a pair of theories doesn't satisfy the compatibility criteria of the mosaic and yet manages to simultaneously persist within a mosaic. It's important to realize that, whenever a new theory is accepted, the conjunction of the zeroth and first laws dictates that any incompatible theories be immediately rejected. For example, suppose a community accepts the theory "drug X is the single best drug for treating headaches". If the community accepts a new theory the next day "drug Y is the single best drug for treating headaches," the drug X theory will be immediately rejected, thanks to the zeroth law. But if two incompatible theories were simultaneously accepted *without bringing about any mosaic splits*, then the zeroth law would clearly be violated. Say a small community accepts the following theory "Zhi's absolute favourite colour is blue", and their *compatibility criterion* says "the same person can't have two absolute favourite colours". Then, a day later – in accord with the second law – they accept "Zhi's absolute favourite colour is yellow" but, for whatever reason, they fail to reject "Zhi's absolute favourite colour is blue". This hypothetical scenario – however absurd – would be a violation of the zeroth law, since it's a persistent incompatibility between two theories in the same mosaic. We would, of course, require an explanation, invoking non-scientific influences to explain the violation of the law.

Hopefully even without knowing the laws of scientific change many of these examples would have seemed unscientific, or at least odd. But knowing the laws of scientific change, and how they might be violated, helps us to articulate exactly what about each of these changes makes them unscientific. In short, *a change* in a mosaic strikes us as *unscientific* if it violates one or more of the laws of scientific change. A change in the mosaic can be considered *pseudoscientific* if, while violating at least one of the laws, it is also presented as though it has followed them all.

Summary

There are two distinct questions of demarcation. While discussions concerning the demarcation of science and non-science have typically centred on figuring out what makes *theories* scientific, we have suggested that the most fruitful way forward is to switch the discussion from theories to *changes* in a mosaic. As such, the problem of demarcation can be specified by two distinct questions. The first question concerns scientific *theories*:

What are the characteristics of a scientific theory?

As we have explained, this question doesn't have a universal and transhistorical answer as the criteria of demarcation are changeable. The question that we suggest is this:

What are the characteristics of the process of scientific change?

We have argued that in order to determine whether a certain change was or wasn't scientific, we should check to see if it violated the laws of scientific change.

Now that we have outlined some of the central problems in the philosophy of science, it is time to proceed to the history of science and consider snapshots of some of the major scientific worldviews.

7.

Aristotelian-Medieval Worldview -- Introduction to History and Philosophy of Science

Chapter 7: Aristotelian-Medieval Worldview

Intro

So far, we've been studying some of the central philosophical questions about science: Can we know anything with absolute certainty? Is there a universal and unchangeable method of theory evaluation? What is the mechanism of scientific change? Is there scientific progress? What is the difference between science and pseudoscience? Now we are moving to the history of science, and we will focus on some examples of major scientific worldviews through history.

But why would anyone bother studying the scientific worldviews of the past? After all, from the perspective of contemporary science, most elements of past mosaics would be simply unacceptable. So, what's the point of studying theories and methods that were rejected such a long time ago? There are at least two reasons to do this – one historical and one philosophical.

The *historical* reason for studying the worldviews of the past is that it helps us better understand the events of the past. It is appreciated nowadays that one cannot properly make sense of past events unless they are considered in their proper historical *contexts*. Any historical context includes not only a certain social organization, economic and political structure, but also – and importantly – a certain set of accepted beliefs about the world. We've already seen in chapter 4 that it is impossible to make sense of the case of Galileo without knowing the accepted theories and employed methods of the Aristotelian-Medieval mosaic of his time. This can be generalized: understanding a certain historical transition requires knowledge of the respective historical context, including the mosaic or mosaics of the time.

In addition, there is an important *philosophical* reason for studying worldviews of the past: it helps us improve our picture of science by better understanding the process of scientific change. Imagine what sort of picture of science would emerge if we were to ignore pre-eighteenth-century science. We do not need to guess. It was customary until the 1960s or even 1970s to dismiss the whole of pre-Newtonian/pre-Galilean science as not really belonging to the history of science: hundreds of years of Aristotelian-Medieval science were cast aside! As a result, most notable philosophers of that time ended up believing that there is one fixed and universal method of science. While we no longer accept the idea of a fixed and universal method of science, even nowadays some philosophers nonetheless still believe that there is one universal method of science. This is usually an unfortunate outcome of their decision to disregard the whole history of knowledge before the 17th century. If we dismiss the science of the past simply because it employed a different method of theory evaluation, then we have to be prepared for our science to be dismissed once our current methods, such as the HD method, gets replaced by some new method. A more sensible option is to acknowledge the importance of past mosaics and study them with utmost care.

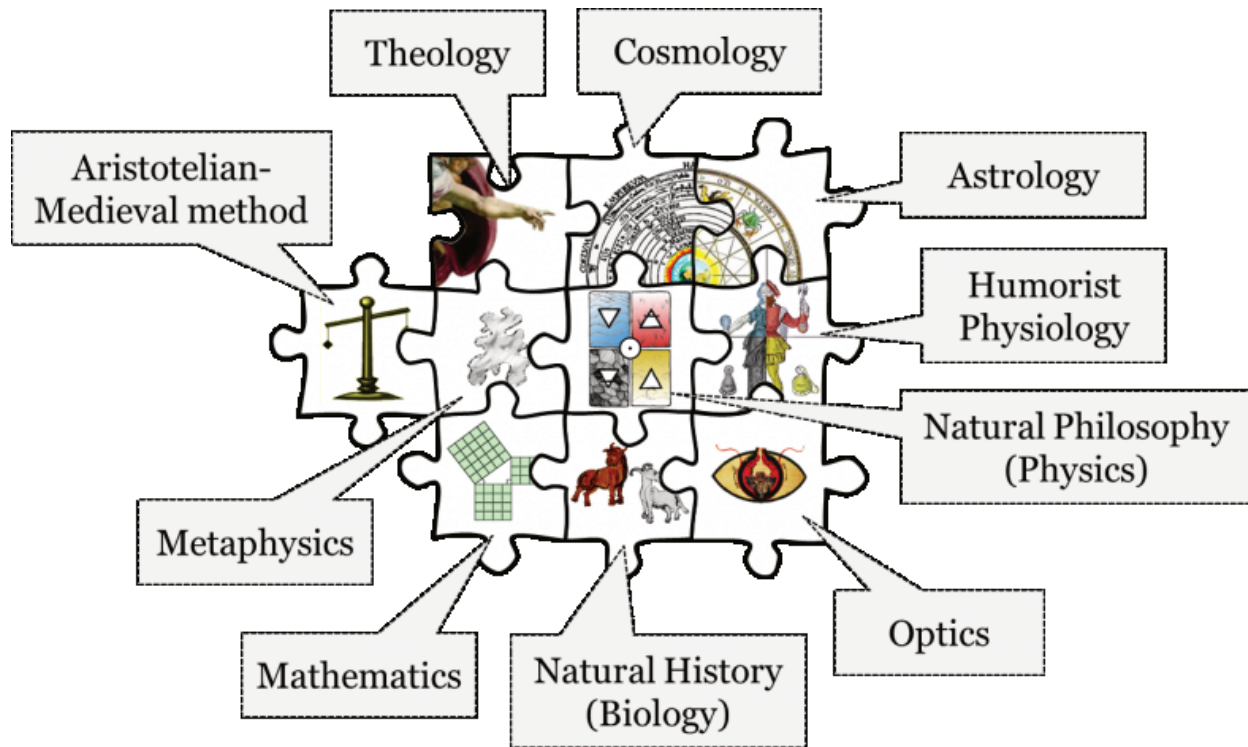
A few important clarifications are in order before we proceed. Firstly, any worldview is a complex entity with thousands of different accepted theories and a wide range of methods for evaluating theories. Thus, a comprehensive exposition of a single worldview can take volumes. In our discussion, however, we focus on a small subset of accepted theories and employed methods to give a cursory snapshot of some of the key elements of a worldview.

Secondly, throughout history, there have been many distinct epistemic communities with distinct worldviews; of these, we only focus on four worldviews which constitute a historical *sequence* – Aristotelian-Medieval, Cartesian, Newtonian, and Contemporary. This selection accomplishes two goals: it gives a snapshot of some of the most influential scientific worldviews of all time, and it helps illustrate how theories, methods, and whole worldviews change through time.

Thirdly, scientific mosaics are not static entities. They constantly evolve; theories and even methods change all the time. Thus, no matter what year and what epistemic community we choose, we inevitably end up with a snapshot of that community's mosaic at that particular time. There is much more to the history of science than these four worldviews.

Fourthly, our reconstructions of these four worldviews are done from a bird's-eye perspective: they skip any mention of specific epistemic communities that bore these mosaics. While this is acceptable for an introductory textbook, a proper historical study should start from identifying the epistemic communities of the time under scrutiny and then proceed to extracting the content of their mosaics.

In this chapter we will consider the first of our four worldviews – the Aristotelian-Medieval mosaic. What were some of the key elements of the Aristotelian-Medieval mosaic? As with any other mosaic, the Aristotelian-Medieval mosaic included a vast array of different theories and methods. Among these are the Aristotelian natural philosophy (physics) of one celestial and four terrestrial elements, humoral physiology and medicine, the geocentric cosmology of concentric spherical shells, natural history (biology), and mathematics, as well as the Aristotelian-Medieval method of intuition schooled by experience. In addition to these elements, we also find a number of surprising inclusions such as astrology, the study of celestial influence upon terrestrial events, and theology, the study of God. Both astrology and theology were considered legitimate topics of scientific inquiry and their theories were part of the mosaic of the time. In fact, different medieval communities often accepted different theologies, such as those of Christianity (catholic, orthodox, protestant, etc.), Islam, or Judaism. In any event, a theology was virtually always an indispensable component of any Aristotelian-medieval mosaic.

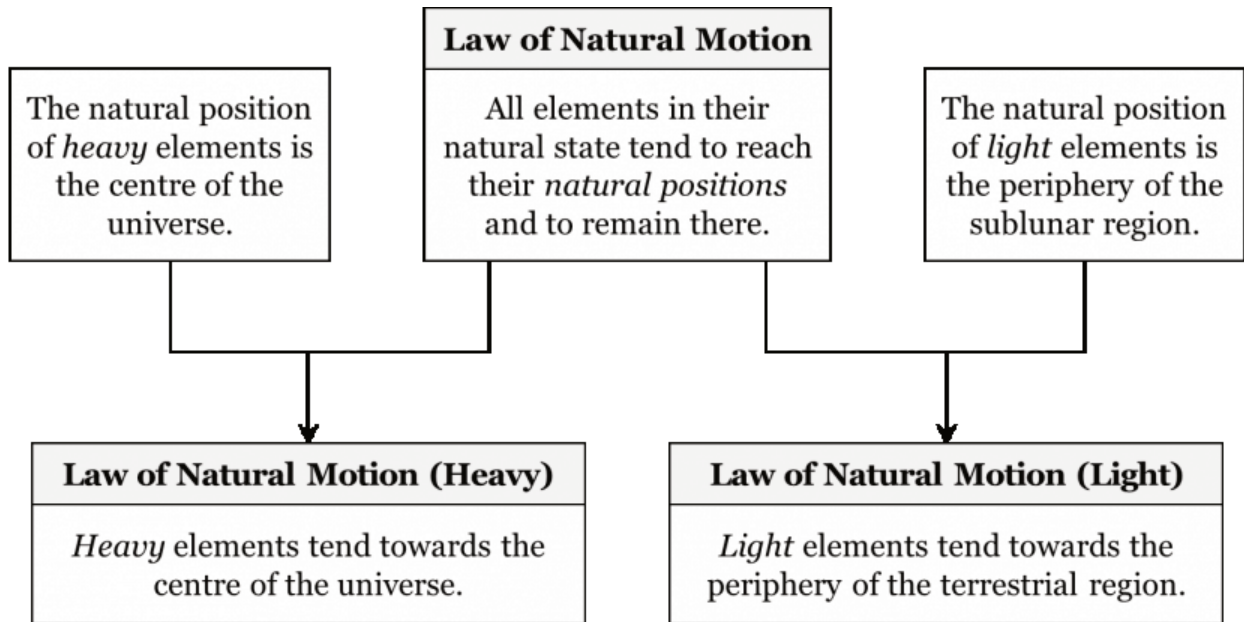


While there were many different varieties of the Aristotelian-Medieval mosaic, all of them shared a number of key elements. In one form or another, these central elements remained accepted by different epistemic communities until ca. 1700. It is these elements that we are going to focus on in this chapter.

Aristotelian Physics / Natural Philosophy

One of the key elements of this mosaic was Aristotelian physics, or natural philosophy as it was also often referred to at that time. Implicit to this physics was the idea that the universe consists of two distinct regions – terrestrial (or sublunar) and celestial. Everything within the terrestrial region was believed to be made out of the four terrestrial elements – *earth*, *water*, *air*, and *fire*. As we already know, two of these elements – earth and water – were thought to be *heavy*, while the other two elements – air and fire – were thought to be *light*. It was also believed that earth is heavier than water, while fire is lighter than air.

All of these elements were believed to have a natural position in the universe and each element was thought to have a natural tendency to get to that natural position and remain there. The natural position of the heavy elements earth and water was thought to be the centre of the universe and, thus, they were thought to have a natural tendency to descend towards the centre of the universe. As for the light elements air and fire, it was believed that their natural position is the periphery of the sublunar (terrestrial) region. Therefore, they were thought to have a natural tendency to ascend along the radius of the universe towards the periphery of the terrestrial region. The basic Aristotelian laws of natural motion can be formulated as follows:

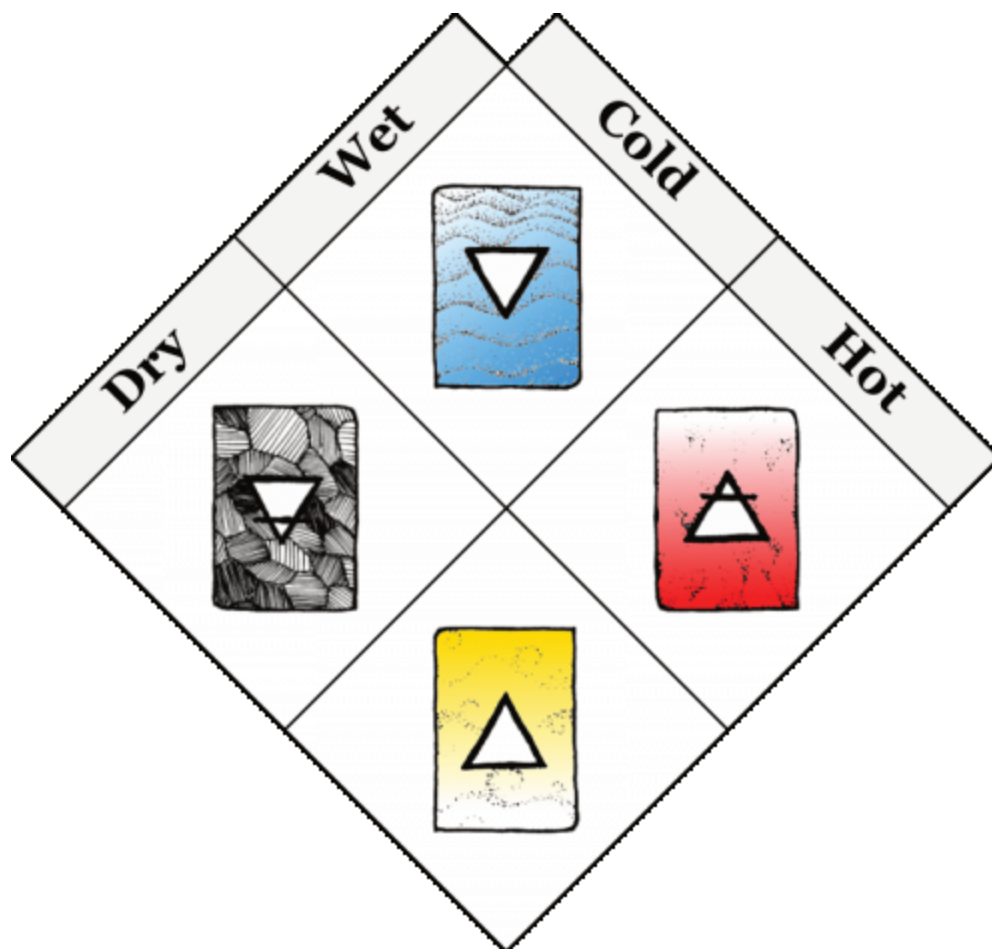


These seemingly basic laws allowed Aristotelians to explain a wide array of terrestrial phenomena. Why does fire penetrate through air while ascending? Because fire has a natural tendency to ascend toward the periphery of the sublunar region and because fire is lighter than air, through which it penetrates. Why does a stone sink in water? It is because the stone is made predominantly out of the element earth and because earth is heavier than water.

One straightforward consequence of this physics was the idea that earth, water, air, and fire should be arranged in that very same order starting from the centre of the universe all the way to the edge of the terrestrial region. As the heaviest element, earth should naturally collect at the centre of the universe, with a layer of water on top of it. After that comes the layer of air followed by a layer of fire up in the skies; a layer that we don't usually see but know is there, since that's where all the burning fire ascends. Thus, the idea of geocentrism is implicit in this theory: the Earth, which is a combination of the elements earth and water, must necessarily be located at the centre of the universe. Importantly, geocentrism wasn't the result of a random choice, but was dictated by the accepted physics of the time.

Now, why would they *accept* the existence of precisely those four terrestrial elements? As we already know from previous chapters, Aristotelian-Medieval science was a science of *common sense*; a theory was thought to be acceptable if it was grasped *intuitively* by an expert. But what is so *intuitive* about the idea of four elements? If we look around us, we see that some things are solid, others are liquid, and yet others are gaseous. We also notice that there is fire which is different from the previous three. Thus, it seems intuitive to suggest that all things are made of earth (solid), water (liquid), air (gas) and fire. So, it is perhaps for this reason that many different epistemic communities in different geographic regions including Japan, India, and Egypt, arrived at a similar idea. While we no longer accept the idea of four elements, nowadays we do accept that *solid*, *liquid*, *gas*, and *plasma* are four states of matter. We can easily see why the idea of four elements appeared intuitively true.

It was also accepted in the Aristotelian-Medieval worldview that the four terrestrial elements are characterized by a pair of sensible qualities of cold/hot and wet/dry. Thus, fire is hot and dry, air is hot and wet, water is cold and wet, and earth is cold and dry.



Because the elements were believed to bear a specific pair of these sensible qualities and because it was known that these qualities can transform into one another, it was also accepted that the elements themselves can *transform* into one another. Such a transformation would require changing at least one of the underlying sensible qualities. For instance, since water is cold and wet, if we were to turn its quality of cold into hot, we would thus transform water into vapor, i.e. air. According to Aristotelian physics, that is exactly what happens during the heating of water. The idea of transformability also explains the process of combustion of wood. Since wood falls down when dropped, we can conclude that it is made out predominantly of *heavy* elements – earth and water. Both of these elements have the sensible quality of cold. In combustion, this quality of cold is turned into hot, which results in the earthly parts being transformed into fire (hot and dry), and watery parts into air (hot and wet).

This idea of the transformability of the elements legitimized the study of these transformations, *alchemy*. Since the four elements transform into one another, it only makes sense to study how exactly they do so. While alchemy was concerned with transformations of all kinds, it was specifically interested in obtaining the so-called *philosopher's stone* and *the elixir of life* – two elusive yet potentially possible substances. The philosopher's stone was believed to be an alchemical substance – a stone or powder – that could help transform any base metal into gold or silver. As for the elixir of life, it was thought to be a potion that would restore the perfect balance of bodily fluids and, potentially, grant eternal life. The possibility of both the philosopher's stone and the elixir of life was based on the idea that metals as well as bodily fluids, by way of being composed of different combinations of terrestrial elements, could transform to other metals and fluids. Needless to say, the alchemists failed to produce either of these substances. It is still very much an open question whether any alchemical theory was actually an *accepted* part of the Aristotelian-Medieval mosaic, or whether alchemical theories were merely *pursued*. What is well known is that the very possibility of alchemy was beyond question.

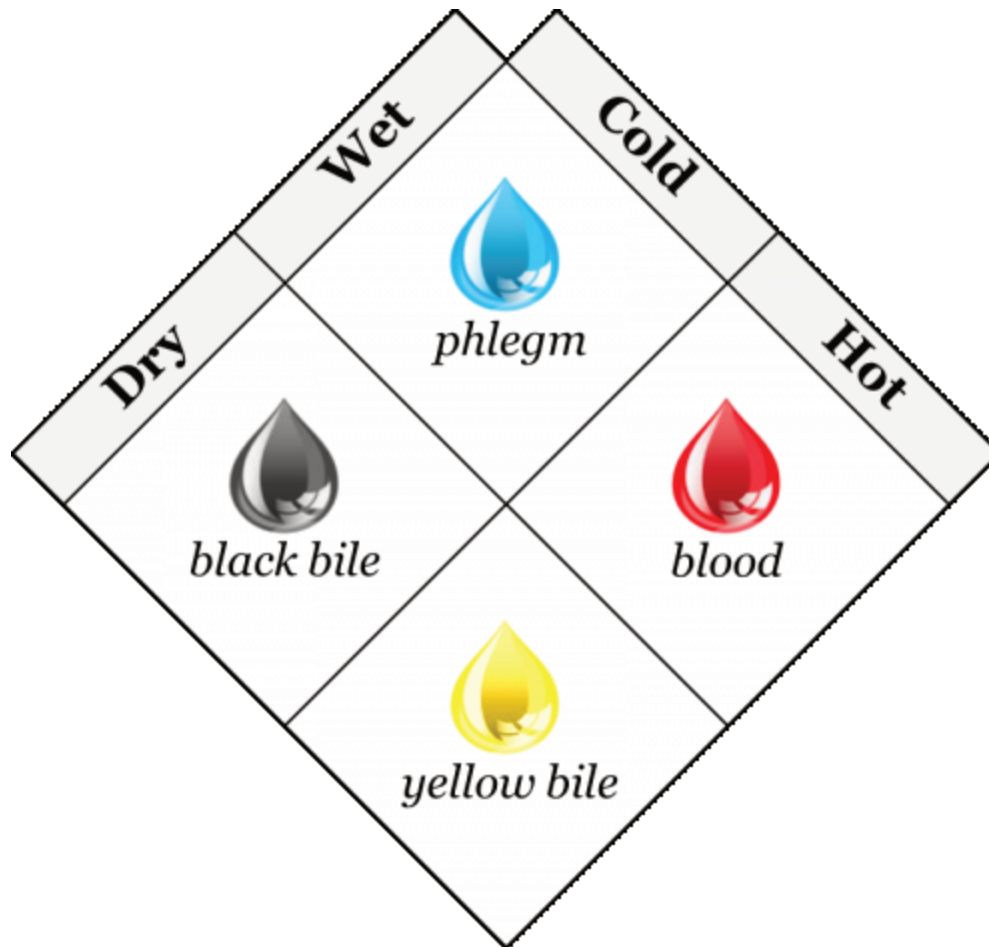
Humoral Physiology and Medicine

In the absence of any miraculous elixirs, medieval physicians had to rely on more down-to-earth remedies to restore the balance of bodily fluids. But what is this *balance*, and why was it so important to restore it? The answer has to do with the humoral physiology and medicine that was developed by, among others, Hippocrates and Galen and was an essential part of the Aristotelian-Medieval mosaic.

According to humoral theory, the human body contains four bodily fluids, the so-called *humors*: *blood*, *yellow bile*, *black bile*, and *phlegm*. Why were there exactly four humors? The reason is that everything in the terrestrial region, including the humors in the human body, was believed to consist of different combinations of the four elements. Thus, one crucial difference between any two humors should be those specific combinations of elements which make up these humors. Naturally, the four elements can combine differently and, therefore, different humors will have a different predominant element. In blood, the predominant element is air. The predominant element in yellow bile is fire. In black bile, the predominant element is earth. Finally, the predominant element of phlegm is water.

Humor	Predominant Element	Qualities
yellow bile	fire	hot & dry
blood	air	hot & wet
phlegm	water	cold & wet
black bile	earth	cold & dry

By way of having a predominant element, each of these four humors was also characterized by a pair of sensible qualities of hot/cold and dry/wet. Thus, yellow bile was believed to be hot and dry, whereas blood was believed to be hot and wet. Phlegm was characterized as cold and wet, and black bile as cold and dry.



Importantly, as the constituents of the human body, these humors were thought to be responsible for health and disease. Health was understood as the state of the perfect balance between the four constituent humors; when the humors are in the correct proportion, then the body is healthy. Thus, a disease occurs when this natural balance is impaired by either an excess or a deficiency of a certain humor. This imbalance could be due to an array of different factors, such as *unhealthy regimen* (lack of sleep, exercising, sex, bathing, etc.), *environmental factors* (bad weather, corrupted air, etc.), or *unhealthy diet*.

But if a disease is a state of imbalance of humors, then curing amounts to restoring the body's natural balance of humors. It was accepted that the human body has a natural healing ability, i.e. an ability to generate humors in the body and bring them to the state of natural balance. Thus, to cure a patient, a medieval physician was supposed to help the body restore the balance of humors. There was a wide range of means of curing available to the physician. For one, different food was believed to affect the balance of humors in a certain way. Depending on the nature of a disease, a physician could recommend a specific diet. For example, if a deficiency of blood was diagnosed, a physician could advise eating green vegetables. Similarly, moderate consumption of wine could be prescribed for normalizing black bile, just as a diet of fresh fruits could be advised in case of the deficiency of phlegm. In addition, many natural substances were thought to have therapeutic qualities and were often used as drugs for restoring humoral balance. Opium, for instance, was sometimes prescribed for reducing an excess of blood. It was also accepted that the human body has a natural healing ability, so a proper regimen was often thought to be the best treatment. Exercising, sleep, and bathing were believed to have a potency to increase or decrease respective humors. Moderate exercise, for instance, was believed to be conducive to building blood, while cold baths were believed to help increase phlegm. In extreme cases, if a specific humor had to be reduced immediately, a physician could opt for cleansing the body through a variety of body-purging therapies, such as laxatives, diuretics, expectorants (coughing), emetics (vomiting), enema, cautery (application of hot iron) as well as the all-time favourite – bloodletting.

In addition to natural causes of disease, it was accepted in the Aristotelian-Medieval mosaic that a disease can also be due to supernatural causes, i.e. it can be a divine punishment for a sin or a stimulus for spiritual growth. Similarly, in addition to a variety of natural therapies, curing could be achieved by supernatural means, such as God's direct intervention. Thus, praying for the patient's healing was considered an important addition to administering specific body-purging therapies or prescribing a specific diet and regimen.

It is important to note that the healthy natural balance of humors was believed to be different for different people. The natural balance of one person – the balance they were born with – could easily differ from the natural balance of another person. Thus,

different people would have individual balances where different humors were predominant. For some people, blood would be the predominant humor of their natural balance. For others, it would be yellow bile, or black bile, or phlegm. It was accepted that each individual is born with a specific balance of humors.

An individual's specific balance of humors was believed to determine their *temper*, or *temperament*. Four basic temperaments were believed to exist, each characterized by a predominant humor. *Choleric* temperament was a result of the predominance of yellow bile and was characterized as irritable, bold, and confrontational. In contrast, *sanguine* temperament was believed to be due to the predominance of blood in the person's balance; this temperament was described as optimistic, enthusiastic, and sociable. *Phlegmatic* temperament was a result of phlegm's predominance and was portrayed as compassionate, sensitive, and passive. Finally, the predominance of black bile was believed to result in *melancholic* temperament which was portrayed as practical, serious, and pessimistic.

Temperament	Predominant Humor	Description
Choleric	yellow bile	irritable, bold, confrontational
Sanguine	blood	optimistic, enthusiastic, sociable
Phlegmatic	phlegm	compassionate, sensitive, passive
Melancholic	black bile	practical, serious, pessimistic

Since different people were believed to be born with different balances of humors, i.e. with different temperaments, it was essential for a physician to be able to determine the patient's temperament in order to provide a proper diagnosis, administer efficient treatment, as well as recommend a specific diet. For example, since sanguines were believed to suffer from excess heat and moisture, they were also thought to be prone to overeating and overdrinking. Thus, they would typically be advised to stay away from fatty foods, sweets, as well as excessive drinking. In contrast, melancholics were believed to suffer from excess cold and dryness, resulting in coughs, sensitive skin, and arthritis in the joints. Therefore, melancholics would be advised to stay away from dry foods, consume enough healthy fats, and drink a sufficient amount of water. Similarly, choleric were thought to be predisposed towards excess heat and dryness that would result in headaches, overexertion, and injuries. So, a recommended regimen for a choleric would involve limiting fatty, salty, and hot foods, as well as taking physical activity in moderation. Finally, phlegmatics were believed to suffer from excess cold and moisture that would manifest in colds, allergies, and anemia. Consequently, a phlegmatic would be advised to consume foods that are conducive to heat and dryness, i.e. to regularly eat meat rather than exclusively vegetarian food.

But why are different people born with different temperaments? What determines which specific temperament a person is born with? Surely, this cannot be random and must depend on something. So, what would determine a person's temperament? The short answer is: their *natal horoscope*.

Cosmology and Astrology

To understand how exactly a person's natal horoscope was believed to determine their temperament, we have to take a step back and appreciate Aristotelian-Medieval views on the structure of the universe – their cosmology.

As we already know, it was accepted in the Aristotelian-Medieval mosaic that the universe has two distinct regions – terrestrial and celestial. While the terrestrial region was believed to consist of the four terrestrial elements of earth, water, air, and fire, the celestial region was believed to contain only one element – *aether*, also known as *quintessence* (Latin for “fifth element”). The celestial region was believed to be filled with aether, and all stars, planets, and spheres to which they are attached were believed to be made out of aether. Whereas the terrestrial elements had a natural tendency to either ascend to the periphery of the terrestrial region, or to descend towards the centre of the universe, it was believed that the natural tendency of aether is to revolve in circles around the centre of the universe.

Law of Natural Motion (Aether)

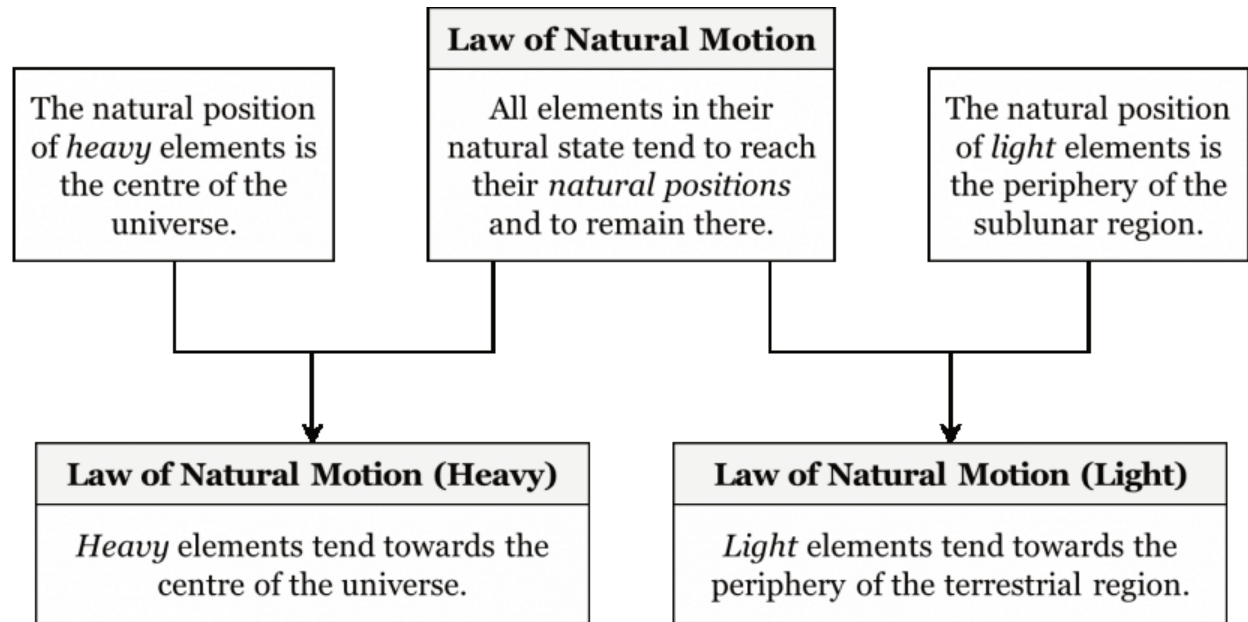
Celestial element *aether* moves in a uniform circular motion around the centre of the universe.

Since aether revolves around the centre of the universe, all planets and stars must also revolve around the centre of the universe, i.e. around the Earth. According to the then-accepted theory of concentric spheres developed by Eudoxus, Callippus, and Aristotle, each planet is nested in a solid crystalline orb – a spherical shell, which too is made out of aether and revolves around the Earth. The whole celestial region was believed to consist of these tightly nested concentric shells. The revolution of these spherical shells around the centre of the universe was believed to explain why all planets appear to be revolving around the Earth. There were believed to be seven planets – Moon, Mercury, Venus, Sun, Mars, Jupiter, and Saturn. In order to explain the apparent motion of planets around the Earth, the theory of concentric shells posited that each planet is carried by a number of nested spheres, each revolving slightly differently. The innermost planetary sphere would have the planet nested in it and would have its poles attached to a slightly larger sphere. This slightly larger sphere would have its poles attached to an even larger sphere and so on all the way to the innermost sphere of the next planet. The outermost sphere was the sphere of fixed stars, i.e. the sphere to which all stars were attached. Collectively, this system of nested concentric spheres would reproduce the apparent paths of the planets. This system is depicted in the following drawing:



Importantly, it was believed that the celestial element aether is immutable, in the sense that it cannot transform into any other element, it cannot be generated or corrupted (except by the Creator himself, of course), i.e. it doesn't come to be or cease to be. The only change in the celestial region had to do with the revolution of the celestial spheres; no new star, planet, or sphere was believed to be possible. This is in sharp contrast with the terrestrial region, where the four elements were believed to be mutable and where generation and corruption were the order of the day.

Now, what would happen if there were no external influences on the terrestrial region? What kind of an arrangement of elements could we reasonably expect if the laws of natural motion for heavy and light elements were not affected by anything else? If there were no external influences on the terrestrial region, then according to the laws of natural motion, the terrestrial elements would be arranged in ideal concentric spherical shells with the element earth being in the centre, followed by a layer of water, a layer of air, and a layer of fire. In such a hypothetical scenario, there would be no place for continents or islands, for all earth would have to be uniformly covered by water. Yet, we observe that not all earth is covered by water. In fact, we see that in the terrestrial region, the four elements are not arranged in perfect layers, but are mixed and combined to form all sorts of things. But how is this possible? It seemed reasonable to suspect that the terrestrial region was somehow influenced by something external to it, i.e. by the celestial region. In brief:



The conclusion that the celestial region affects the terrestrial region was in tune with several well-known phenomena. For one, it was known that the Sun exerts considerable influence on the Earth by being the main source of heat and light as well as the cause of the seasons. Similarly, the Moon – as the heavenly body closest to the Earth – was thought to cause a wide range of terrestrial phenomena. The Moon was thought to increase or decrease the flow of rivers, to cause the ebb and flow of the tides, and even to influence some biological processes in plants and animals. All these phenomena clearly suggested that terrestrial processes are somehow affected by processes in the heavens.

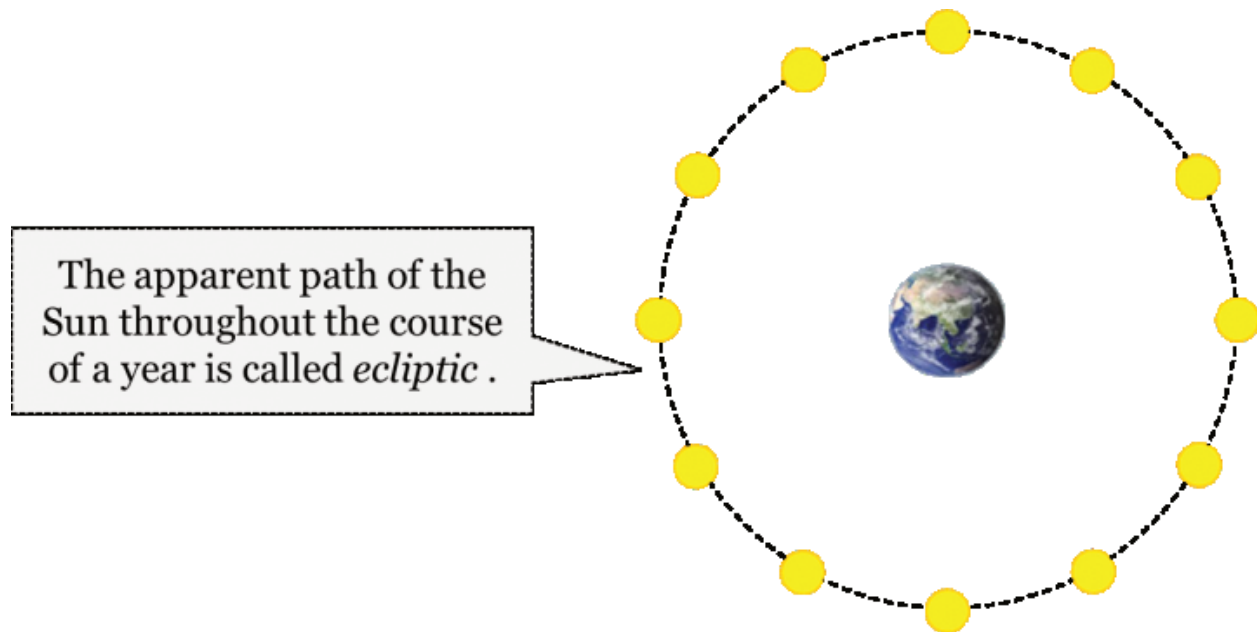
To explain *how* exactly the motion of the celestial spheres affects the terrestrial region, it was assumed that the motion of the outermost sphere of stars is due to God himself; the sphere of stars revolves eternally around the centre of the universe due to its love for God and as an attempt to mimic God’s eternity. The motion of the sphere of stars brings into motion the inner spheres; the motion is thus transferred to each subsequent sphere which eventually causes all sorts of change in the terrestrial region. While there was considerable disagreement as to what sort of celestial arrangements cause this or that type of event in the terrestrial region, the very existence of celestial influences was beyond doubt. There were believed to be numerous ways in which the motion of celestial bodies caused changes in the terrestrial regions (e.g. the Moon causing the tides, the Sun bringing heat and light). Therefore, it was reasonable to study these influences and find out exactly how changes in the celestial region bring about different terrestrial events. Enter the science of *astrology*.

Astrology

The presence of astrology in the Aristotelian-Medieval mosaic could initially strike one as surprising; after all, the very notion of studying horoscopes at a university sounds absurd to a 21st-century student. Yet, it should be clear by now that there was nothing strange in astrology’s inclusion into the curricula of medieval universities: the existence of celestial influences upon terrestrial phenomena was well known and, thus, there was nothing strange in devoting time and effort to studying these influences.

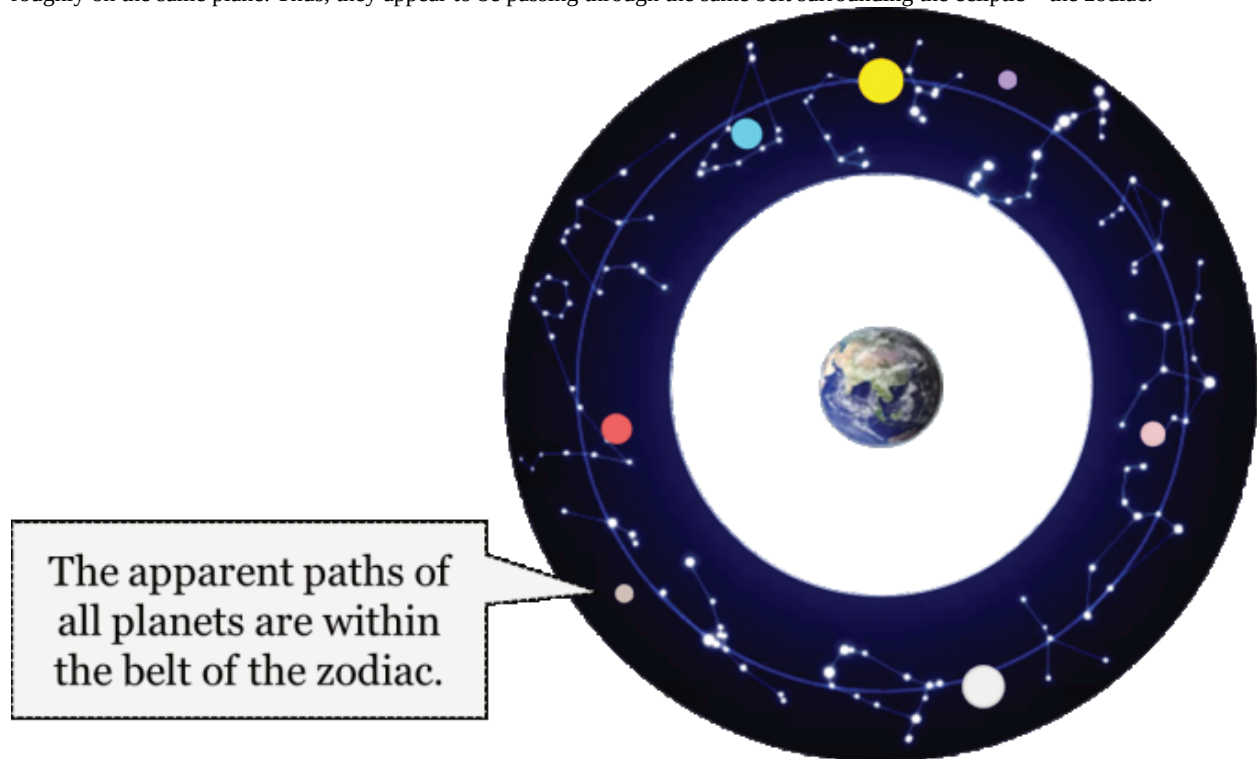
Note that in the context of the Aristotelian-Medieval worldview, the terms “astrology” and “astronomy” were sometimes used interchangeably: e.g. a treatise labelled as “astrological” could easily deal with astronomical subjects, and vice versa. However, to avoid confusion, we will differentiate the study of the motion of the planets (i.e. astronomy) from the study of the celestial influences on terrestrial affairs (i.e. astrology).

When an observer on the Earth studies the positions of the celestial bodies, it becomes clear that all of these bodies make a complete revolution around the Earth within the course of a day. Further observations reveal that while the relative positions of most of these bodies – the stars – remain the same, there are several curious celestial bodies – planets – which change their position relative to other celestial bodies. If we take the stars as our background, then we can notice that the positions of the planets gradually change relative to that starry background, i.e. the planets *wander* about the heavens. Carrying our observations further, we soon notice that the Sun completes precisely one full revolution per year and traces the same path across the heavens year after year. This apparent path of the Sun through the course of a year is called the *ecliptic*.



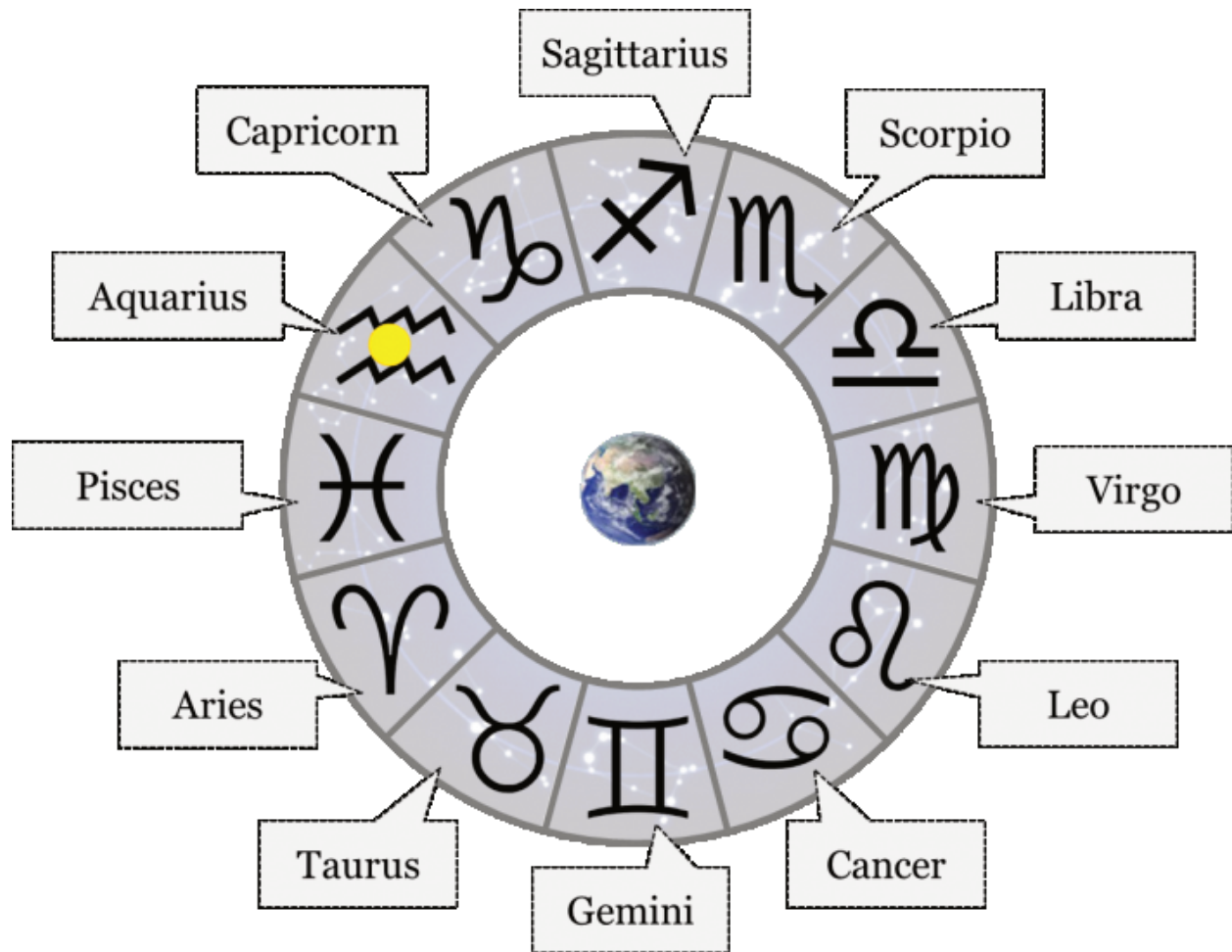
The apparent path of the Sun throughout the course of a year is called *ecliptic*.

The ecliptic is in the centre of a belt that we call the *zodiac*. Interestingly enough, the apparent paths of all the planets are positioned within this belt of the zodiac. In other words, the circles that the planets trace in the heavens are not random but are all roughly on the same plane. Thus, they appear to be passing through the same belt surrounding the ecliptic – the zodiac.



The apparent paths of all planets are within the belt of the zodiac.

Since the times of ancient Babylon, it has been customary to divide the ecliptic into 360 degrees – just as with any other circle. The main reason for this is that there are 365 days in a year, but 365 is not easily divisible by whole numbers and, thus, 360 – the closest number that can easily be divided by a great many whole numbers to yield a whole number result – was chosen instead. It also maps nicely onto 30-day synodic lunar cycles – the cycles between each meeting of the Moon and the Sun. There are twelve such cycles within each 360-day period, i.e. twelve months. This also explains why the belt of the zodiac was divided into twelve 30-degree divisions; each division contains one constellation, i.e. one astrological *sign*: Aries, Taurus, Gemini, Cancer, Leo, Virgo, Libra, Scorpio, Sagittarius, Capricorn, Aquarius, and Pisces.



Throughout the course of a year, the Sun passes through each of these twelve signs spending about 30 days in each of them. It is not surprising that an astrologer would adopt this twelve-month or twelve-sign division while searching for correlations between celestial and terrestrial phenomena. It was reasonable to suspect that each of these twelve signs has its own distinctive features and, therefore, its own specific influence upon terrestrial events. The only remaining task was to understand how these different signs could influence different types of events in the terrestrial region.

While astrology had many fathers, it was Claudius Ptolemy (c. 100–170 CE) who compiled astrological knowledge into a comprehensive system and provided rational explanations for many of its tenets. His astrological treatise, *Tetrabiblos*, was considered the utmost authority on astrological matters throughout the Aristotelian-Medieval period. Among the key tenets of astrology was the idea that each individual is born under one of the twelve signs. The individual's sign was believed to be determined by the position of the Sun at the time of their birth. For instance, an individual was said to be born under the sign of Aquarius if the Sun was in the constellation of Aquarius at the time of the person's birth. Each of the twelve signs was believed to have a specific effect on both the person's character and physique. But how could the planets exert this influence?

Since everything in the terrestrial region was believed to be made out of the four elements, it was reasonable to assume that celestial bodies would influence the Earth by means of increasing or decreasing the relative proportions of the four elements in people's bodies as well as in their environment. Each planet was believed to exert specific effects on the sublunary region: the Sun heats and dries, the Moon moistens, Saturn mainly cools, Jupiter somewhat heats and moistens, etc. It was also believed that the powers of the planets could be strengthened or weakened depending on their relative positions. Since each of the seven planets could find itself in each of the twelve signs, it was also believed that these signs can influence exactly how each planet affects terrestrial phenomena. Why is that? This is because each sign was expected to have a different effect on different elements; as a result, each sign was believed to be associated with one of the four elements. Aries, Leo, and Sagittarius were said to be the *fire signs*. Gemini, Libra, and Aquarius were the *air signs*. Cancer, Scorpio, and Pisces were the *water signs*. And finally, Taurus, Virgo, and Capricorn were the *earth signs*. Different outcomes were expected depending on the specific positions of different planets in each of the twelve signs.

Among many other things, the positions of the planets at the moment of the individual's birth were believed to shape the person's temperament. While all the planets were believed to exert a certain influence on a person's balance of humors, it was the position of the Sun at the moment of the person's birth that was believed to play the central role in determining the person's temperament.

The person's temperament essentially depended on the sign in which the Sun was found at their birth. Since each sign has an associated element, and thus an associated humor, one of the four temperaments was believed to obtain depending on the position of the Sun:

Signs	Element	Temperament	Humor
Aries, Leo, Sagittarius	fire	Choleric	yellow bile
Gemini, Libra, Aquarius	air	Sanguine	blood
Cancer, Scorpio, Pisces	water	Phlegmatic	phlegm
Taurus, Virgo, Capricorn	earth	Melancholic	black bile

For instance, a person born under the sign of Capricorn would have earth as her predominant element. As a result, black bile would be her predominant humor and she would therefore have a melancholic temperament. Similarly, a person born under the sign of Cancer would have water as the predominant element and, thus, phlegm as the predominant humor and would therefore be a phlegmatic.

While the location of the Sun was believed to determine the temperament of the person, the locations of other planets were thought to chisel the finer features of the person's character. For example, if at the moment of the person's birth the Moon was in the sign of Gemini, the person would be restless and be predisposed to traveling and learning new things. If, however, the Moon was positioned in the sign of Taurus, the person would be predisposed to a steady, patterned way of life. Other planets were also believed to have considerable influence on different aspects of the person's character. Studying these different effects was one of the main tasks of the science of astrology.

It should be clear by now why medieval physicians were supposed to know their astrology. Since a person's temperament, health, and disease were seen as being constantly affected by celestial influences both directly and through the environment, then it would be unreasonable to assume that medicine could be practiced without astrology. To begin with, it was important to know the patient's temperament to understand the severity of the illness. Similarly, in order to prescribe effective treatment, the medieval physician had to take into consideration the respective positions of planets. The planetary configurations would determine the dosage of drugs as well as the timing of their administration. It was also believed that surgical therapies, such as blood-letting or cautery, had to be conducted only at favourable times dictated by astrology. As a result, the knowledge of astrology, and therefore also of astronomy, was indispensable for a practicing medieval physician.

Aristotelian-Medieval Method

We already know a thing or two about the Aristotelian-Medieval criteria of *theory acceptance* from chapters 3 and 4. In particular, we know that the requirement of intuition schooled by experience was an essential part of the method employed for most of the history of knowledge in the West until ca. 1700. Yet, there was much more to the requirements of the medieval scholars than mere intuition schooled by experience. It is important to recall that any epistemic community can employ at least three different types of criteria. Employed *demarcation criteria* decide whether a theory is considered scientific or non-scientific. *Acceptance criteria* are employed to determine whether a theory is accepted into the mosaic. Finally, *compatibility criteria* are employed to decide whether a given pair of theories can be simultaneously accepted in the mosaic.

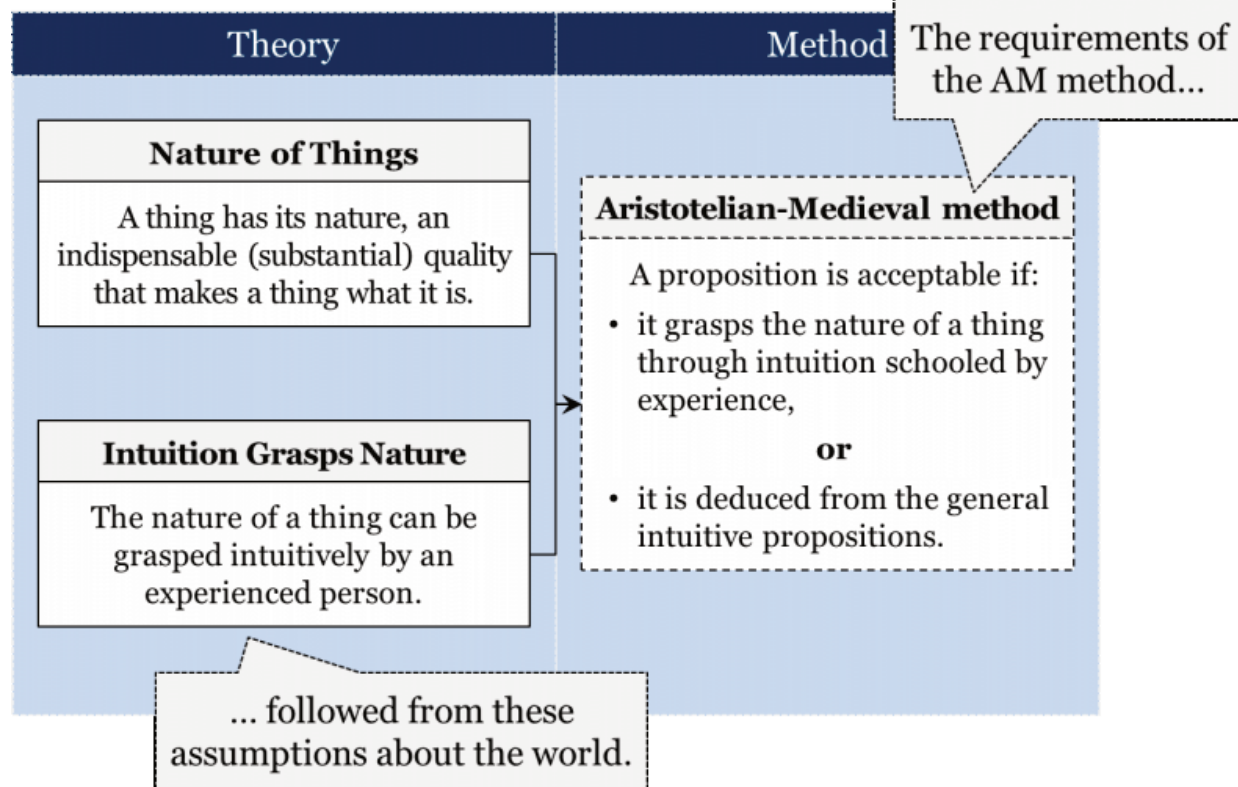
While further research may reveal other types of criteria, it is safe to say that, at the very minimum, any epistemic community employs these three types of criteria. The Aristotelian-Medieval community is no exception. Recall that the Aristotelian-Medieval criteria of *demarcation* between science and non-science that we discussed in chapter 6 were an integral part of their implicit expectations. It is also clear that they had to have certain expectations concerning *compatibility* or incompatibility of different theories. As for their *acceptance* criteria, it is true that the requirement of intuitive truth was an essential part of it. However, there were also other important requirements, which determined what was and wasn't allowed into the medieval mosaic. It is these additional requirements of theory acceptance that we will focus on in this section.

Let's start with our explication of the most general acceptance criteria of the Aristotelian-Medieval method:

Aristotelian-Medieval Method

A proposition is acceptable if it grasps the nature of a thing through intuition schooled by experience or is deduced from the general intuitive propositions.

As we already know from chapter 4, this method was based on two assumptions about the world. The first assumption was the idea that every natural thing has its *nature*, an indispensable quality that makes the thing what it is. For example, the indispensable quality of an acorn, its nature, is the capacity to grow into an oak tree, the nature of a lion cub is to become a full-grown lion, and the indispensable quality of humans is their capacity of reason. The second assumption was the idea that the nature of a thing can be grasped intuitively by an experienced person: the more time one spends observing the thing under study, the better one is positioned to grasp the nature of that thing. It is important not to confuse these *assumptions* with the *requirements* of the method; the requirements are those expectations that a theory should satisfy in order to become accepted, while the assumptions are those beliefs about the world from which these requirements deductively follow (by the third law). In the case of the Aristotelian-Medieval method, the requirements are the intuitive truth of a theory or its deducibility from what is intuitively true.



To illustrate the Aristotelian-Medieval requirements, consider an example from geometry:
 The sum of the interior angles of a triangle is equal to that of two right angles.
 This proposition is true, since it reveals the very nature of a triangle – that is what it is to be a triangle. Taking this proposition as our axiom, we can now deduce something more specific:
 The sum of the interior angles of an equilateral triangle is equal to that of two right angles.
 This theorem is true since any equilateral triangle is also a triangle.
 The idea that’s implicit here is that it is easier to notice those properties that are common to *all* objects of a class. Therefore, any inquiry was supposed to start from an identification of the most general first principles, *axioms*, from which the rest of the system – the *theorems* – could be deduced. For instance, since the idea of four terrestrial elements – two heavy and two light – was considered to be intuitively true, it was accepted as one of the axioms of Aristotelian-Medieval natural philosophy. From

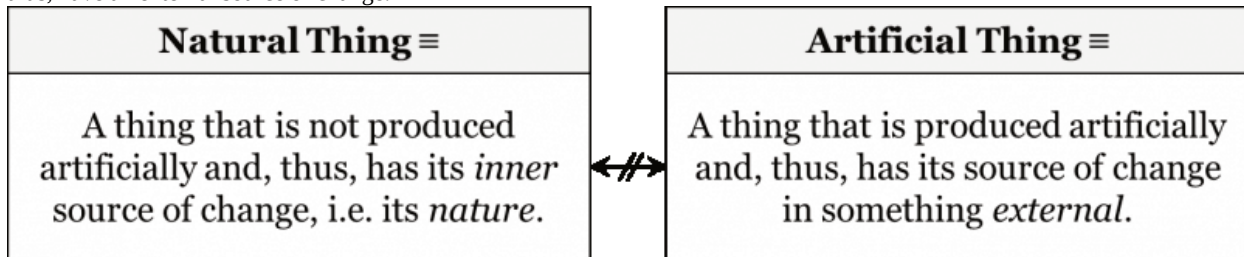
this axiom many theorems were deduced, including the idea that the Earth, a combination of the elements earth and water, should necessarily be in the centre of the universe. All of the sciences of the Aristotelian-Medieval mosaic would proceed in a similar fashion: they would attempt to grasp the most general, indispensable features of things under study and would then proceed to uncovering more specific features of things.

But what does it mean for something to be *intuitively* true? Yes, the axioms of our knowledge should be grasped as intuitively true, but surely not every intuition is an acceptable intuition. According to Aristotle and his followers, intuitions only count if they are schooled by years of experience in the respective field. Thus, if one wishes to know what the nature of bees is, one doesn't ask a random person, but only an *expert* – an apiarist, who has spent a lifetime observing bees in proximity. Similarly, if one's goal is to know the nature of human beings, one should, according to Aristotelians, refer to the intuitions of a philosopher who has spent years observing humans and reflecting on their traits. The key point here is that one must have a lifetime of experience with a thing under study to be properly positioned to grasp the nature of that thing; therefore, only an expert's intuitions count. Aristotelian-Medieval science was a science of *common sense*, but the common sense of an expert, not a layman.

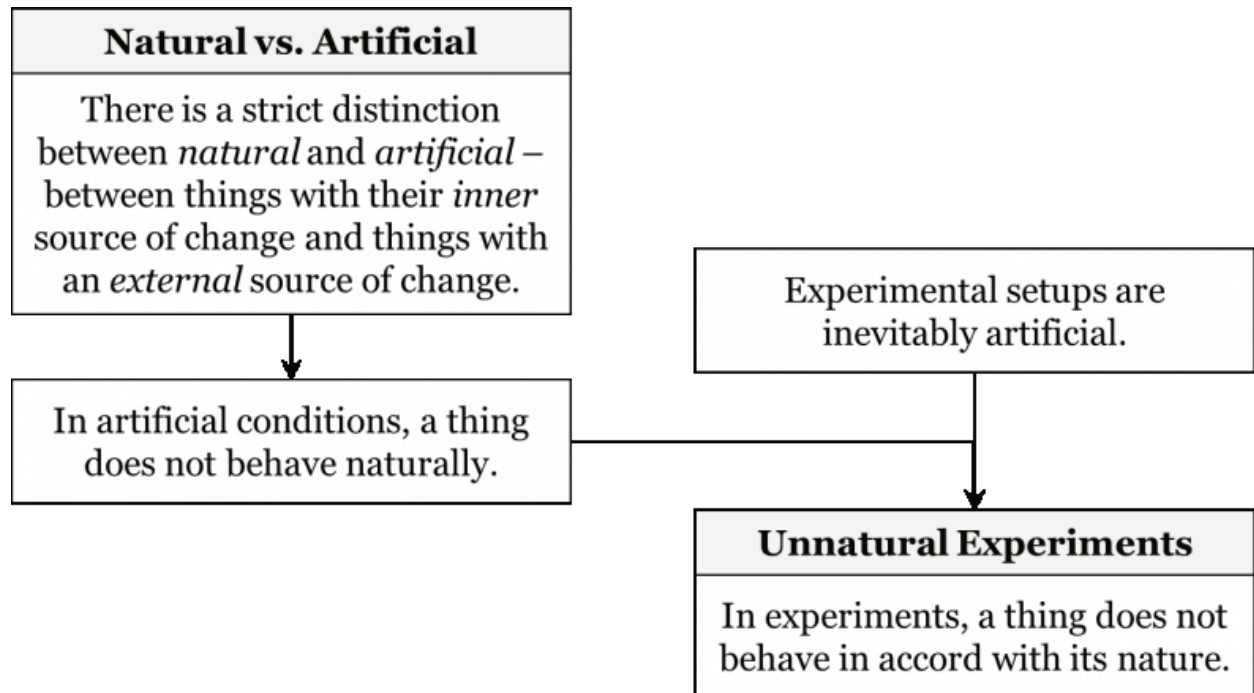
In addition to the requirement of intuitive truth, the Aristotelian-Medieval method of theory acceptance included a few other ingredients. One important ingredient of that method was the requirement that we should only rely on *observations* when evaluating a theory, i.e. that the results of *experiments* can be effectively ignored. Now, how could anyone dismiss the results of experiments and rely solely on the results of observations?

The answer has to do with the Aristotelian-Medieval distinction between *natural* and *artificial*. What is a natural thing? A thing is said to be *natural* if it is not produced artificially and, thus, has an *inner* source of change. A rock, for instance, is a heavy object and, by virtue of its heaviness, has an inner tendency to descend towards the centre of the universe. The same applies to all natural things. Animals reproduce because that's their inner tendency dictated by their very nature. Humans engage in studying the world around them because that's their inner tendency dictated by their nature – the capacity of reason. In all of these cases, the source of change is the very nature of the thing.

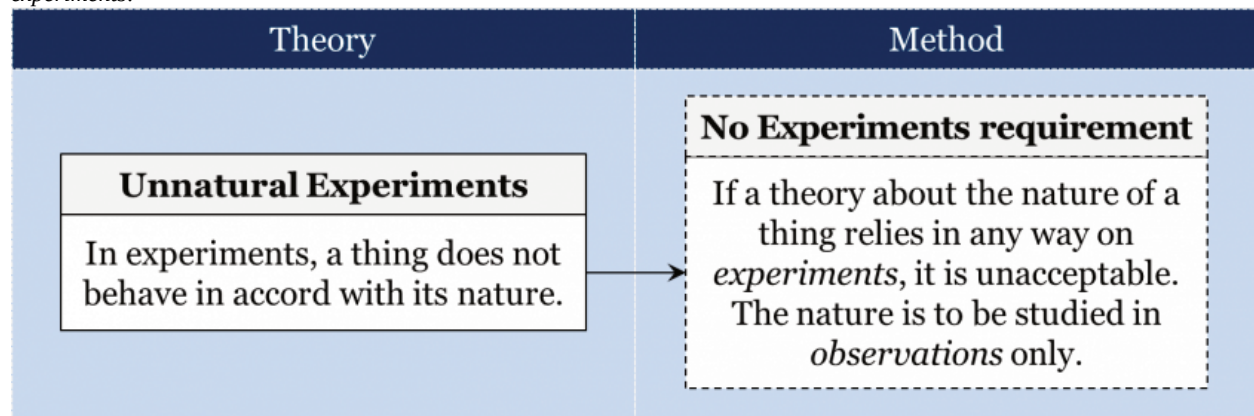
Compare this with something *artificial*, say, a compass. A compass is created by a craftsman to aid with navigation by showing the direction North. As such, it exists for the sake of something *external* to itself, i.e. for the sake of navigation. The same goes for a clock that is built to serve a purpose that is external to the clock itself – the purpose of timekeeping. It doesn't have an inner source of change, as it exists not for its own sake but as an aid for keeping time. According to the Aristotelian-Medieval way of thinking, this applies to all artificial things which, by definition, are the ones crafted by humans to perform some useful task and, thus, have an external source of change.



It was accepted in the Aristotelian-Medieval mosaic that there is a strict distinction between *natural* and *artificial* – between things with their *inner* source of change and things with an *external* source of change. Albeit foreign to us in the 21st century, this strict distinction between natural and artificial was accepted in the Aristotelian-Medieval mosaic and played an instrumental role in shaping their criteria of theory evaluation. Specifically, it follows from this distinction that in artificial conditions things cannot behave naturally; natural behaviour presupposes natural conditions. But any experiment, by definition, assumes a certain artificial setup; it is this artificial setup that differentiates experiments from observations. This explains why in the Aristotelian-Medieval mosaic experiments were not considered a trustworthy source of knowledge. It was believed that if we were to uncover the nature of a thing under study, subjecting it to experiments would be of no use. In other words, one had to study things in their natural unaltered condition, for experiments are artificial and do not reveal the nature of things. For example, if we are to unearth the nature of honeybees, we should not lock them inside a house where they have no access to pollen and nectar. That would put honeybees in an environment that is not natural to them and would therefore prevent us from seeing how they behave in their natural environment. As a result, we wouldn't be able to grasp their nature. The only right way of studying the bees is by observing them in their natural condition, when they behave in accord with their nature. Generally speaking, it was accepted at the time that any sort of experimentation is deficient because of its inevitable introduction of an artificial setup:

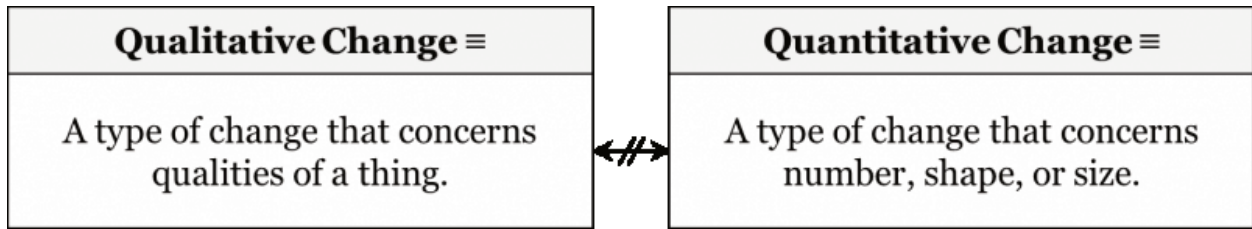


The *third law of scientific change* suggests that this belief had to shape their employed method and, sure enough, it did. The Aristotelian-Medieval community would not accept a theory about the nature of a thing if the latter relied in any way on *experiments*:



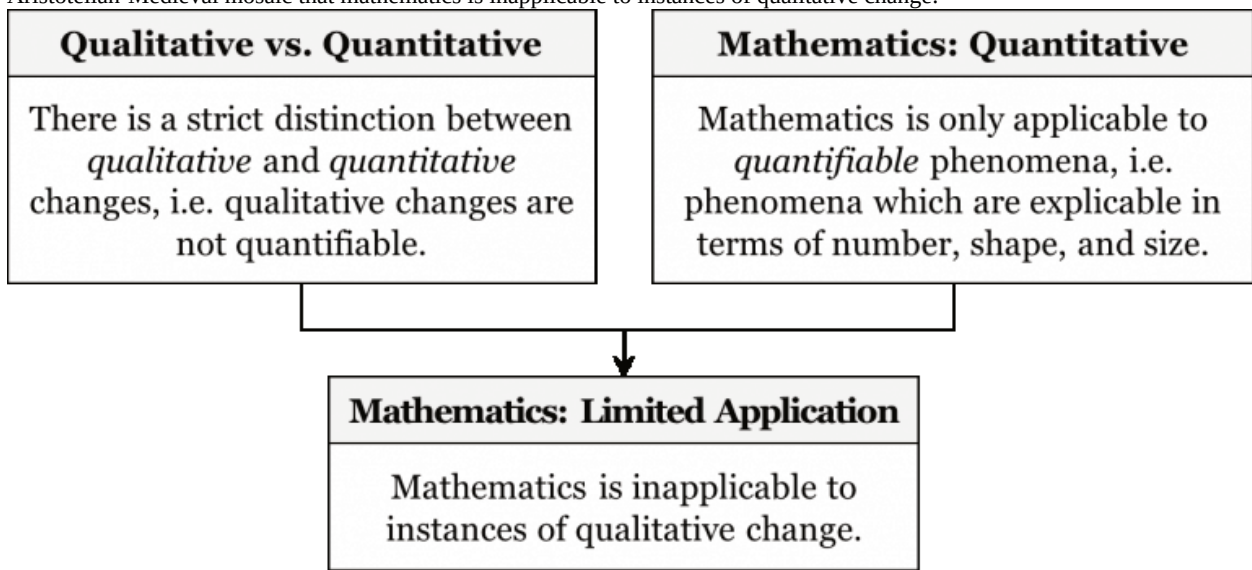
This should make it evident that the medieval dismissal of experiments as a reliable source of knowledge wasn't the result of a random whim but was based on the then-accepted strict distinction between natural and artificial. This restriction persisted until the end of the 17th century. It was removed with the acceptance of the Cartesian and Newtonian mosaics, which we will discuss in chapters 8 and 9.

Now, let's consider another important ingredient of the Aristotelian-Medieval method – their requirement concerning the limited applicability of mathematics. To appreciate this requirement, we should start from the then-accepted distinction between *quantitative* and *qualitative* changes. It was accepted that while some changes concern the number, shape, or size of things, other changes concern their qualities, and these changes cannot be expressed quantitatively. When a man grows in height or gains weight, these are instances of quantitative change, as they can be expressed in numbers. However, when a person learns to read or write, or masters a second or third language, we are no longer dealing with instances of quantitative change; these changes are qualitative since they cannot be expressed through numbers. The acquisition of the ability to speak another language is not about the number of memorized words but about the acquisition of a new quality. Similarly, when a caterpillar gradually transforms into a chrysalis and then into a butterfly, it acquires entirely new qualities, such as the capacity to fly.

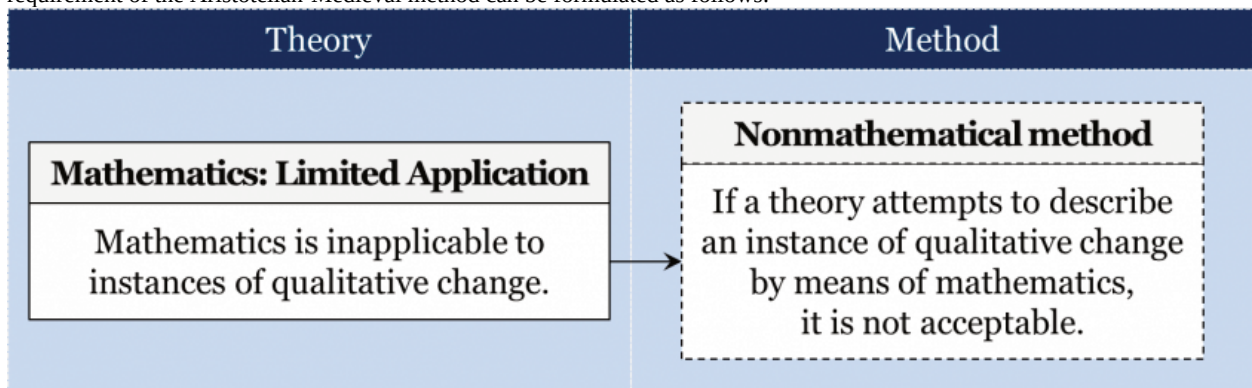


Because qualitative changes were thought to be unquantifiable, it was accepted in the Aristotelian-Medieval mosaic that the distinction between *qualitative* and *quantitative* changes is strict. This is not something we would agree with nowadays in the 21st century. For example, we might argue that the learning of a new language consists, fundamentally, of quantitative changes in the strengths of vast numbers of synaptic connections between nerve cells in the learner's brain. But, at the time, it was believed that qualitative changes are really not about changes in shape, size, or number, but about the acquisition or loss of qualities.

Now, what does this tell us about the applicability of mathematics? By definition, mathematics is only applicable to *quantifiable* phenomena, i.e. phenomena which are explicable in terms of number, shape, or size. Indeed, in order to apply mathematics to a certain type of change, we must first manage to quantify that change, i.e. to express that change through numbers. But if qualitative changes are not expressible numerically, then no mathematics can possibly be applied to them. That's why it was accepted in the Aristotelian-Medieval mosaic that mathematics is inapplicable to instances of qualitative change:



By the third law, it followed from this belief that no mathematical description of a qualitative change can be accepted. This requirement of the Aristotelian-Medieval method can be formulated as follows:



This restriction affected all those fields of inquiry that dealt with qualitative changes, including biology (natural history) and physics (natural philosophy). Of course, there was no reason why mathematics could not be applied to those biological phenomena that were quantifiable; nobody would object to counting, say, the number of limbs or teeth of a certain creature. In the same fashion, there was no objection to applying mathematics to the study of locomotion, since any change of position is readily expressed geometrically. However, when it came to describing the instances of qualitative change, the application of mathematics was believed to be impossible. Consider a lion cub that gradually learns and finally acquires the ability to hunt, or consider a

human being that learns a certain craft – such changes were believed not to be amenable to mathematical treatment. Once again, we see how a simple line of reasoning that stemmed from a strict distinction between quantitative and qualitative had serious consequences for the practice of science.

Summary

Summarizing a worldview is not an easy task. Perhaps the best way of doing it is by comparing one worldview to another; that's what we will be doing in chapters 8, 9, and 10, when we study the Cartesian, Newtonian, and Contemporary worldviews respectively. In the next few chapters, we will be revisiting the Aristotelian-Medieval worldview on several occasions and will show how it was different from the other worldviews.

For now, let's appreciate how interconnected different elements of the Aristotelian-Medieval mosaic were. Having as its root the requirement to accept only what is intuitively true, what is commonsensical, Aristotelian-Medieval science managed to achieve a high degree of mutual agreement between its different elements. Such a degree of mutual agreement has not since been replicated.

8.

Cartesian Worldview -- Introduction to History and Philosophy of Science

Chapter 8: Cartesian Worldview

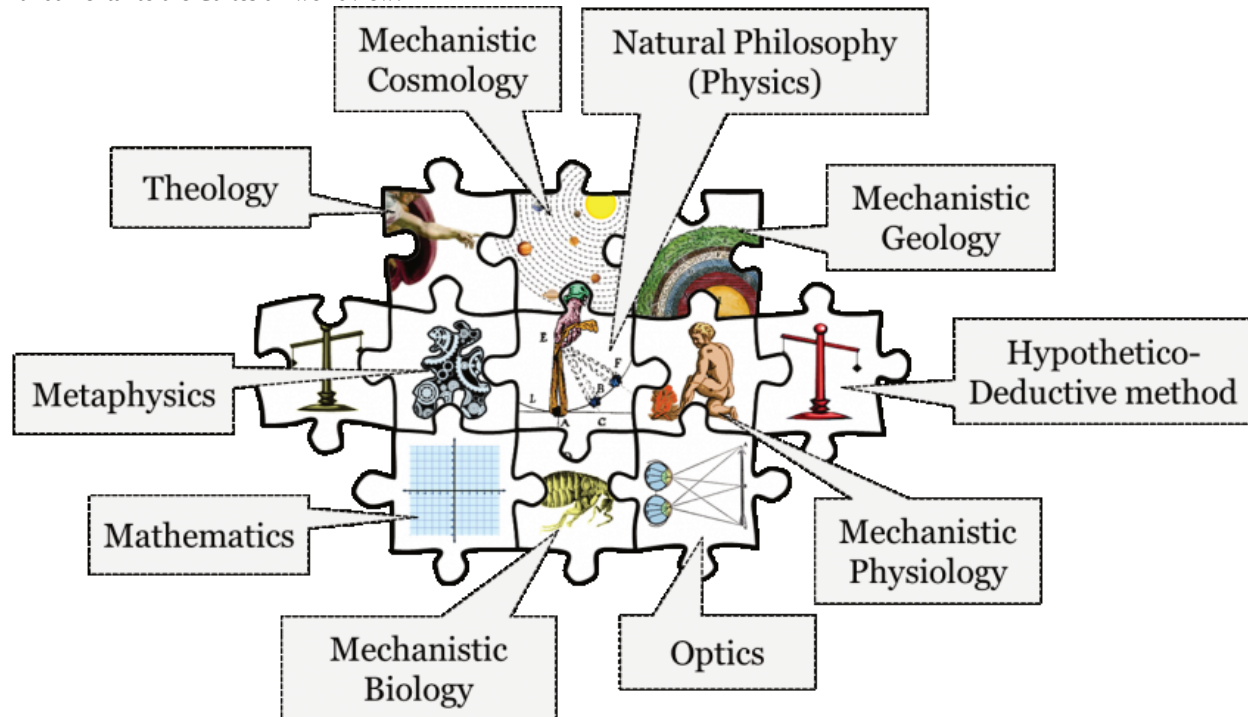
Intro

Usually, the seventeenth century is treated as an intellectually revolutionary period in the history of science. Filling the gap between the medieval Aristotelians and the massively influential Isaac Newton were the likes of Galileo Galilei, who tried and failed to convince the Catholic Church to accept a heliocentric cosmology, Robert Boyle, who experimented with gases in his air pump, and Pierre Gassendi, who renewed interest in ancient Greek atomic theory. However, what is often forgotten about all these theories is the context in which they were proposed; they were only ever *pursued* in the seventeenth century. In fact, as we learned in chapter 7, the Aristotelian-Medieval worldview was accepted by most scientific communities of Western Europe until the very end of the 1600s. So why didn't the scientific communities of Western Europe accept any of the theories of these revolutionary figures? Simply, they all failed to satisfy the requirements of the then-employed Aristotelian-Medieval method.

Well, all except for one. So far as we can tell, René Descartes (1596–1650), a French philosopher and mathematician, was the first to assemble an entire mosaic of theories that challenged the central tenets of the Aristotelian-Medieval worldview while, at the same time, satisfying the demands of the Aristotelian-Medieval method. By the beginning of the eighteenth century, textbooks and encyclopedias began to document and indicate the acceptance of this new worldview. We have come to call this the Cartesian worldview. It is the topic of this chapter.

Most of the key elements of the Cartesian mosaic were accepted or employed for a very short time: just four decades, from about 1700 to 1740. We say that this mosaic was accepted “on the Continent”, and by this we mean on the European continent; in places like France, the Netherlands, and Sweden, for example. This is to differentiate the acceptance of the Cartesian worldview “on the Continent” from the acceptance of the Newtonian worldview “on the Isles” – meaning in Britain (the topic of our next chapter), since both worldviews replaced the Aristotelian-Medieval worldview at the same time.

The key elements of the Cartesian mosaic include mechanistic metaphysics, theology, mechanistic physiology, natural philosophy or physics, mechanistic cosmology, mathematics, mechanistic geology, optics and biology, as well as the hypothetico-deductive method. You'll notice the recurrent theme of *mechanical philosophy*, or *mechanicism*, in these elements, an idea fundamental to the Cartesian worldview:



In this chapter, we will elucidate some of the most important elements of the Cartesian mosaic and start to uncover the metaphysical principles or assumptions underlying these elements. Let's begin with the metaphysics of the Cartesian mosaic, and what is perhaps Descartes' most well-known quotation, *cogito ergo sum*, or “I think, therefore I am”.

Cartesian Metaphysics

Descartes first wrote these words, “*I think, therefore I am*”, in his *Discourse on the Method* in 1637. What Descartes was trying to do with this simple phrase was actually much more ambitious than proving his own existence. He was trying to create a foundation for knowledge – a starting point upon which all future attempts at understanding and describing the world would depend. And of course, to know anything requires a *knower* in the first place, right? In other words, we need a *mind*. Descartes recognized this, so he sought to found all knowledge on the existence of a thinker’s own mind.

To prove that his mind exists, Descartes didn’t actually start with the act of thinking. He started instead with the act of radical doubt: he doubted his knowledge of the external world, his knowledge of God, and even his knowledge of his own mind. The reason for this radical doubt was to ensure that he didn’t accept anything without justification. So, the first step was to doubt every single belief he held.

For one thing, Descartes understood that sometimes our senses deceive us, so we must doubt everything we experience and observe. He further understood that even experts sometimes make simple mistakes, so we should doubt everything we’ve been taught. Lastly, Descartes suggested that even our thoughts might be complete fabrications – be they the products of a dream or the machinations of an evil demon – so we should also doubt that those thoughts themselves are correct.

Yet, in the process of doubting everything, he realized that there was something that he could not doubt – the very act of doubting! Thus, Descartes could at least be certain that he was doubting. But doubting is itself a form of thinking. Therefore, Descartes could also be certain that he was thinking. But, surely, there can be no thinking without a thinker, i.e. a thinking mind. Hence, Descartes proved to himself the existence of his own mind: I doubt; therefore, I think; therefore, I am.

Importantly for Descartes, this argument is intuitively true, as anyone would agree that doubting is a way of thinking and that thinking requires a mind. Thus, according to Descartes, the existence of a doubter’s own mind is, for the doubter, established beyond any reasonable doubt.

It’s important to notice what Descartes is claiming to exist through this line of reasoning. He’s not suggesting that his physical body exists. So far, he has only proven to himself that his mind exists. If you think about it, the second “I” in “I think therefore I am” should really be “my mind” – hence, “I think, therefore *my mind* is”.

But, what about the existence of everything else? How does he know that, say, the buildings of the city around him truly exist? For this, Descartes turns to *theology*, a central piece of the Cartesian mosaic. You see, since our senses can deceive us, we need some sort of guarantor to ensure that what we perceive is, in fact, the case. For Descartes, this guarantor is none other than God. If Descartes can prove that God exists, then it will follow from God’s benevolence that the objects that appear clearly and distinctly to us actually exist in the external world.

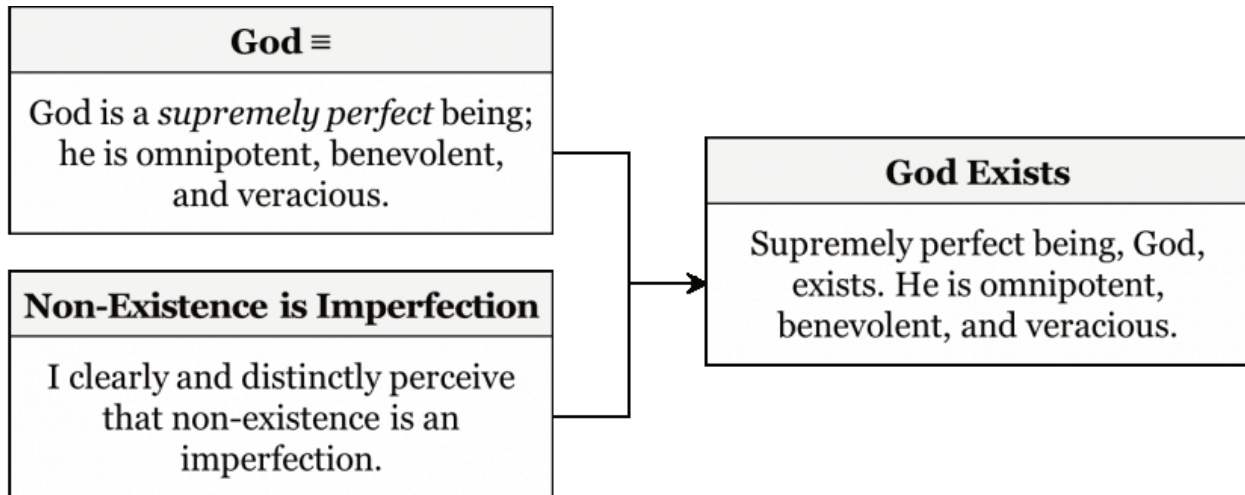
Descartes starts with a definition of God, as a supremely perfect being:

God ≡
<i>God is a supremely perfect being; he is omnipotent, benevolent, and veracious.</i>

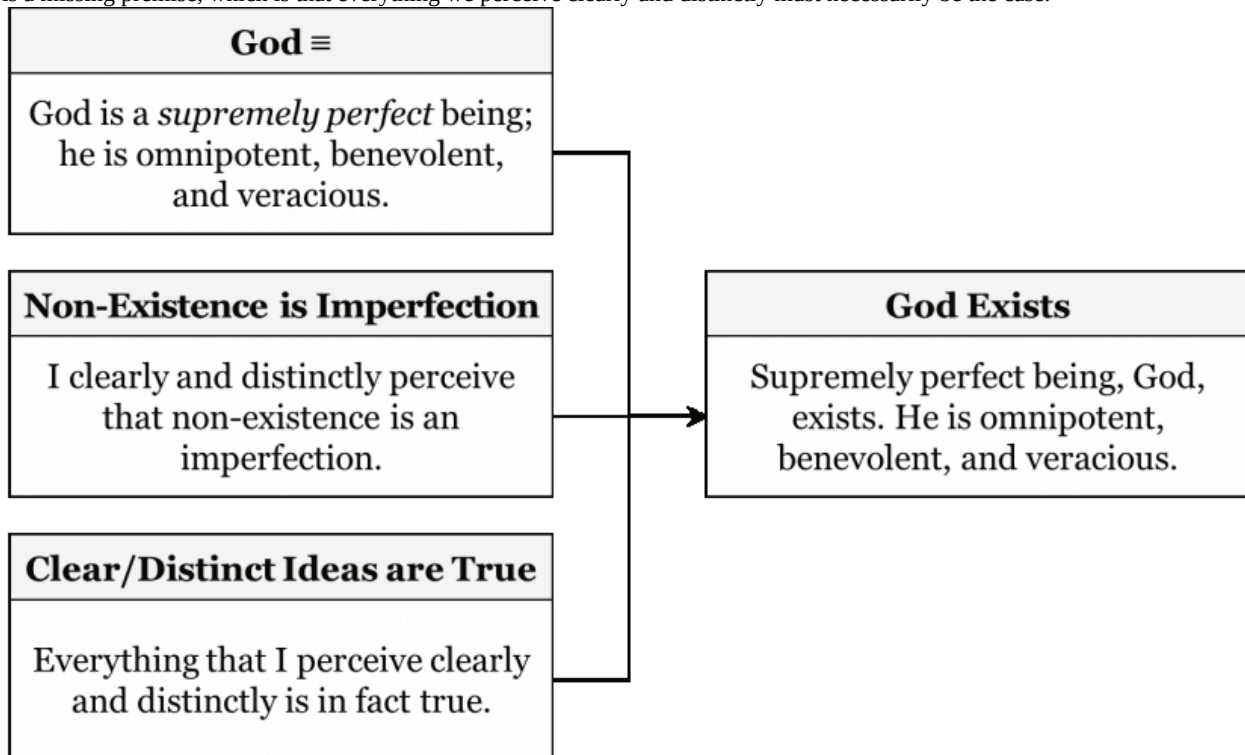
Theologians of the time would readily accept Descartes’ definition. After all, the idea of God as a supremely perfect being is implicit in all Abrahamic religions. So here Descartes simply spells out the then-accepted definition of God as a being of *all* perfections: he is *omnipotent* or perfectly powerful, he is *omniscient* or perfectly knowing, he is *omnibenevolent* or perfectly good, and so on. Any other perfection, God necessarily has by definition. If being *omniungular*, or having perfect nails, were a form of perfection, then God would have it.

Descartes then suggests that *existence* is part of being perfect. If God is perfect in all ways, then he must exist as well. But how does Descartes, or any Cartesian theologian, for that matter, know that existence is a requirement for perfection? Well, they would say that it would be impossible to be perfect without existing. Existence alone doesn’t make you perfect, but non-existence clearly makes you imperfect. Descartes suggests that we clearly and distinctly perceive this non-existence as an imperfection.

Therefore, according to Descartes, the most perfect being, i.e. God, must necessarily exist. This proof for the existence of God is also known as the *ontological argument*. It was first proposed by Saint Anselm, Archbishop of Canterbury centuries before Descartes. Briefly, the gist of the argument goes like this:

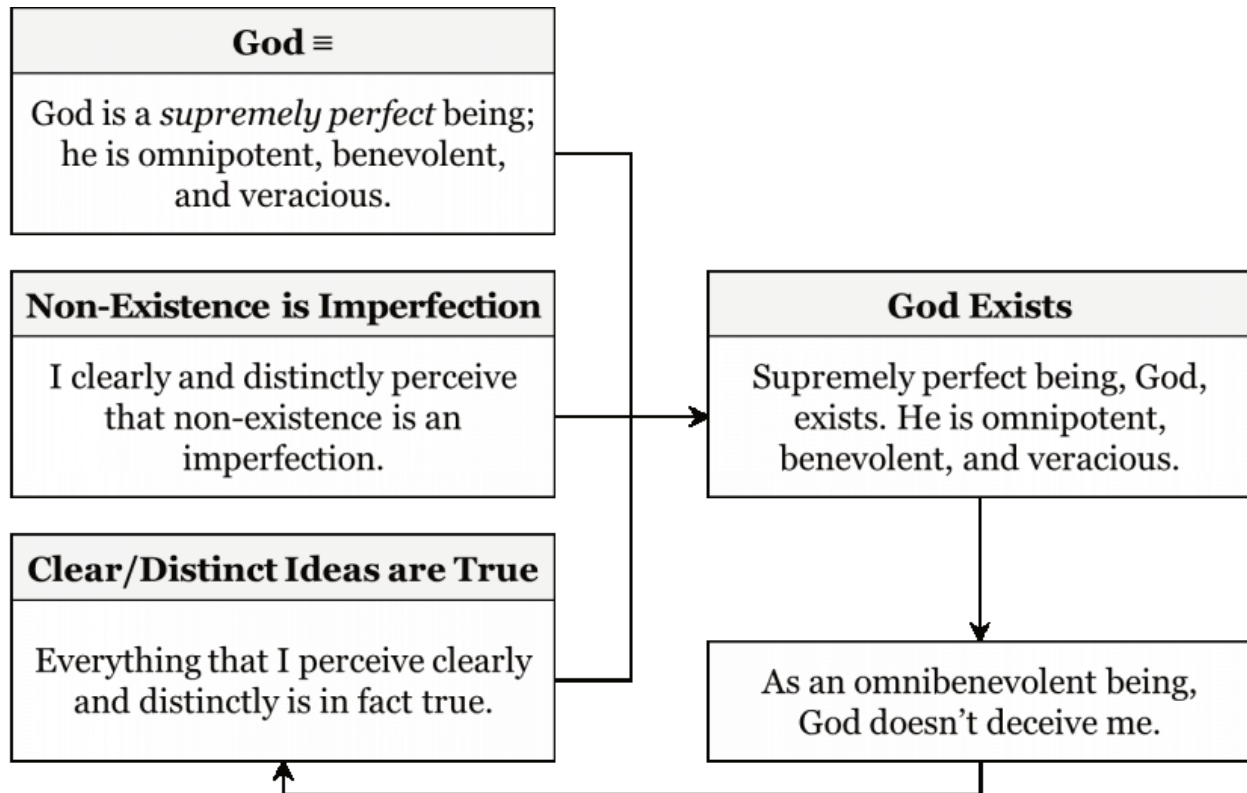


There are several objections to Descartes' ontological argument – we're going to consider only one of them. This objection is that the truth of the clear and distinct perception that non-existence is an imperfection is actually a circular argument. Specifically, God's existence doesn't follow merely from the definition of God and our perception that non-existence is an imperfection. There is a missing premise, which is that everything we perceive clearly and distinctly must necessarily be the case.



In fact, Descartes clearly accepts this premise, but he justifies it by pointing out that God is not a deceiver, from which follows that everything we perceive clearly and distinctly must be the case. Since God is all-good, he is not a deceiver. Therefore, everything we clearly and distinctly perceive must be true. Indeed, if God allowed our clear and distinct ideas to be false, he would be a deceiver, which would go against his very nature as a benevolent being.

So, the truth of Descartes' clear and distinct perception that non-existence is an imperfection proves that God exists. The perfect goodness of God ensures that Descartes' clear and distinct perception (that non-existence is an imperfection) is true. However, to prove that God exists, Descartes relies on the assumption that everything he perceives clearly and distinctly must be true. And, to justify this assumption, Descartes relies on God's benevolence, and thus his existence. This is a *circular argument*: it is problematic because the truth of at least one of the premises relies on the truth of the conclusion.



The existence of such objections to Descartes' ontological argument for the existence of God didn't hurt the prospects of Descartes' metaphysics because the community of the time already accepted God's existence. This is not surprising because Christian theology was part of virtually every European mosaic.

Having established that God exists, Descartes then proceeds to proving the existence of the world around him. Descartes makes this argument in two steps: first, he shows that the world is *external* to, independent from, or not fabricated by his mind, and second, he shows that this external, mind-independent world is a *material* one.

To show that the external world is real and is not a figment of his imagination, Descartes invokes something called the *involuntariness argument*. He notices that his sensations come to his mind involuntarily, i.e. that he has no control over what exactly he perceives at any given moment. But, clearly, everything happening in his without the mind's control should necessarily have its source outside his mind. Therefore, he clearly and distinctly perceives that something external to his mind exists. Since everything that he perceived clearly and distinctly must be true, it is the case something external to his mind exists.

As an example, Descartes suggests that we feel heat whether or not we want to, i.e. we feel it involuntarily. When the feeling of heat is outside our control, we suspect that it is not a product of our mind but instead the product of something external to our mind, like a fire. Since it's clear to us that sensations come from something outside the mind, then, it's also clear that things outside the mind must exist. But because we know that God is not a deceiver, everything we perceive clearly and distinctly must be true. Therefore, there exists something other than the mind.

Descartes still must show that this external world, which is the source of our sensations, is material. Descartes takes it as a premise that we have a clear and distinct perception that external things are material. He next refers to God's benevolence to show that external things are therefore actually material. Indeed, if the source of our sensations were not material, God would effectively be deceiving us by allowing for a clear and distinct idea of something that is not the case. Clearly this is impossible because God is benevolent and not a deceiver. Thus, Descartes eliminates both the possibility that God himself, or that something else, like an evil demon, is causing his sensations. If external sensations were caused by either God or something else like an evil demon, then Descartes would require some God-granted faculty or capacity to recognize such causes. Imagine having Spiderman's "spidey sense" for detecting danger, but for detecting God-caused or evil demon-caused sensations instead. This faculty would allow you to determine whether something you perceive is caused by God, an evil demon, or anything else immaterial. For instance, if you saw the Spiderman villain Green Goblin dancing on a table with this "spidey sense", you would be able to tell whether this sensation was caused by the actual material Green Goblin dancing on the table, or whether this sensation was produced by God or some other immaterial being. But we don't have this "spidey sense". Instead, Descartes suggests that God has given him a very different capacity, one that causes him to believe that his sensations are caused by material things. Again, if Descartes believed that his sensations were caused by material things when they were in fact not, then God would be a deceiver. Since God is benevolent and not a deceiver, the only possibility that remains for Descartes is that sensations are caused by the external material world, and therefore, the external, mind-independent, *material* world exists.

Now that Descartes has established the existence of both *mind* and *matter*, he can then establish what the indispensable properties of these two substances are. What are the qualities without which these substances are inconceivable?

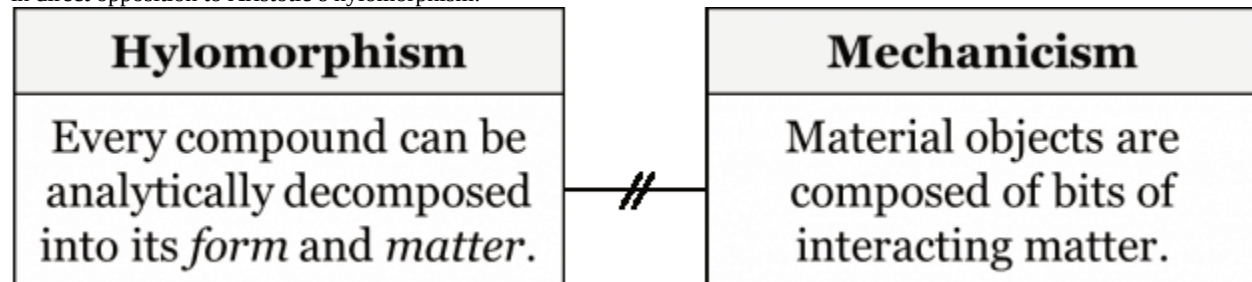
The indispensable property of mind is *thought*. For Descartes, it was impossible to conceive of a mind that lacked the capacity for thinking. What would it even mean for a mind to be non-thinking? The indispensable property of matter is *extension*, i.e. the capacity to occupy space. Just like a non-thinking mind, Descartes argues that it is impossible to imagine any material object that does not occupy some space. Does the mind occupy space? Can you pack it up in a box when moving apartments? Not at all, as the mind is an immaterial substance. Does a chair think? Nope. How about a table? Not at all. Chairs and tables merely occupy space, thus are extended substances, i.e. material.



By identifying the indispensable properties of mind and matter, Descartes was strengthening the foundations of the Cartesian worldview. It is from these two fundamental principles that other important Cartesian metaphysical elements follow. Since metaphysical elements are central to distinguishing one worldview from another, we will now introduce some metaphysical elements from the Aristotelian-Medieval worldview alongside those of the Cartesian worldview. In doing so, we can more clearly understand the most significant differences between the two worldviews.

Aristotelians subscribed to an idea known as *hylomorphism*. Hylomorphism says that all substances can be analytically decomposed into matter (from the Greek word *hyle* meaning *wood* or *matter*) and form (from the Greek word *morphē* meaning *form*). We know what matter is. A form is essentially an organizing principle that differentiates one combination of matter from another. The idea behind hylomorphism is not that you can literally separate any substance into its constituent parts, like a salad dressing into oil and vinegar, but that, in principle or conceptually, any substance contains both matter and form. For instance, a human body is composed of both some combination of the four humors (its matter) and a soul (its form). For Aristotle, the soul is a principle of organization responsible for digestive, nutritive, and vegetative functions as well as thinking and feeling. It organizes the humors. Blood is composed of some combination of the four elements (its matter) along with what we might call “blood-ness” (its form). Similarly, a Maple tree composed of some combination of elements (presumably earth and water) and the form of “maple-tree-ness” which organizes these elements into a Maple tree.

In the Cartesian worldview, the idea of hylomorphism was replaced by a new idea, *mechanicism*. Mechanicism says that, since the only attribute of matter is extension, all material things are composed solely of interacting material parts. Thus, every material thing could be understood as a mechanism of greater or lesser complexity. For instance, Descartes would describe the human body as a complex machine of levers, pulleys, and valves, or the heart as a furnace, causing the blood to expand and rush throughout the body. You might be wondering, what happened to the *form* of material things? Descartes actually doesn’t need forms. Since every material object could be explained in terms of the arrangement of extended matter in different ways, there was no need for a form to act as an organizing principle and do any arranging. The mere combination of constituent parts of a thing is sufficient, according to Descartes, to produce the organization and behaviour of that thing. Even the bodies of living organisms, including human bodies, could in principle be explained as complex arrangements of material parts. Digestion, development, emotions, instincts, reflexes and the like could therefore all be explained as the result of mechanical interactions alone. Descartes’ mechanicism was therefore in direct opposition to Aristotle’s hylomorphism:

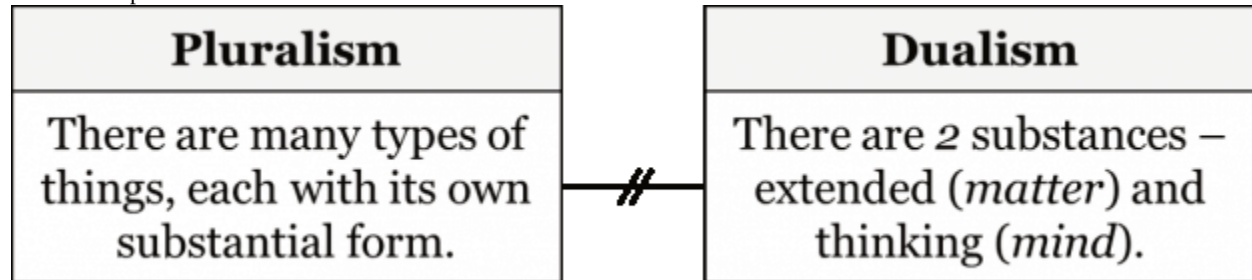


Descartes mechanistically explained the physiological functions that Aristotle assigned to the soul. But he felt that the human capacity for reason and language, what Aristotle called the rational soul, could not be explained this way. Thus, Descartes supposed that the human mind, or soul, was completely immaterial. This leads us to the next key metaphysical element of the Cartesian worldview – *dualism*.

But in order to understand the idea of dualism, we should start from the Aristotelian stance on the number of substances, known as *pluralism*. According to Aristotelian *pluralism*, there are as many substances as there are types of things in the world. Since Aristotelians accepted that each type of thing had its own substantial form, for them there is a plurality of substances. Lions, tigers,

and bears might all amount, materially, to being composed of some combination of the elements earth and water, but they are also *substantially distinct* through their forms of lion-ness, tiger-ness, and bear-ness, respectively.

For Descartes, there is not a plurality of substances, but only two independent substances, matter and mind. This is the conception of *dualism*. Since Descartes also showed interest in sword fighting, you might call him a dualist *and* a duellist! Duelling aside, Descartes argued that extended, material substance is completely distinct from thinking substance. Among material things are those lions, tigers, and bears, as well as rocks, plants, and other animals. Thinking substances, on the other hand, included entities with minds but without extended bodies, entities like God and angels. Thus, his dualism was meant as a replacement to Aristotelian pluralism:

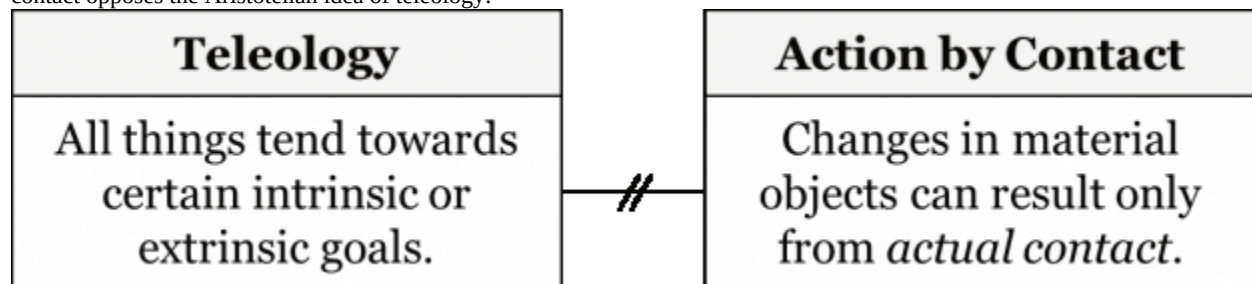


For Descartes, however, human beings remained a special case. While we are clearly material beings – having bodies extended in space – we also have the capacity for *thought*. Descartes felt that thought could not be explained mechanically, because all the machines known to him, such as clocks or pumps, inflexibly performed a single function, whereas human reason was a flexible general purpose instrument. As such, human beings were considered ‘citizens of two worlds’ because they clearly have a mind *and* occupy space.

The final metaphysical elements of the Aristotelian-Medieval and Cartesian worldviews to be considered here are those dealing with changes in objects. Why do material objects change from one shape to another, as a tadpole changes into a frog? Why do material objects move from one place to another, as in a rising flame? More generally, why do objects move or change?

Aristotelians answered such questions with reference to the purpose, aim, or goal of these objects, also known as their final cause. Accordingly, they subscribed to the principle of *teleology* (from the Greek word *telos* meaning *end*, *purpose*, or *goal*), which says that all things tend towards certain intrinsic or extrinsic goals. According to Aristotelians, a frog changes from a tadpole to its amphibious state because it has an intrinsic or innate goal to live by water and land. A flame flickers upwards because of its intrinsic goal to move away from the centre of the universe. Some goals are also extrinsic, or human imposed. A pendulum clock, for instance, shows the correct time because that is the goal imposed upon it by a clockmaker. Importantly, each goal is dictated by the substantial form of the object. So, a frog will never become an opera singer, for it lacks the substantial form of an opera singer.

Cartesians, however, not only rejected the notion of a substantial form, but denied that goals, aims, or purposes played any part in the behaviour of material objects. For them, matter has only one primary attribute – extension. So how do extended, material substances move or change? In place of teleology, Cartesians accepted the principle of *action by contact*, believing that changes in material objects can only result from actual contact between bits of matter in motion. So, the tadpole becomes a frog following the mechanical rearrangement of bits of its constituent matter; the flame moves upward because bits of fire are colliding and exchanging places with the bits of air above it. Similarly, a pendulum clock is not driven by any extrinsic goals but shows the correct time merely due to the arrangement of its levers, clogs, springs, and other mechanical parts. Clearly, this idea of action by contact opposes the Aristotelian idea of teleology:



Let’s recap the key Cartesian metaphysical elements we’ve discussed so far in the chapter. For Cartesians, the principal attribute of mind is thought and of matter is extension. They also accepted the principles of mechanicism, that all material objects are composed of bits of interacting and extended matter; dualism, that there are two substances that populate the world – mind and matter; and action by contact, that any change in material objects only results from actual contact between moving bits of matter. These Cartesian principles replaced the Aristotelian principles of hylomorphism, pluralism, and teleology respectively.

If changes in material objects can only result from actual contact, then how might we make sense of changes in living organisms? To better understand this, we must consider the Cartesian view of physiology in more detail.

Cartesian Physiology

In the seventeenth century, a great interest was shown in the inner workings of the body and the physiology of living things. Slowly, anatomists and surgeons began questioning the humoral theory of the Aristotelian worldview. Though not an anatomist himself, Descartes, too, wrote about physiology in his *Treatise on Man*, describing the balance and imbalance of the humors in mechanical terms. In fact, Descartes viewed the human body as a machine of interconnected and moving parts, piloted by the human mind. Seeing the body as a machine makes even more sense when we consider the Cartesian principle of action by contact: again, if changes in material objects can only result from actual contact, then changes in living organisms must be caused by colliding bits of matter.

As an example, in an episode of the children's cartoon *The Magic School Bus*, a very Cartesian view of the sense of smell is conveyed, although unintentionally so. In the episode, the class wants to better understand the olfactory system to figure out another student's almost prize-winning concoction of scents. So, they shrink with their bus to a very, very tiny size, and fly up into the student's nose. Once inside the nose, the class wants not only to smell what the student smells, but to *see* what she smells too. They all put on smell-a-vision goggles, which enable them to observe the particles of different types of smells float around and eventually land at specific spots inside the nose. Though the teacher, Ms. Frizzle, and her class were probably not really Cartesians, *actual* Cartesians would have accepted that smell works in a very similar way. They would have described tiny bits of matter, or the particles of a certain smell, landing somewhere inside the nose, and then mechanically transmitting that smell to the brain using tugs on a tiny thread running through the centre of the olfactory nerve. The same idea applies to other senses, like sight being the product of tiny bits of matter impressing on the eyes and then transmitting an image to the brain.

Cartesians also understood other organs in mechanical terms. Most obviously, they viewed the heart as a furnace, causing expanding heated blood to rush into the arteries and throughout the body. The blood was refined by its passage through blood vessels at the base of the brain into a fiery, airy fluid called *animal spirits*, which could flow outward through the nerves to the muscles. Note that this substance is simply a material fluid and has nothing to do with spirits in the *spooky* sense. Muscle contraction occurred when animal spirits flowed into a muscle and inflated it. Descartes even posited a system of valves to ensure that opposing pairs of muscles did not simultaneously contract.

If humans are citizens of two worlds, the worlds of matter and of mind, then how did Cartesians understand the interactions between these two separate substances? The problem is that, in order to interact, substances need to have some common ground. How can something immaterial, like the mind, affect something material, like the body, and vice versa? Surely there are no particles being transferred from the mind to the body because the mind doesn't have any particles – it's immaterial. For the same reason, the particles of the body cannot be transferred to the mind. Yet it is clear that the body and the mind act synchronously, there is a certain degree of coordination between the two. Consider the example of a contracting muscle. It is evident that we can control the bending of our arms; if you will your arm to move, it moves. Once the body receives a command from the mind to move, it proceeds purely mechanically: animal spirits flow through your nerves to your arm, and then cause the contraction of your bicep by inflating it. But how does a material body receive a command from an immaterial mind? In the Cartesian worldview, there wasn't an accepted answer to this question. Instead, at least two major solutions were pursued – *interactionism* and *parallelism*.

According to *interactionism*, the mind and the body causally influence each other. It appears to be intuitively true that the two do in fact influence each other. After all, that's what we seem to experience in everyday life. For Descartes, this interaction took place within the pineal gland: this is where the will of the mind was supposedly transferred into the motion of material particles. Movements of this tiny gland within the cavities of the brain were supposed to direct the flow of animal spirits accordingly. But Descartes' contemporaries and even he himself came to realize that this doesn't explain how the immaterial mind can possibly affect the material body. Since, according to their dualism, the two are supposed to be completely different substances – thinking and extended – the mechanics of their possible interaction remained a mystery.

The opposite view, called *parallelism*, suggested that the world of the mind and the world of matter exist parallel to one another, and there is no actual, causal interaction between them. They simply run in pure harmony with one another, like two synchronized clocks, which show the same time and appear as though there is coordination between them, while in reality, they're merely preprogrammed to show the right time. But how could two distinct substances act synchronously? According to one version of parallelism, this harmony between the two substances was pre-established by God, like the time on two different clocks being pre-established by a clockmaker. One problem with parallelism which left it a pursued theory was that it seemingly denied the existence of free will because it presupposed that both matter and mind are preprogrammed.

Since both of these views had serious drawbacks, there was no accepted Cartesian answer to the problem of seeming body-mind interaction. Yet, this didn't dissuade Cartesians from their dualism. A few mechanists, like the French physician Julien Offroy de La Mettrie were willing to question Descartes' dualism and countenance the possibility that the mind itself was mechanical. La Mettrie's argument for a mechanical mind was simple. How could mental processes like reason and language be affected by physical causes like a fever, or getting drunk if the mind itself wasn't a physical mechanism? But a mechanical view of the mind didn't become accepted until the twentieth century, when the advent of computers and neural network simulations made it possible to imagine how the mind could be a product of material processes. In the context of the Cartesian mosaic, mind-body dualism remained one of the central elements along with the ideas of mechanism and action by contact. They believed that all material phenomena are explicable in terms of material parts interacting through actual contact.

But is it possible to explain *all* phenomena through this mechanistic approach? While many material objects, like clocks, and

analogous parts of the human body, like beating hearts and moving arms, can be readily explained as mechanisms working through action by contact, it's not quite clear that all material processes are amenable to such a mechanistic explanation. For instance, how can such phenomena as *gravity* or *magnetism* be explained mechanistically? After all, two magnets seem to attract or repel each other without actually touching. To appreciate the Cartesian explanation of gravity and magnetism, we need a better understanding of Cartesian physics and cosmology.

Cartesian Physics and Cosmology

In his own times, Descartes' writings formed an integrated system of thought. Among modern philosophers, Descartes is well-known for his epistemological and metaphysical writings, but his complex and ingenious physical and cosmological theories, which were an equally important part of his system, are known only to historians and philosophers of science. Foremost among his physical theories are his laws of motion, first set out in 1644 in his *Principles of Philosophy*.

Recall that the only capacity of material objects is to occupy space, i.e. to be extended, and that interactions between extended substances necessarily occur by contact. From the principle of action by contact, we can deduce *Descartes' first law of motion*. The law states:

Every bit of matter maintains its state of rest or motion unless it collides with some other bit of matter.

The metaphor of billiard balls is most apt for appreciating Descartes' first law. A billiard ball will stand still on a table up until the point that it is knocked in some direction by another billiard ball. Once knocked, the ball will continue moving until it hits another billiard ball or the edge of the table. According to Descartes, the same principle applies to all matter: every bit of matter stands still or continues moving until it contacts some other bit of matter.

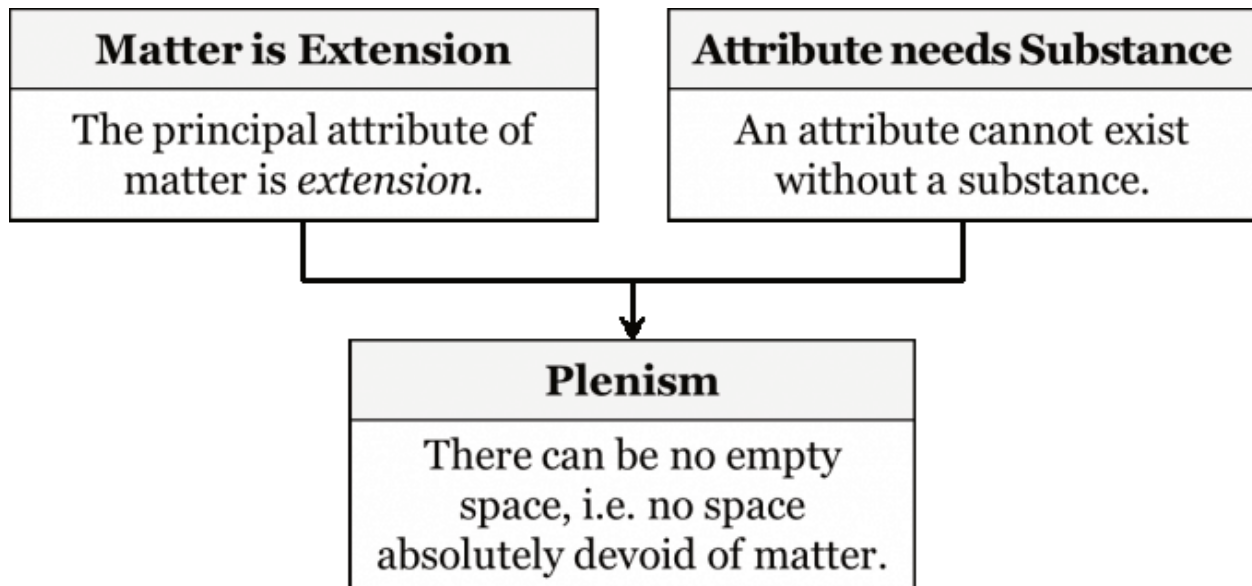
Descartes' *second law of motion* builds on the first. It states:

Every bit of matter, regarded by itself, tends to continue moving only along straight lines.

The law suggests that the unobstructed motion of a material object will always be rectilinear. A billiard ball, for instance, continues to move along a straight line until it reaches an obstacle. When it finally reaches the obstacle, it changes its direction, but then continues in another straight line. Compare this with a stone in a sling. As you swing the stone around, its motion follows a circular path because it is being obstructed by the sling, i.e. it is not regarded by itself. This is the gist of Descartes' second law.

A question arises: how about those numerous instances of non-rectilinear motion that we come across every day? For Cartesians, the moon was made of the same sort of matter as things on Earth, but it travelled in a roughly circular path around it. And, what about the parabolic motion of a projectile? There doesn't seem to be any obstruction in either case, and yet the motion is far from rectilinear. To answer this question, we need to appreciate the Cartesian take on gravity.

If extension is the attribute of matter, then, according to Descartes, extension cannot exist independently of material things. Thus, for Cartesians there is no such thing as empty space. Rather, Cartesians accepted the principle of *plenism* (from *plenum* (Latin) meaning "full") according to which the world is full and contains no emptiness. In this sense, a seemingly empty bedroom – one without a bed, desk, chair, dresser, or posters of Taylor Swift – is not, in fact, empty; it is still full of invisible bits of matter, like those of air. In fact, the entire universe is full of matter, according to Cartesians. So, the seeming empty space between the moon and the earth, and the planets and stars is not actually nothing, it's a whole lot of invisible matter. Descartes' reasoning for this stance is that space is not a separate substance; it is an attribute of matter. But an attribute (a property, a quality) cannot exist without a certain *thing* of which it is an attribute. For instance, redness cannot exist on its own unattached to anything red; it necessarily presupposes something red. Similarly, extension or space cannot exist without anything material, because extension is just an attribute of matter.

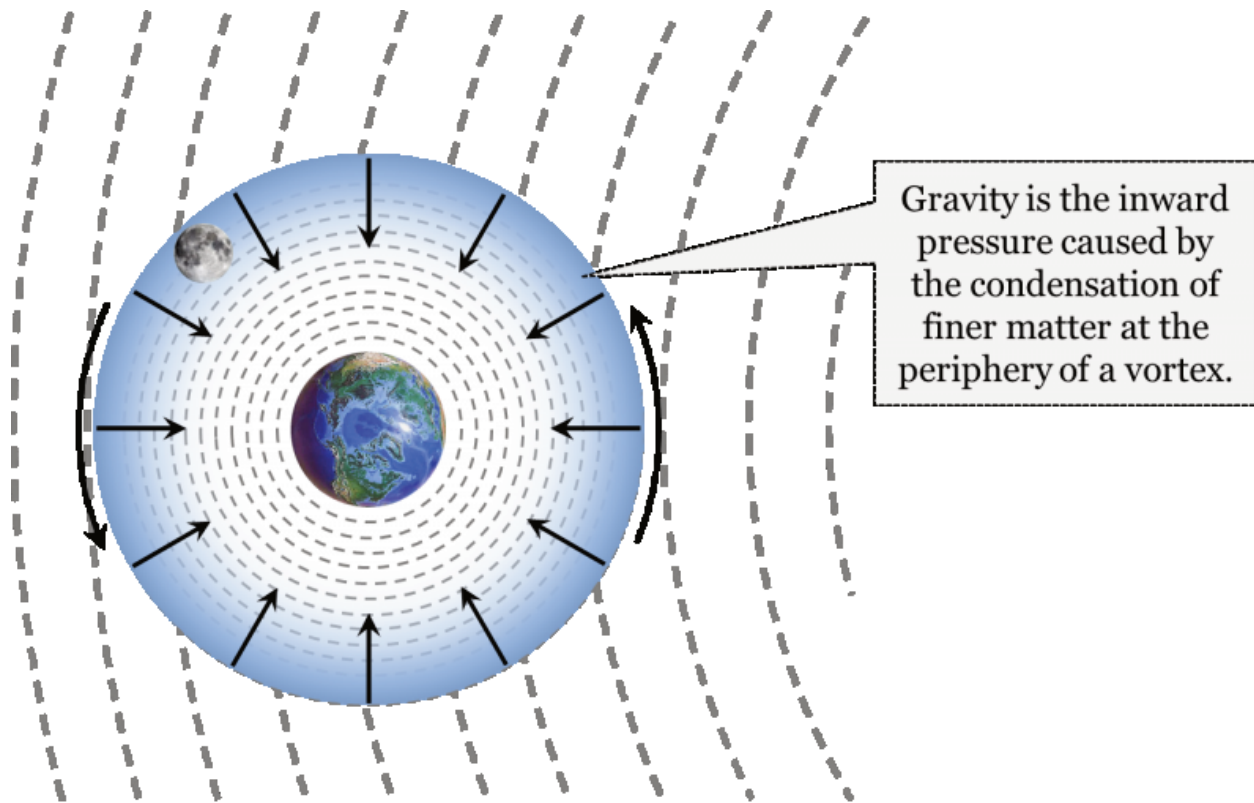


Since the universe is full, i.e. it is a plenum; the ideal of unobstructed, rectilinear motion described by Descartes' second law becomes an unlikely occurrence. In practice, all motion turns out to be an interchange of positions, that is, all motion is essentially *circular*. Imagine two scenarios involving particles. In the first, empty space does exist. If a particle moves in this scenario, it can move in a straight line into some empty space and would leave a new empty space in its place. In the second scenario, however, empty space does not exist. If a particle moves in this plenum, it might be blocked by all of the particles around it. It is also impossible for this particle to leave any empty space upon vacating its initial location. Its only option is to *exchange* places with the particles that surround it. Neither particle can pass through the other, as they are both extended, so they do a little dance, spinning around each other until they occupy their new places. For Cartesians, motion along a straight line is what a bit of matter would do if it wasn't obstructed. But the world being a plenum means that matter is always obstructed, and that *actual* motion always ends up being that of circular displacement.

The idea that all motion is eventually circular is the basic principle underlying the Cartesians' so-called vortex theory. Among many phenomena, this theory explains gravity, the revolution of moons and planets, and even the motion of comets. In a word, vortex theory says that all planets are lodged in the centre of their own vortex of visible and invisible matter, each of which is, in turn, carried about an even greater vortex centred on the sun. But the complexities of the theory warrant a closer look.

According to Cartesians, every star, like our Sun, revolves around its own axis. This revolution creates a whirlpool of mostly invisible particles around the star. As these particles are being carried around the star in this giant whirlpool, or vortex, they gradually stratify into bands, some of which later form into planets also carried by the vortex around the star. But because each of these planets can also rotate around its own axis, it creates its own smaller vortex. This smaller vortex can carry the satellites of a planet, which is how Cartesians explain the revolution of the Moon around the Earth, or Jupiter's moons around Jupiter. The idea of vortices, for Cartesians, could potentially be extended beyond solar systems; there might be even greater vortices that carry individual stellar vortices. In sum, for Cartesians, the whole universe is like a huge whirlpool that carries smaller whirlpools that carry even smaller whirlpools, ad infinitum.

This vortex theory provides a simple mechanistic explanation of gravity. Cartesians accepted that the Earth stood in the centre of a vortex with its edges or a periphery extending to just beyond the moon. So, between the surface of the Earth and the far side of the moon are layer after layer of invisible bits of matter. Now, what happens to all of that matter when the Earth rotates around its axis? Every bit of matter is carried away from, or flees, the centre of the vortex. That is, the rotating Earth generates a *centrifugal force*. In effect, the innermost layer of invisible particles is constantly being pushed towards the periphery of the vortex. But the inner layers cannot leave any voids as they move away from the centre, since according to Cartesians there is no empty space. Since there is no empty space beyond the terrestrial vortex, the particles flying away from the Earth due to centrifugal force eventually accumulate at the periphery of the vortex, and thus create an inward pressure. Thus, Cartesians reasoned that it is only the lightest bits of matter that actually make it to the edge of the vortex. Meanwhile, the inward pressure pushes heavier bits of matter towards the centre of the vortex. According to Cartesians, this inward pressure is gravity. In essence, Cartesians described gravity mechanistically in terms of the inward pressure caused by the displacement of heavy matter, which is itself a result of the centrifugal force generated by a rotating planet at the centre of a vortex.

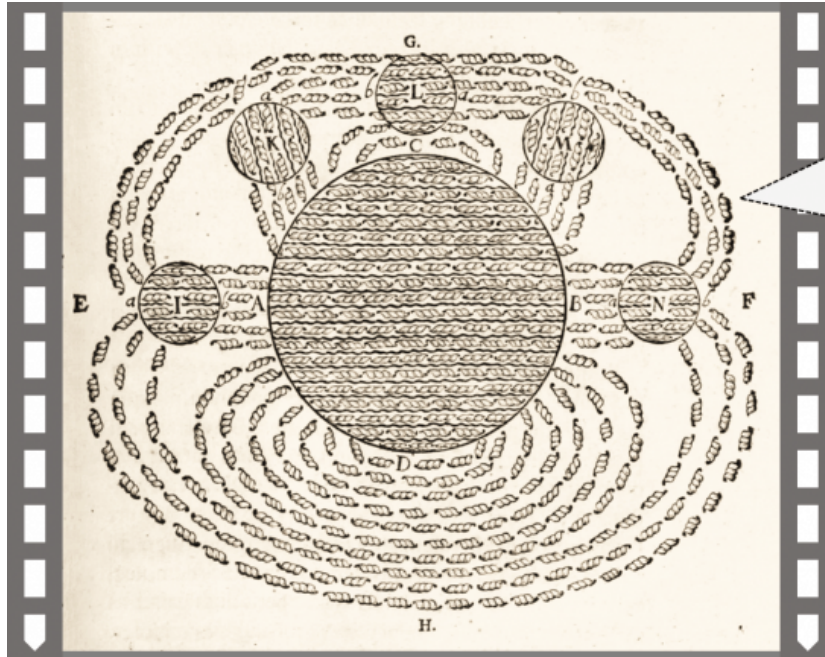


As an analogy, consider a bucket of water with a handful of sand in it. As you begin stirring the water inside the bucket, you notice the grains of sand being carried around the bucket by the whirlpool you've created. Initially, the grains of sand are spread throughout the bucket. Gradually, however, they begin to collect in the centre of the whirlpool, being displaced by the finer particles of water along the periphery. This is similar to what happens in any Cartesian vortex.

Of course, not all matter is moving back and forth through the ebb and flow of a vortex. Very noticeably, the Moon seems to maintain an average distance from the centre of the terrestrial vortex, while rotating around it. Descartes explained that the moon was in a position of relative equilibrium between the centrifugal force created by a rotating Earth and the inward pressure created by the accumulation of matter at the periphery of the vortex. He extended the same logic to the Earth's steady distance from the Sun.

Another natural phenomenon seems to challenge the mechanistic worldview of the Cartesians. This is *magnetism*. Seventeenth- and eighteenth-century physicists had observed and experimented with magnetic materials like lodestones for quite some time. They commonly placed iron filings around a lodestone to reveal its magnetic effects. At face value, magnets seem to be affecting iron without any actual contact, at a distance. But, how was this phenomenon accounted for by Cartesians?

The Cartesian explanation of magnetism starts with the idea that there are numerous parallel pores running between the north and south poles of any magnet. Let's take the Earth as an example. Imagine parallel tunnels running straight through the Earth, but tunnels so tiny that they're invisible to the naked eye. Each and every one of these tunnels or pores is corkscrew-shaped, and constantly passing through them is a stream of invisible, also corkscrew-shaped particles. The stream of particles runs in both directions: straight from the North to the South Pole, then out and around the Earth back up to enter from the North Pole again, and vice versa.



Magnetism is a result of the circulation of tiny helical particles through parallel threaded pores and through space around magnets.

Cartesians accepted the continuous flow of these corkscrew-shaped particles as the mechanistic cause of magnetic effects. Were Descartes to take, for instance, a lodestone, and observe that it turns to align with the Earth in a certain way, he would claim that the lodestone behaves like a magnet because it has the same corkscrew-shaped pores as the Earth. For Cartesians, the aligning of the lodestone with the Earth is in reality the turning of the lodestone until the corkscrew-shaped particles can flow smoothly through both the Earth and the lodestone.

Descartes explained numerous observations of magnetic phenomena with this mechanistic idea. Let's take the attraction, and then the repulsion, between two magnets as examples. For Descartes, two magnets with opposite poles aligned would *attract* one another for a couple reasons. First, since the magnets were already emanating corkscrew particles in the same direction, it was easy enough for their streams of corkscrew particles to merge, one with the other. Second, as soon as both magnets shared the same stream of corkscrew particles, those particles would displace any air particles from between the magnets, swiftly pushing the magnets even closer together.

The same poles of two magnets *repelled* one another for a couple reasons as well. First, the opposing streams of corkscrew particles pushed the magnets in opposite directions, like two hoses spewing water at each other. Second, Descartes suggested that the pores running through each magnet also contained one-way valves or flaps, which opened if the corkscrew particles passed through the pore in the right direction but closed if not. Magnets repelled when the corkscrew particles met a closed valve and pushed the magnet away.

In summary, Cartesians accepted the following physical and cosmological theories: they accepted Descartes' *first* and *second laws of motion*, which state that matter maintains a state of rest or motion unless it collides with some other bit of matter, and that it moves along a straight line when unobstructed, respectively. Cartesians accepted the metaphysical principle of *plenism*, that space is a secondary attribute of matter and, as a result, there is no such thing as empty space. Since the universe is full or a plenum, unobstructed motion becomes unrealistic, meaning that *all motion is, in practice, circular motion*. Finally, we discussed that the Cartesian views of gravity and magnetism were necessarily mechanistic, amounting to the products of *vortices* and *corkscrew particles* respectively.

Cartesian Method

Up until this point, we've gone over numerous metaphysical assumptions and accepted theories of the Cartesian worldview. What remains is to look at the methods Cartesians employed to evaluate their theories.

We first discussed the Cartesians' transition from the *Aristotelian-Medieval (AM) method* to the *hypothetico-deductive (HD) method* in chapter 4, but let's quickly review. Two important ideas follow or are deducible from the Cartesian belief that the principal attribute of matter is extension. One, any phenomenon can be produced by an infinite number of different arrangements of bits of extended matter. This first deduction opens the door to the possibility of post-hoc explanations. Two, all other attributes or qualities of matter, like colour, taste, or weight, are secondary attributes that result from different arrangements of bits of matter. This second deduction implies that the world is far more complex than it appears in observations. Both of these ideas, the principles of post-hoc explanations and complexity, became accepted as part of the Cartesian worldview, which satisfied the requirements of the AM method. The acceptance of the Cartesian worldview resulted in a change to the method by which theories were evaluated.

By the third law of scientific change, the HD method becomes employed because it is a deductive consequence of the principles of post-hoc explanations and complexity.

While Cartesians did reject the AM method as the universal method of science, they continued to employ a method of intuition in certain cases. Specifically, Cartesians continued to expect intuitive truths for formal sciences, like logic and mathematics, and for identifying their general metaphysical principles. For example, the existence of the mind is accepted because it's shown to be intuitively true by means of Descartes' "I think, therefore I am". Similarly, the principle that matter is extension is accepted because it also appeared to be intuitively true.

In the Cartesian worldview, the method of intuition can only reveal so much: it can only help unearth the most general characteristics of mind and matter, including the principles of logic, mathematics and metaphysics. But when it came to the explanations of more specific phenomena, the method of intuition was not thought to apply. Instead, the requirements of the HD method were employed when evaluating the explanations of specific phenomena.

Thus, any hypothesis about the inner structure of this or that process would require a confirmed novel prediction. Again, this is because any phenomenon can be explained by hypothesizing very different internal structures. For instance, the Cartesian theories of gravity or magnetism, with their not-so-intuitive vortices and corkscrew particles, would be unlikely to satisfy the requirements of the method of intuitive truth. Instead, their acceptance in the Cartesian worldview depended on the ability of these theories to successfully predict observable phenomena.

But the requirement for confirmed novel predictions was only one of the requirements of the time. There were additional requirements that theories in different fields were expected to satisfy. These distinct requirements have to do with the underlying assumptions of their respective worldviews. Let's consider three of them.

We know from chapter 7 that Aristotelians accepted that every non-artificial thing has its own nature – some indispensable quality that makes a thing what it is. This belief shaped one of the requirements of the AM method – namely, the requirement that any genuine explanation must reveal a thing's nature.

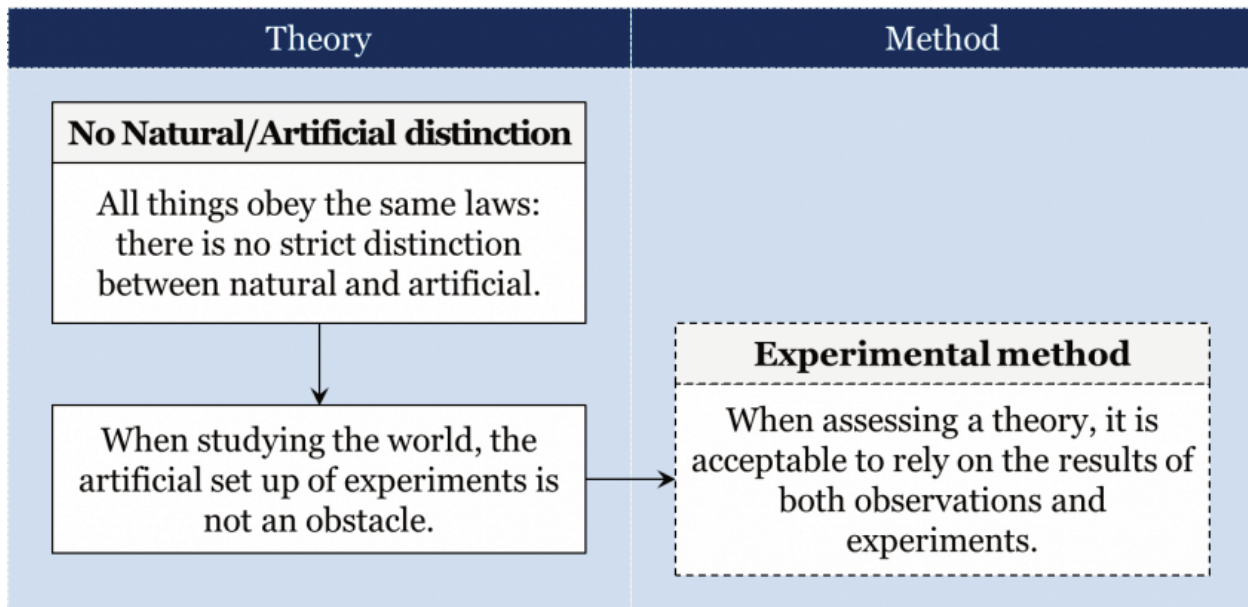
Cartesians, however, accepted a different fundamental assumption about the world. They accepted instead that every material thing is essentially akin to a mechanism, albeit often very complex. Thus, they required that any genuine explanation be a mechanical one. That is, Cartesians employed a *mechanical explanations-only method*. Consequently, even those phenomena which *seemed* non-mechanical and suggested the possibility of action at a distance would be considered mechanical in reality.

Say, a Cartesian at home alone observes a door slam shut on its own. She did not witness the mechanism by which the door shut – there was no other person present, nor any gust of wind. Still, as a good Cartesian she will only accept that some mechanical process led the door to close, lest she undermines the fundamental assumptions of her worldview.

There's an important caveat to the Cartesian requirement for mechanical explanations: it only requires mechanical explanations for material processes. As dualists, Cartesians would have a different method for mind-processes, or thoughts. Since thought is the principal attribute of mind, and since the fundamental science of thought is logic, then for Cartesians any explanation of mind-processes must be logical and rational.

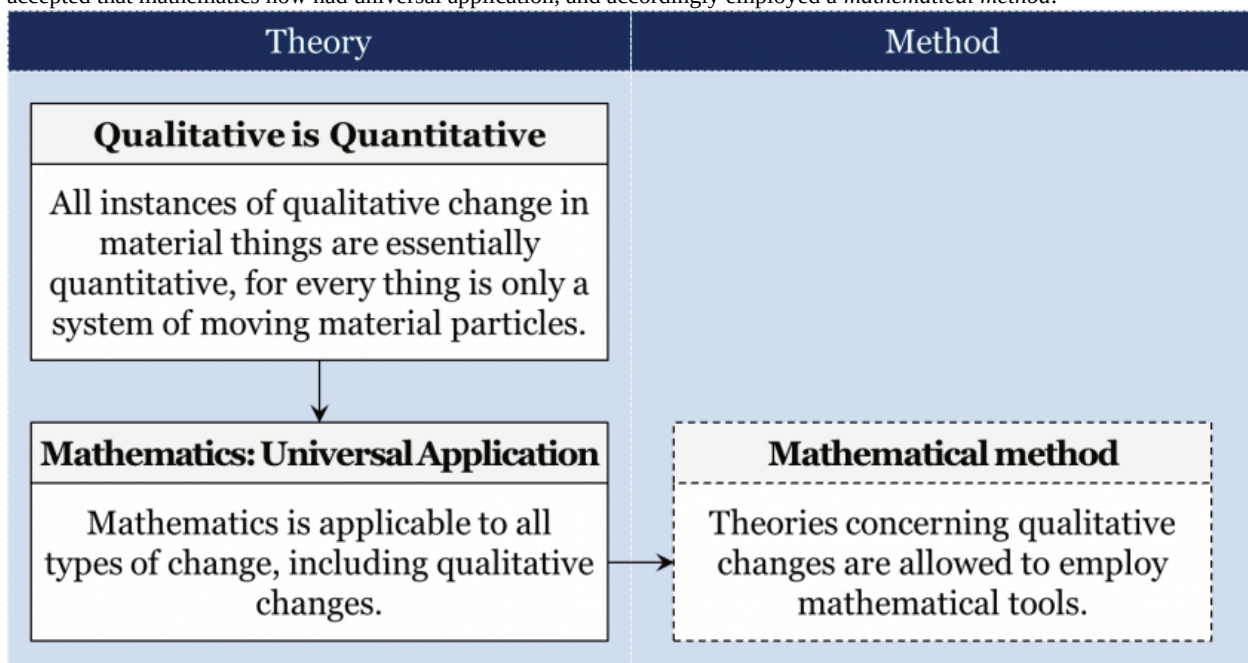
Additionally, Aristotelians clearly distinguished between *natural* things and *artificial* things. A natural thing, for Aristotelians, was a thing with an inner source of change. Anything not made by humans would be natural, as it would have a certain nature which would be the internal source of change for that object. For instance, a rock's quality of heaviness would be its inner source of change that would naturally direct the rock towards the centre of the universe. Similarly, an acorn's tendency to grow into an oak tree would be its inner source of change. Artificial things, on the other hand, are guided by something external to them, and don't have an inner source of change. A ship travels from port to port due to the commands of the captain and the actions of the crew. A clock shows the right time because it has been so constructed by its maker. Aristotelians accepted a strict distinction between natural and artificial things. Because the two have very distinct sources of change, explanations of natural and artificial things would have to be very different. This is the reason why Aristotelians didn't think that the results of experiments could reveal anything about the nature a thing under study. Experiments, by definition, presuppose an artificial set up. But since the task is to uncover the *nature* of a thing, studying that thing in artificial circumstances wouldn't make any sense, according to Aristotelians. You can certainly lock a bird in a cage, but that won't tell you anything about the nature of the bird. Thus, the proper way of studying natural phenomena, according to Aristotelians, was through observations only.

In contrast, Cartesians rejected the distinction between natural and artificial things, and, consequently, were very interested in experiments. For them, all matter is merely a combination of bits of interacting matter. Therefore, both artificial and natural things obey exactly the same set of laws. Since one aim of science is to explain the mechanism by which these bits of matter interact, it is now allowed to rely on experiments when uncovering these mechanisms. In other words, Cartesians employed an *experimental method*, which states that when assessing a theory, it is acceptable to rely on the results of both experiments and observations.



Another important distinction implicit in the Aristotelian-medieval worldview was that between *quantitative* and *qualitative* changes. A change was said to be qualitative if it concerned the qualities of a thing. For instance, when someone learns to read and write, she acquires a new quality. Similarly, when a caterpillar transforms into a butterfly, it thus acquires a host of new qualities, such as the ability to fly. In contrast, a change is said to be quantitative if it concerns number, shape, or size, i.e. something expressible in numbers. But because mathematics can only be applied to quantifiable phenomena, Aristotelians also believed that mathematics cannot be applied to instances of qualitative change. This brings Aristotelians to the non-mathematical method: if a theory about certain qualitative changes uses some mathematics, it is unacceptable. This explains why, in the Aristotelian mosaic, mathematics was only permitted in those fields of science which were concerned with quantitative phenomena, and had virtually no place in their biology, physiology, or humoral theory.

Cartesians rejected the strict distinction between quantitative and qualitative change. Again, Cartesians accepted that all matter is simply a combination of bits of interacting parts. So, any instance of qualitative change is in reality the rearrangement of these parts – a rearrangement that can be measured and expressed quantitatively. As long as qualitative changes remained indistinguishable from quantitative ones, mathematics could, in principle, be used to describe any material process whatsoever. That is, Cartesians accepted that mathematics now had universal application, and accordingly employed a *mathematical method*.

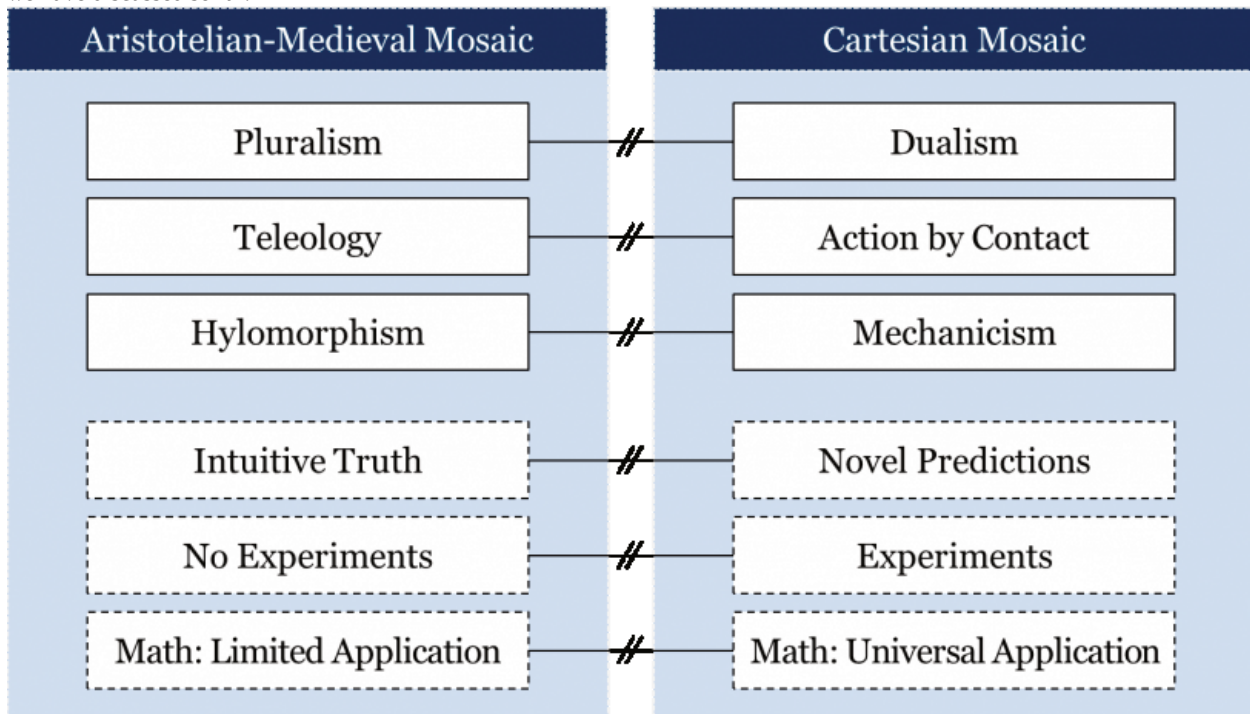


Let's recap this section on the methods employed in the Cartesian worldview. Foremost, Cartesians transitioned from employing

the AM method to employing the HD method, requiring confirmed novel predictions to accept any claim about the inner, mechanical structure of the world. They did continue to employ the method of intuition, but in a more limited way than the Aristotelians – only to evaluate the fundamental principles of logic, mathematics, and metaphysics. Cartesians also required that all explanations of material processes be mechanistic, that the results of experiments be considered when evaluating theories, and that mathematics is universally applicable.

Summary

The following table compares some of the key metaphysical assumptions of the Aristotelian-Medieval and Cartesian worldviews we have discussed so far:



We will revisit both the Aristotelian-Medieval and Cartesian worldview in the next two chapters and discuss some of their other metaphysical assumptions.

9.

Newtonian Worldview -- Introduction to History and Philosophy of Science

Chapter 9: Newtonian Worldview

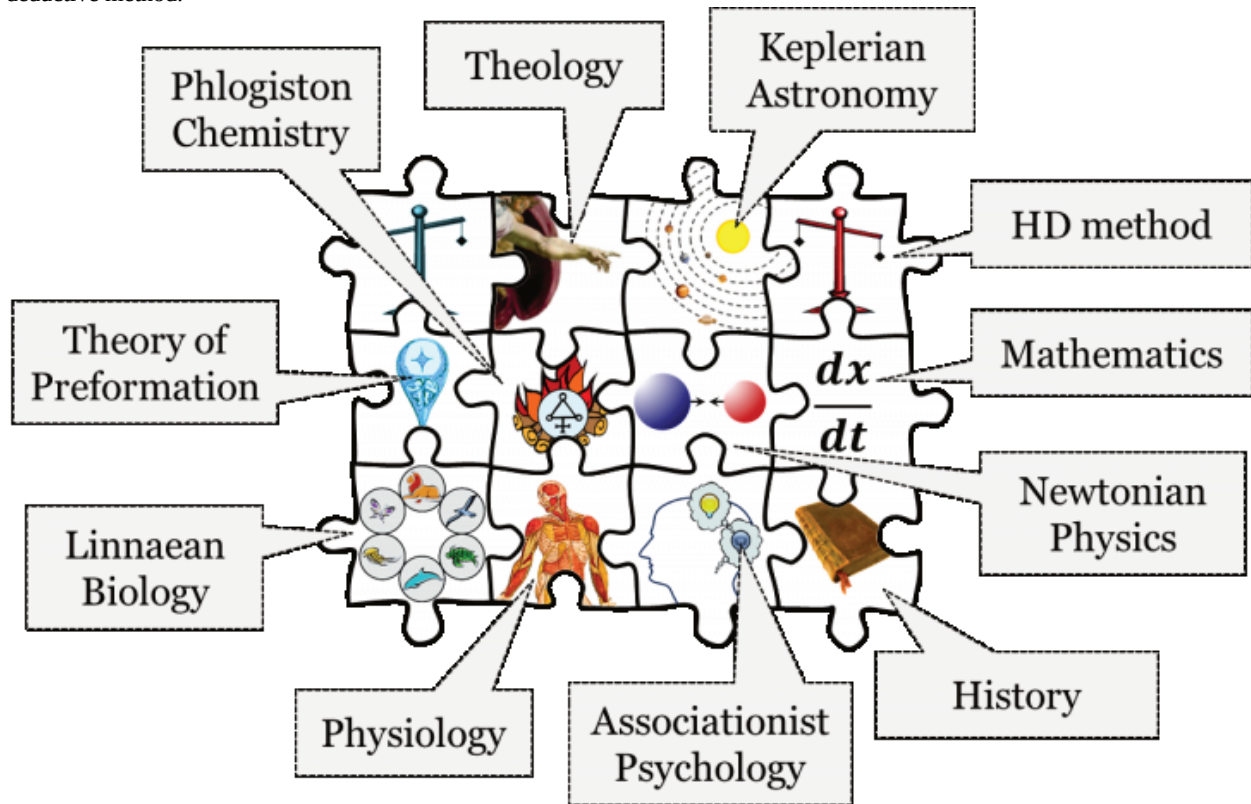
Intro

Back in chapter 4, we deduced a couple theorems from the laws of scientific change. One such theorem is called the *mosaic split theorem*, where the acceptance of two incompatible theories leads to a split in a mosaic. A very noteworthy split happened to the Aristotelian-Medieval mosaic around 1700, when theories of both the Cartesian and the Newtonian worldviews equally satisfied the expectations of Aristotelians. For a period of around 40 years between 1700 and 1740, two incompatible sets of theories were accepted by two very different communities. We covered the Cartesian worldview, which was accepted on the Continent, in chapter 8. In this chapter, we will cover the *Newtonian worldview*.

The Newtonian mosaic was first accepted in Britain ca. 1700. Continental Europe accepted the Newtonian mosaic around 1740, following the confirmation of a novel prediction concerning the shape of the Earth that favoured the Newtonian theory of gravity. The once-split Cartesian and Newtonian mosaics merged, leaving the Newtonian worldview accepted across Europe until about 1920.

One thing we must bear in mind is that the Newtonian mosaic of 1700 looked quite different from the Newtonian mosaic of, say, 1900; a lot can happen to a mosaic over two centuries. Recall that theories and methods of a mosaic do not change all at once, but rather in a piecemeal fashion. We nevertheless suggest the mosaic of 1700 exemplifies the same *worldview* as that of 1900 because, generally-speaking, both mosaics bore similar underlying metaphysical assumptions – principles to be elaborated on throughout this chapter.

That said, we can still understand and appreciate the key elements of the Newtonian mosaic at some particular time. In our case, we're going to provide a snapshot of the mosaic ca. 1765. Its key elements at that time included revealed and natural theology, natural astrology, Newtonian physics and Keplerian astronomy, vitalist physiology, phlogiston chemistry, the theory of preformation, Linnaean biology, associationist psychology, history, mathematics (including calculus) as well as the hypothetico-deductive method.



Let's start with the most obvious elements of the Newtonian mosaic – Newtonian physics and cosmology.

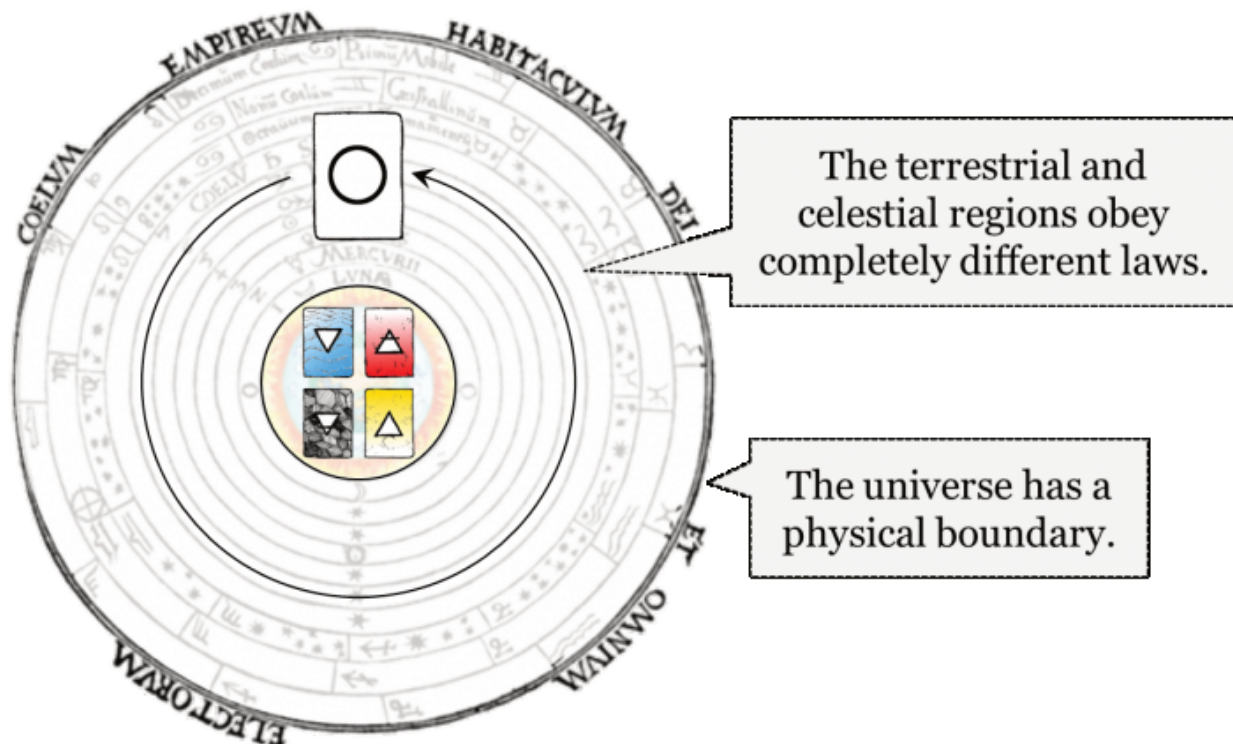
Newtonian Physics and Cosmology

In 1687, Isaac Newton first published one of the most studied texts in the history and philosophy of science, *Philosophiæ Naturalis Principia Mathematica*, or the *Principia* for short. It is in this text that Newton first described the physical laws that are part and parcel of every first-year physics course, including his three laws of motion, his law of universal gravitation, and the laws of planetary motion. Of course, it would take several decades of debate and discussion for the community of the time to fully accept Newtonian physics. Nevertheless, by the 1760s, Newtonian cosmology and physics were accepted across Europe.

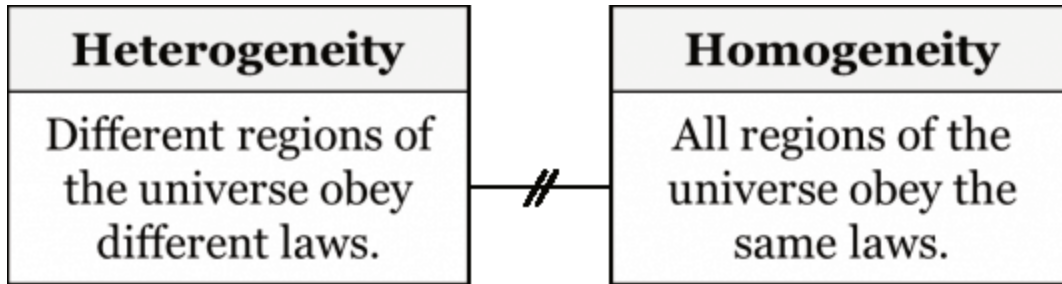
As we did in chapter 8, here we're going to cover not only the individual theories of the Newtonian mosaic, but also the *metaphysical elements* underlying these theories. Since any metaphysical element is best understood when paired with its opposite elements (e.g. hylomorphism vs. mechanism, pluralism vs. dualism, etc.), we will also be introducing those elements of the Aristotelian-Medieval and Cartesian worldviews which the Newtonian worldview opposed.

Recall from chapter 7 that, in their accepted cosmology, Aristotelians separated the universe into two regions – terrestrial and celestial. They believe that, in the terrestrial region, there are four elements – earth, water, air, and fire – that move *linearly* either towards or away from the centre of the universe. The celestial region, on the other hand, is composed of a single element – aether – which moves *circularly* around the centre of the universe. Since Aristotelians believed that terrestrial and celestial objects behave differently, we say that Aristotelians accepted the metaphysical principle of *heterogeneity*, that the terrestrial and celestial regions were fundamentally different.

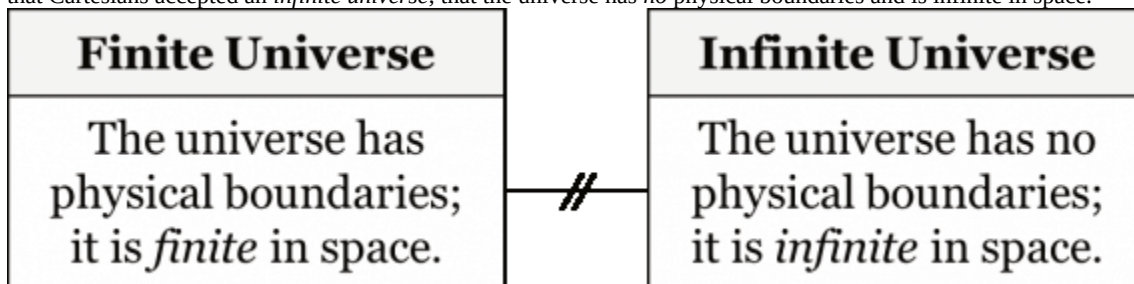
Additionally, Aristotelians posited that the celestial region was organized in a series of concentric spheres – something like a Matryoshka, or Russian nesting doll – with each planet nested in a spherical shell. The outermost sphere was considered the sphere of the stars, which was believed to be the physical boundary of the universe. According to Aristotelians, there is nothing beyond that sphere, not even empty space. Thus, they also accepted that the universe is *finite*.



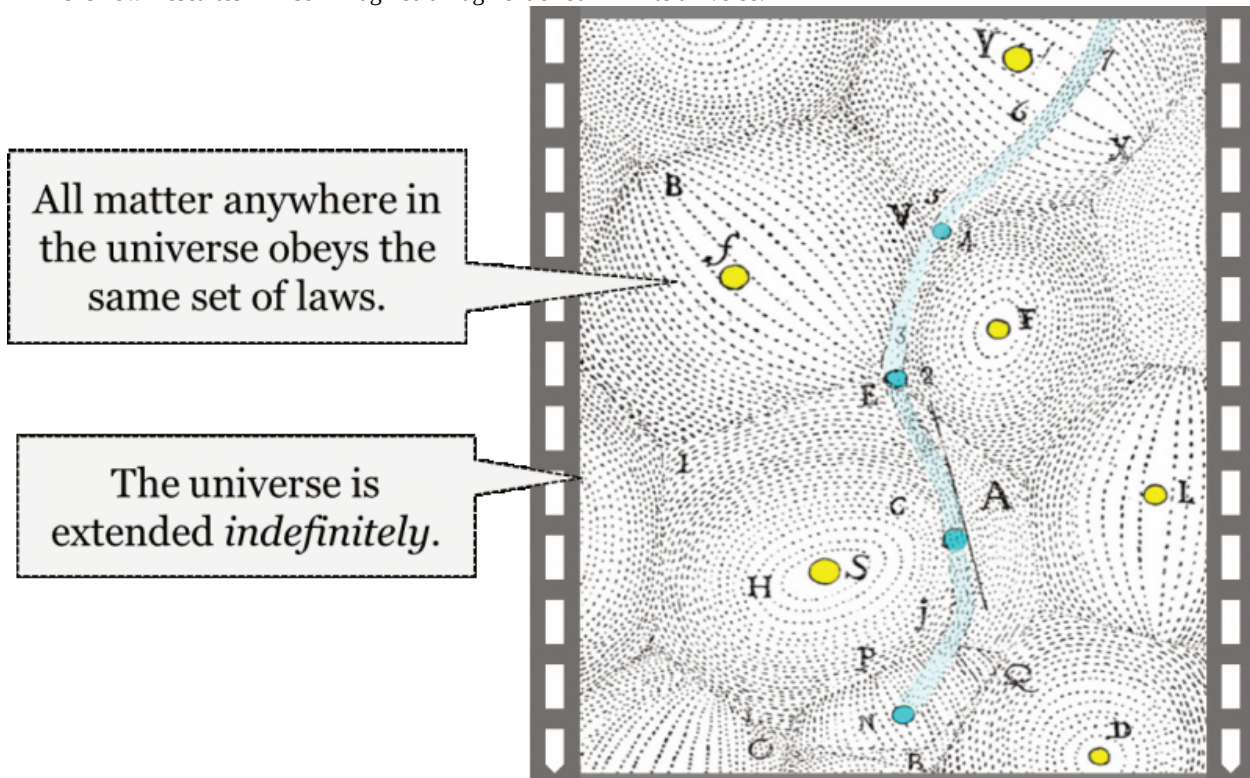
Cartesians rejected the Aristotelians idea of heterogeneity of the two regions as well as their idea of a finite universe. First, let's recall one of the central tenets of the Cartesian worldview: the principal attribute of all matter is extension. For Cartesians, it makes no difference whether that is the tangible matter of the Earth or the invisible matter of a stellar vortex – it must always be extended, i.e. occupy space. Since all matter, both terrestrial and celestial, is just an extended substance, the same set of physical laws applied anywhere in the universe. That is, Cartesians accepted the *homogeneity* of the laws of nature, that all regions of the universe obey the same laws.



Additionally, if extension is merely an attribute of matter, i.e. if space cannot exist independently of matter, then a question emerges: what would a Cartesian imagine existing beyond any boundary? Surely, they would never imagine an edge to the universe followed by empty space and nothingness – that would violate their belief in plenism. Instead beyond every seeming boundary is simply *more extended matter*, be it spatial matter or the matter of other planetary systems. Descartes would say that the universe extends ‘indefinitely’, meaning potentially infinitely, because he could *imagine* (but be uncertain about) a lack of boundaries to the edge of the universe and because he reserved the true idea of infiniteness (rather than indefiniteness) for God. So, we would say that Cartesians accepted an *infinite universe*, that the universe has *no* physical boundaries and is infinite in space.



This is how Descartes’ himself imagined a fragment of our infinite universe:



All matter anywhere in the universe obeys the same set of laws.

The universe is extended *indefinitely*.

The drawing shows a number of stars (in yellow) with their respective stellar vortices, as well as a comet (in light blue) wandering from one vortex to another.

What about the Newtonian attitude towards the scope of the laws of nature and the boundaries of the universe? Let’s start with the Newtonian view on the boundaries of the universe. While Cartesians accepted that space was an *attribute* of matter, i.e. that

it is indispensable from matter, Newtonians accepted quite the opposite: that space can and does exist *independently* from matter. For Newtonians, space is like a stage on which the entire material universe is built. But space can also exist *without* that material universe. This idea is known as the conception of *absolute space*: space is independent of material objects; it is an empty receptacle in which physical processes take place.

Bearing in mind that Newtonians accepted the existence of absolute space, then it remained possible, from the Newtonian point of view, for this absolute space to exist *beyond* any perceived boundary of the universe. Effectively, such boundaries would not even exist in the Newtonian worldview; space is essentially a giant void filled with solar system after solar system. If space is a void, then the universe must be infinite. Therefore, we say that Newtonians accepted the metaphysical idea of an *infinite universe*.

What about the laws of nature in the Newtonian worldview? Newton introduced three laws of motion as well as the law of universal gravitation to describe physical processes. Let's go over these laws and see what they suggest about homogeneity or heterogeneity. First, consider *Newton's second law*, which states:

The acceleration (a) of a body is directly proportional to the net force (F) acting on the body and is inversely proportional to the mass (m) of the body:

$$\mathbf{a} = \frac{\mathbf{F}}{m}$$

The law is also often stated as $F = ma$. To understand what this law states, imagine we used the law to describe an arrow being launched from a bow. A Newtonian would say the *acceleration* of the arrow after being launched from the bow would depend on the *mass* of the arrow, as well as the *force* of the bowstring pushing the arrow.

Now, what would happen to this arrow if there were *no* additional forces applied to it after being launched? For Newtonians, the answer is given by *Newton's first law*, or *the law of inertia*, which states that:

If an object experiences no net force, then the velocity of the object is constant: it is either at rest or moves in a straight line with constant speed.

So, after being launched but while remaining subject to no additional forces, the arrow would just keep moving in a straight line because of inertia. In other words, an object will remain at rest or in constant motion until some external force acts upon it. In reality, projectiles do not fly in a vacuum, for other than inertia, they are also subject to gravity. Newton accounted for the falling of objects by a force of mutual gravitational attraction between every pair of objects. His *law of universal gravitation* states that any two bodies attract each other with a force proportional to the product of their masses and inversely proportional to the square of the distance between them.

Any two bodies attract each other with a force (F) proportional to the product of their masses (m_1 and m_2) and inversely proportional to the square of the distance (r) between them:

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

To apply Newton's theory to the flying arrow, a Newtonian would need to know the *distance* between the arrow and the centre of the Earth, as well as the *masses* of the arrow and the Earth. Knowing these two values, they could calculate what the *force of gravity* between the arrow and the Earth is. Although gravity is a force of *mutual* attraction, the force pulling the Earth towards the arrow is immeasurably tiny compared to the force pulling the arrow towards the Earth. This is because the mass of the Earth is vastly greater than that of the arrow.

Finally, we have *Newton's third law*, which states that:

If one body exerts a force on a second body, then the latter simultaneously exerts a force on the former, and the two forces are equal and opposite.

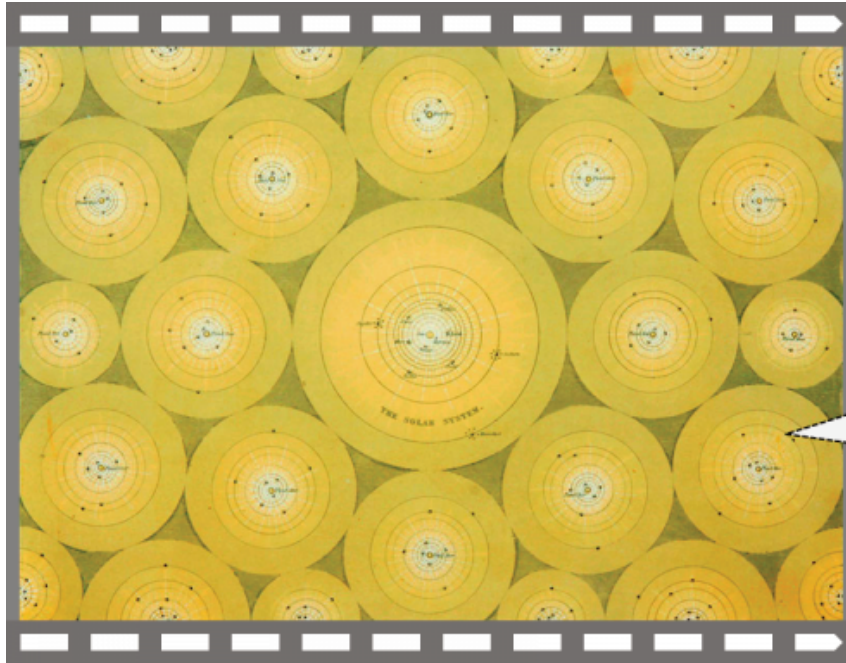
In other words, the third law is the law of equal and opposite reactions. So, every time an object interacts with another object – either by colliding with it, or by exerting some attractive force on it – the other object will experience the same force, but in the opposite direction. Thus, as the bowstring exerts its propulsive force on the arrow, the arrow exerts an equal and opposite force on the bowstring.

The two main factors a Newtonian would have to consider determining a flying arrow's trajectory are *inertia* and *gravity*. Were the arrow launched at a forty-five degree angle, the Newtonian would explain the forward and upward motion of the arrow, after leaving the bow, as due to its inertia. They would explain the fact that the arrow does not move in a straight line at a forty-five degree angle to the ground as due to the action of the force of gravity, which bends the arrow's trajectory by pulling it towards the surface of the Earth. The resulting motion would be due to both inertia and gravitational force. The angle of ascent would decrease and turn into an angle of descent, and the arrow's trajectory would be a *parabola*.

In a thought experiment in his *Treatise of the System of the World* Newton imagined he had a fantastically powerful cannon, on top of an imaginary mountain so high that the force of air resistance on a cannonball would be negligible. If he fired a cannonball from this super cannon, it would hurtle forward due to inertia, but also fall toward Earth due to the force of gravity, eventually crashing into Earth's surface. The faster the cannonball left the cannon barrel, the further it would travel before crashing to Earth. Newton realized that if a cannonball were fired fast enough, its fall due to the force of gravity would be at the same rate as the Earth

curved away beneath it. Rather than crashing to Earth, it would continue to circle the globe forever, just as the moon circles the Earth in its orbit. The orbital path it would follow would be circular or, more generally, *elliptical*; an oval shape with an off-centre Earth. If the cannonball were fired faster than a certain critical velocity, called the *escape velocity*, a Newtonian could calculate that it would escape from the Earth's gravitational pull entirely, hurtling away into outer space, never to return.

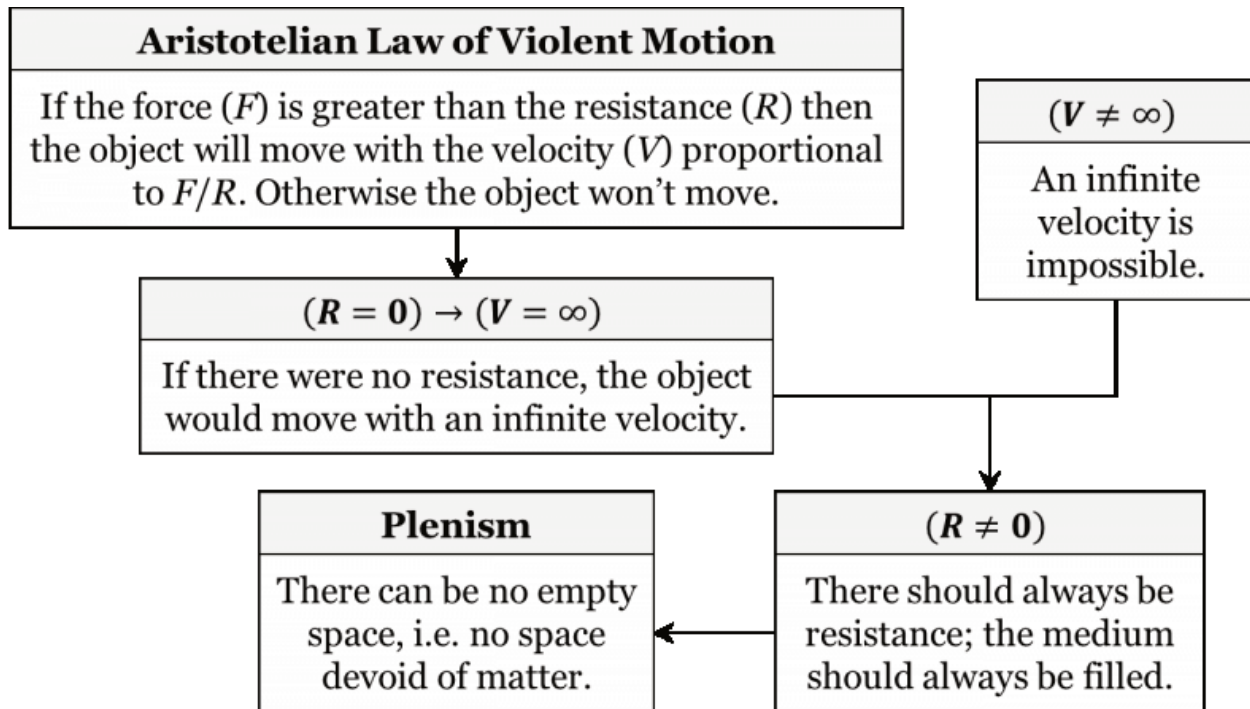
The conclusion we can draw from these examples is that the same laws that govern a projectile here on Earth must also govern a projectile in the heavens, as well as the motion of the planets and the stars. In other words, Newtonians accepted that the same laws of nature applied in the terrestrial regions of the universe as in the celestial. For this reason, they abandoned the distinction between the two regions that characterized the Aristotelian-Medieval worldview, and instead accepted the principle of *homogeneity* of the laws of nature. That is, in addition to the idea of infinite universe, they also accepted that all regions of the universe obey the same laws.



The universe is both homogenous and infinite.

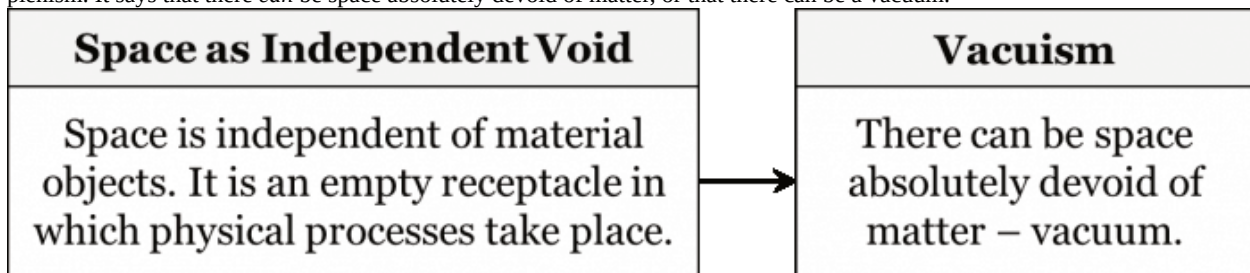
By this point in the chapter, it is hopefully evident that the underlying assumptions of the Newtonian worldview are vastly different from those of the Aristotelian-Medieval worldview. It might also *seem* as though Newtonians shared many assumptions with Cartesians – which they did. But the rival Cartesian and Newtonian communities also saw stark contrasts in the basic characteristics of their worldviews. In some ways, Cartesians shared more with Aristotelians than Newtonians.

Let us consider the idea of *absolute space*. Neither Aristotelians nor Cartesians would ever accept such an idea, for it conflicted with some basic assumptions of their worldviews. First, recall the *Aristotelian law of violent motion*: if force (F) is greater than resistance (R), then the object will move with a velocity (V) proportional to F/R ; if resistance (R) is greater than force (F), then the object will stay put. According to this law, should a moving object experience *no* resistance, the formula calls for us to divide by 0, leaving the object with an infinite velocity. It's tough to imagine what infinite velocity would look like in the real world, but we might imagine it as something like instant teleportation or being in two or more places at once. Aristotelians recognized the absurdity of an infinite velocity, and accordingly denied it was even possible in the first place. It followed from the impossibility of an infinite velocity that *some* resistance is always necessary in any motion. For our purposes, what this means is that, for Aristotelians, the universe is always filled with something that creates resistance. Thus, implicit in the Aristotelian-Medieval worldview was the idea of *plenism*, that there can be no empty space, i.e. no space devoid of matter.



Cartesians, as we know from chapter 8, also accepted plenism, though they justified it in a very different way. Since extension, according to Cartesians, is the principal attribute of matter, and since no attribute can exist without a substance, extension too cannot exist on its own. Extension, according to Cartesians, is always attached to something extended, i.e. to material things. Thus, there is no space without matter. The idea of plenism was one of the essential metaphysical assumptions of the Cartesian worldview.

In contrast, Newtonians rejected plenism. Recall that, in the *Principia*, Newton introduced and defended the idea of *absolute space* – the idea of space as independent from material objects. This implies *vacuism*, which is quite simply, the exact opposite of plenism. It says that there *can* be space absolutely devoid of matter, or that there can be a vacuum.



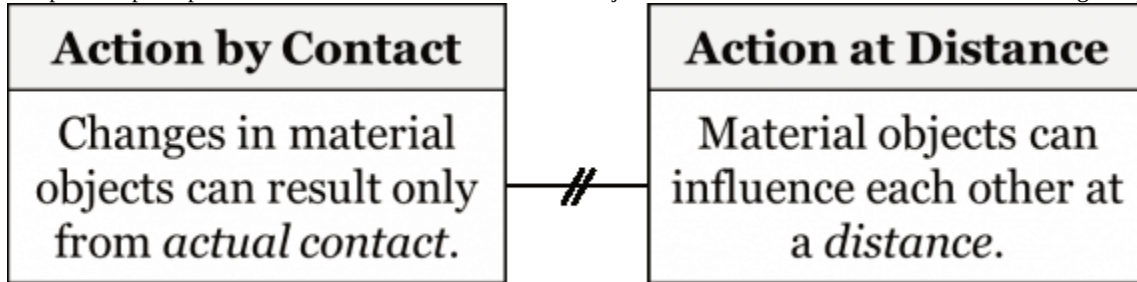
That said, why might Newton have introduced the idea of absolute space to begin with? Our historical hypothesis is that, at the time Newton was writing the *Principia*, scientists across Europe were conducting experiments that seemed to suggest the existence of a vacuum. These included barometric experiments conducted by Evangelista Torricelli and Blaise Pascal. Because the idea of a vacuum contradicted the then-accepted Aristotelian-Medieval idea of plenism, it could only, at best, be seen as a pursued theory at the time. Newton, however, seemed to have taken the results of these experiments seriously and developed a physical theory that could account for the possibility, or even actuality, of empty space.

Let's focus on one such experiment in more detail: the *Magdeburghemispheres*. After the barometric experiments of Torricelli and Pascal, Otto von Guericke, mayor of the German town of Magdeburg, invented a device that could pump the air out of a sealed space, effectively, he claimed, creating a vacuum. Von Guericke's device consisted of two semi-spherical shells or hemispheres, that, when placed together and emptied of air, would remain sealed by both the vacuum within and the air pressure without. So powerful was the vacuum within the sealed hemispheres that, reportedly, two sets of horses could not pry the device apart. Were the universe a plenum, it should have been impossible to create a vacuum inside the device and the horses would have been easily able to displace the halves in different directions. Since this was not the case, it seemed that pumping air out of the device left a vacuum inside of it. The pressure exerted by the air outside the two halves of the device was not balanced by the pressure of any matter within the device making the hemispheres extremely difficult to pull apart.

It is because of experiments like those of von Guericke that Newton seems to have been inspired to base his physics on the idea of absolute space. Effectively, he built his theory upon a new assumption, that space is not an attribute of matter, but rather an

independent void or a receptacle that can, but need not, hold any kind of matter. It is furthermore the reason that *vacuism* was an important element of the Newtonian worldview.

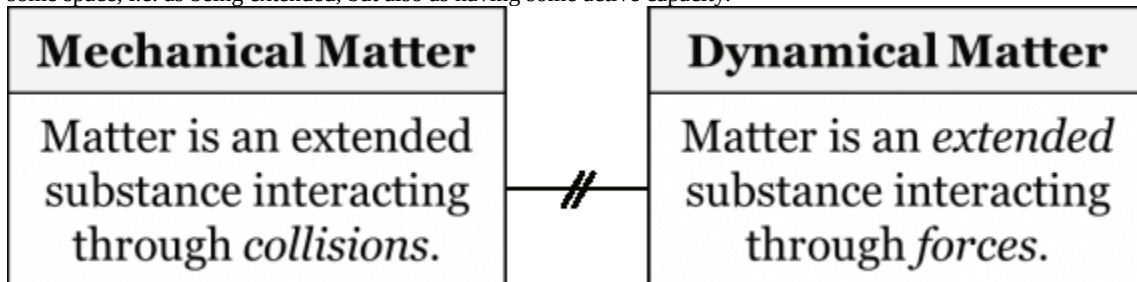
There are other important metaphysical elements that separate the Cartesian worldview from the Newtonian. Recall that Cartesians accepted the principle of *action by contact* – under which material particles can only interact by colliding with one another. Also recall that *Descartes' first law of motion* – that every material body maintains its state of motion or rest unless a collision with another body changes that state – follows from action by contact. Newton, as well, had a first law of motion. Phrased more along the lines of Descartes' first law, *Newton's first law of motion* says that every material body maintains its state of motion or rest unless it is caused to change its state by some force. The key difference between the two first laws is that the Newtonian worldview allows changes to result from the influence of *forces*, while in the Cartesian worldview changes can only result from the actual contact of material bodies. The clearest example of a force is probably the force of gravity. It is entirely possible in the Newtonian worldview for two objects, like the Moon and the Earth, to be gravitationally attracted to one another without any intermediary objects, like the bits of a matter of a vortex. Essentially, this means that in place of action by contact, Newtonians accepted the principle of *action at a distance* – that material objects can influence each other at a distance through empty space.



There is an important clarification to make about action at a distance. Accepting the *possibility* of action at a distance does not necessitate that *all* objects interact at a distance. For instance, were a Newtonian to observe a football (or soccer, if that's what you think the sport is called) player dribbling a ball down a field with their feet, they would not assume that there is some kind of contactless force keeping the ball with the moving player. Rather, they would explain that the ball is being moved by the player's feet contacting the ball and pushing it down the field. Newtonians continued to accept that many objects interact by contact. But they also accepted the idea that objects can influence one another across empty space through forces.

In addition to action by contact, we also know that Cartesians accepted the principle of *mechanicism* – that all material objects are extended substances composed of bits of interacting matter. So, for Cartesians all *seeming* instances of action at a distance, like the revolution of the Moon around the Earth or the Earth around the Sun, must actually be the result of colliding particles – the matter of the terrestrial and solar vortices, in these cases. Effectively, the absence of forces in the Cartesian worldview, and the fact that the source of all motion is external to any particular piece of matter, i.e. caused by its collision with another piece of matter, mean that all matter is *inert*. In other words, material things do not have any capacity to influence other things without actually touching them.

Newtonians conceived of matter in a new way. While Cartesian mechanical matter was inert, Newtonian matter was active and dynamic. It not only occupied space and maintained its state of rest or motion unless compelled to change it by an external force, it also had an active capacity to exert force on other bodies from a distance. Newton's law of gravitation tells us that any two objects are gravitationally attracted to one another. Effectively, Newtonians replaced the Cartesian conception of mechanicism with *dynamism*, the idea of matter as an extended substance interacting through forces. Thus, they saw all matter as not only occupying some space, i.e. as being extended, but also as having some active capacity.



These new ideas were actually troubling to Newton himself. Newton disliked that his theory of universal gravitation suggested that objects have a mysterious, almost magical, ability to interact from a distance. Action at a distance and dynamic matter were, for Newton, occult ideas. Mathematically, the law of gravity worked, and it allowed one to predict and explain a wide range of terrestrial and celestial phenomena. But because Newton was educated in a mechanistic tradition, he initially believed that proper explanations should be mechanistic, i.e. should involve physical contact. In fact, he searched for, but ultimately failed to provide, a mechanical explanation for gravity. In other words, it seems as though Newton himself likely accepted mechanicism. However, because of his failure to provide such a mechanical explanation, and the many successes of his theory, the Newtonian community

eventually accepted the notion of the force of gravity as acting at a distance. Furthermore, the strong implication that action at a distance remained the most reasonable explanation for the motion of the planets and the stars led the Newtonian community to accept that matter must be dynamic.

We bring up the conflicting perspectives of Newton and the Newtonians to emphasize the importance of distinguishing between individual and communal beliefs when studying the history of science. On the one hand, we can write fascinating intellectual biographies of great scientists such as Newton. But when we write such histories – histories of individuals – we risk misrepresenting the accepted worldviews of the communities in which those great individuals worked. In such a case, we trade a focus on the discoveries and inventions of individuals – what we would call newly pursued theories – for a proper reconstruction of the belief system of the time. If we write our histories from the perspective of the community, we can understand these individuals in their proper context. We can better realize not only how novel their ideas were but also what the response of the community was and at what point, if ever, their proposed theories became accepted. In sum, it is only by distinguishing the history of the individual from the history of the community that we can realize that Newton's personal views on matter and motion did not necessarily align with that of the Newtonians; he most likely accepted mechanicism, while Newtonians clearly accepted dynamism.

The dynamic conception of matter underlies more than just the cosmology and physics of the Newtonian mosaic. We see dynamism implicit in the theories of other fields as well. For instance, Newtonians rejected the Cartesian idea of corkscrew particles to explain magnetism, accepting the idea of a *magnetic force* in its place. In chemistry, Newtonians accepted that the chemicals they believed to exist – like mercury, lead, silver, and gold – combine and react more or less effectively because of something they called a *chemical affinity*. Chemical affinity was interpreted as an active capacity inherent in different chemical substances which caused some to combine with others in a way that had clear parallels to the Newtonian conception of gravity. For example, following numerous experiments and observations, they concluded that mercury would combine better with gold than silver, and they explained this in terms of mercury's strong chemical affinity to gold. Even in physiology, Newtonians posited and accepted the existence of a *vital force* that brought organisms to life.

Physiology

In the Cartesian worldview, the accepted physiological theories were *mechanistic*. That is, Cartesians saw human bodies – indeed, all living organisms – as complex machines of interconnected and moving parts. Though they were uncertain how the mind commands the body to operate, they were confident that all biological processes acted mechanistically through actual contact, similar to a clock with its various gears and cogs.

The Newtonian response to mechanistic physiology was known as *vitalism*. In the first few decades of the eighteenth century, physicians found themselves asking about what properties are essential to life. Mechanicists would probably answer that living organisms were carefully organized combinations of bits of extended matter, much like the carefully organized gears, wheels, and pendulum of a clock. But by the mid-to-late eighteenth century, the medical community began observing phenomena that were anomalous for mechanistic physiology. One observation concerned an animal's ability to preserve its body heat, even when the circulation of its blood was stopped. Mechanicists posited that heat is generated by circulating blood, and they could not provide a satisfactory explanation for why heat continued to be generated in the absence of this mechanical cause. Another observation concerned the temperature of a dog's nose. It was noted that a dog's nose is filled with blood and should thus be warm like the rest of its body, and yet most often a dog's nose is as cold as the temperature of the air around the dog. Why did the temperature of the rest of the dog's body not cool to the temperature of the air around it? It seemed that mechanicists could not produce a satisfactory answer to such questions. Vitalists, on the other hand, posited that there was some additional force inherent in living things which, in this case, regulates an animal's body heat. By the late 1700's, vitalist physiology and the idea of a vital force had replaced mechanistic physiology as the accepted physiological theory of the time. In essence, *vitalism* suggested that living matter is organized by an inherent vital force.

Newtonians saw vital forces as the living principles responsible for maintaining health and curing illness. Physicians generally characterized an organism's vital force by two properties, sensibility and contractility. *Sensibility* involved what the different parts of your body could feel. It included both voluntary properties that allowed you to use your senses to interact with your environment, and involuntary properties, like feeling hunger or maintaining a sense of balance. *Contractility* involved how the different parts of your body moved. It was sometimes an involuntary property that ensured the beating of your heart or the digestion of food, and sometimes a voluntary property involved with things like locomotion. In essence, vitalists accepted that organisms would lack both sensibility and contractility in the absence of a vital force.

Accordingly, vitalists suggested that illness and disease derived from damage to one of these vital properties. For instance, vitalists believed proper digestion was a *contractile* property directed by a vital force. Were a person to catch the flu, vitalists suggested that the vital force directing proper digestion was somehow interrupted. In effect, that person's digestive system would not function properly, as demonstrated by flu-like symptoms such as vomiting. Alternatively, consider a *sensible* property guided by a vital force, like hunger or thirst. Falling ill manifests itself not only in changes in existing sensible properties, like a loss of appetite, but sometimes also in the addition of new, unwanted sensations, like itching or tingling.

The treatment of illness, for vitalists, was about administering medicine that activates the vital forces of the body in such a way as to accelerate healing by the proper functioning of these vital properties. Treatment did not always involve straightforward activation of one of these properties. For instance, a physician would avoid giving medicine that heightens contractility to a person

suffering from convulsions – a symptom, the involuntary contraction of the muscles, associated with an unwanted increase in the contractible property.

The vitalist conceptions of illness and treatment were in sharp contrast with those of both Aristotelians and Cartesians. As explained in earlier chapters, for Aristotelians disease was a result of an imbalance of bodily fluids or humors. Consequently, treatment involved the rebalancing of these humors. For Cartesians, the ideas of disease and curing remained largely the same as those of Aristotelians, despite the fact that in Cartesian physiology, humors received a purely mechanistic interpretation. Conversely, vitalists at the time of the Newtonian worldview didn't believe curing was about the mechanical balancing of humors; it was about restoring the vital forces that help maintain a properly functioning body.

Similarly, to the force of gravity, vital force was seen as a *property* of a material body, as something that doesn't exist independently of the body. Importantly, Newtonians did *not* see vital force as a separate immaterial substance. This is true about the dynamic conception of matter in general: the behaviour of material objects is guided by forces that are *inherent* in matter itself.

Although Cartesians and Newtonians had different conceptions of matter – mechanical and dynamic respectively – they agreed that matter and mind can exist independently of each other. Both Cartesians and Newtonians accepted dualism, the idea that there are two independent substances: matter and mind.

In addition to purely material and purely spiritual entities, both parties would agree that there are also entities that are both material and spiritual. Specifically, Cartesians and Newtonians would agree that human beings are the only citizens of two worlds. However, some alternatives to this view were pursued at the time. For instance, some philosophers believed that animals and plants were composed of not only matter, but also mind. They believed this because they saw all living organisms as having inherent organizing principles – minds, souls, or spirits – which are essentially non-material. Others denied that humans are composed of any mind, or spiritual substance, at all; they saw humans, along with animals, plants, and rocks, as entirely material, while only angels and God were composed of a mental, spiritual substance. However, albeit pursued, these alternative views remained unaccepted. The position implicit in the Newtonian worldview was that only humans are composed of both mind and matter.

This dualistic position was very much in accord with another important puzzle piece of the Newtonian mosaic – theology. Different Newtonian communities accepted different theologies. In Europe alone, there would be a number of different mosaics: Catholic Newtonian, Orthodox Newtonian, Lutheran Newtonian, Anglican Newtonian etc. Yet, the theologies accepted in all of these mosaics assumed that the spiritual world can exist independently of the material world and that matter and mind are separate substances. So, it is not surprising that dualism was accepted in all of these mosaics.

Theology

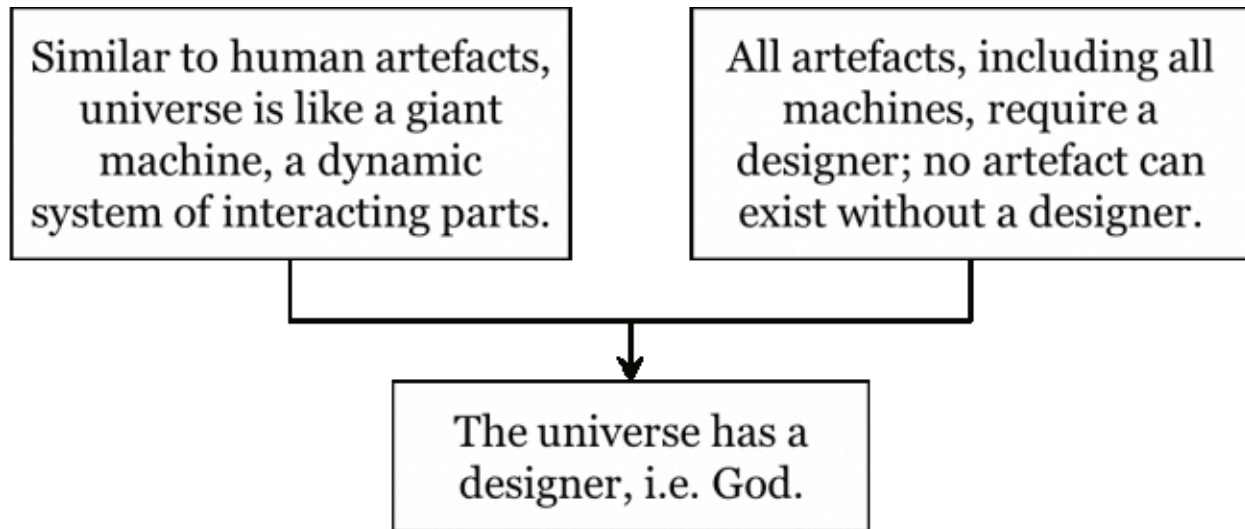
Theology, or the study of God and relations between God, humankind, and the universe, held an important place in the Newtonian worldview. Theologians and natural philosophers alike were concerned with revealing the attributes of God as well as finding proofs of his existence. Nowadays, these theological questions strike us as *non-scientific*. But in the eighteenth and nineteenth centuries, they remained legitimate topics of scientific study.

Just like Aristotelians and Cartesians before them, Newtonians accepted that there were two distinct branches of theology: revealed theology and natural theology. *Revealed theology* was concerned with inferring God's existence and attributes – what he can and cannot do – exclusively from his acts of self-revelation. Most commonly for Newtonians, revealed theology meant that God revealed knowledge about himself and the natural world through a holy text like the Bible. But revelation also occurred in the form of a supernatural entity like a saint, an angel, or God himself speaking to a mortal person, or through a genuine miracle like the curing of some untreatable illness.

It wasn't uncommon for a natural philosopher of the time to practice revealed theology. Newton himself interpreted many passages from the Bible as evidence of various prophecies. For instance, he believed that a passage from the Book of Revelations indicated that the reign of the Catholic Church would only last for 1260 years. But he was never certain on which year the reign of the Catholic Church had actually begun, and so he came up with multiple dates to mark the fulfilment of the prophecy of 1260 years. Nevertheless, Newton's belief in this prophecy stems from his reading of the Bible, i.e. it stems from his practice of and belief in revealed theology.

In contrast, *natural theology* was the branch of theology concerned with inferring God's existence and attributes by means of reason unaided by God's acts of self-revelation. Philosophers were practicing natural theology when they made arguments about God with reason and logic. Descartes' ontological argument for the existence of God from chapter 8 is an example of a theory in natural theology. Others would practice natural theology by studying God through the natural world around them. In any case, what characterizes natural theology is that conclusions regarding God, his attributes, and works were drawn without any reference to a holy text.

Let's consider one formulation of the famous *argument from design* for God's existence. The argument goes like this. On the one hand, the universe seems like a great big machine, a dynamic system of interacting parts. It is, in a sense, analogous to human artefacts; it is akin to a very complex clock, where all the bits and pieces work in a perfect harmony. On the other hand, we know that artefacts including machines have a designer. Therefore, so the argument goes, the universe also has a designer, i.e. God:



In essence, the argument from design assumes an analogy between the universe as a whole and a human artefact, such as a steam engine, a mercury thermometer, or a marine chronometer. Since such artefacts are the product of design by a higher entity (i.e. humans), then perhaps the universe itself, with the planets and stars moving about in the heavens, is the product of design by some even *higher* entity – God. The argument fits under the category of natural theology because it is based on a certain understanding of nature, i.e. the idea that the universe is a dynamic system of interacting parts operating through collisions and forces.

The argument from design was far from perfect. The eighteenth-century philosopher David Hume remained unconvinced by it and pointed out some of its major problems. First, Hume rejected the premise that there is an analogy between the universe as a machine and artefacts, making the entire argument unsound. He noted that the reason we claim artefacts have a designer is that we have experienced humanity designing artefacts from initial concept to final product. But when it comes to the universe as a whole, Hume reasoned, we have never experienced such a conceptual stage. That is, no one has ever seen an all-powerful being create a universe; we’re merely living in an already operational “machine”. So, while artefacts clearly have a designer, this doesn’t imply that the universe has a designer.

Second, Hume points out that this argument for the existence of God says nothing about what God is like. Even if we were to accept the argument, the only conclusion that would logically follow from it is that the universe has *some* designer. Importantly, it wouldn’t imply that this designer is necessarily the omnipotent, omniscient, and omnibenevolent God of the Christian religion. There is nothing in the argument to preclude an imperfect God from designing the universe, or even multiple gods from designing it. To accept an imperfect God, or the existence of multiple Gods, would be *incompatible* with the then-accepted Christian beliefs concerning an all-perfect God.

Regardless of Hume’s criticism, Newtonians accepted some form of the argument from design for the existence of God. More generally, Newtonians accepted both *revealed theology*, or the study of God through his acts of self-revelation, and *natural theology*, or the study of God by examining the universe he created.

Astrology

While theology was an essential part of the Newtonian mosaic, *astrology*, the study of celestial influences upon terrestrial events, suffered a different fate. Newtonians understood that the stars and the planets exerted some kind of influence upon the Earth. However, in the Newtonian worldview, the astrological topics that were accepted in the Aristotelian worldview were either gradually encompassed by other fields of natural science or rejected altogether.

Traditionally, astrology was divided into two branches – judicial astrology and natural astrology. *Judicial astrology* was the branch of astrology concerned with celestial influences upon human affairs. For instance, consulting the heavens to advise a monarch on when to go to war, or when to conceive an heir fell within the domain of judicial astrology. Judicial astrology could also be involved in something as innocent as reading a horoscope to decide when to ask a crush’s hand in marriage. In all of these examples, it was suggested that the influence of the heavens extended beyond the material world to the human mind.

The other branch of astrology was *natural astrology*, and it was concerned with celestial influences upon natural things. For instance, positing a link between the rising and setting of the Moon and the ebb and flow of the tide fell under natural astrology. Similarly, any study of light and heat coming from the Sun and affecting the Earth was considered part of natural astrology. The measuring of time and the forecasting of weather by studying planetary positions would also pertain to natural astrology. Medical prognostications using natal horoscopes would equally belong to this field.

A question arises: why weren’t medical prognostications using natal horoscopes a part of *judicial* astrology; didn’t they concern heavenly influences upon *humans*? To answer this question, we need to recall the distinction between the mind and the body. You

would be right to recall that physicians would require knowledge of the heavens – specifically of a patient’s natal horoscope – in order to properly rebalance a patient’s humors. So, it might seem as though physicians were studying celestial influences over human affairs and therefore practicing judicial astrology. But, more precisely, physicians would simply monitor and make changes to a patient’s *body*; they would not be concerned with any celestial influences over a patient’s *mind*. Medical prognostication did not fall under judicial astrology because physicians accepted that the celestial realm can and does influence the *material* world by bringing humors in and out of balance. They did not accept, as claimed by judicial astrologers, that the heavens could determine what would normally be determined by a person’s mind or by an even more powerful agent, God.

Furthermore, the idea of judicial astrology was in conflict with the Christian belief in free will. According to one of the fundamental Christian dogmas, humans have an ability to act spontaneously and make decisions that are not predetermined by prior events. This goes against the key idea of judicial astrology: if celestial events do, in fact, determine the actions of humans, then in what sense can humans be said to possess free will? If a state of the human mind is determined by the position of stars and planets, then the very notion of human free will becomes questionable. It is not surprising, therefore, that the practice of judicial astrology was considered heresy and therefore banned. As such, judicial astrology was never an accepted part of either the Aristotelian-Medieval, Cartesian, or Newtonian mosaics.

Natural astrology, on the other hand, was an accepted element of the Aristotelian-Medieval worldview and many of its topics even persisted through the Cartesian and Newtonian worldviews. While in the Aristotelian-Medieval mosaic natural astrology was a separate element, it ceased to be so in the Cartesian and Newtonian mosaics; only *some* of its topics survived and were subsumed under other fields, such as astronomy, geology, meteorology, or physics. There were other topics of natural astrology that were simply rejected. Consider the following three questions:

How are tides influenced by celestial objects such as the Sun and the Moon?

How is the weather on Earth influenced by celestial phenomena?

How is human health influenced by the arrangement of celestial objects?

In the Aristotelian-Medieval worldview, all three of these questions were accepted as legitimate topics of study; they all pertained to the domain of natural astrology. In particular, Aristotelians accepted that there was a certain link between the Moon and the tides. They believed that the positions of planets influenced weather on Earth. For instance, a major flood might be explained by a conjunction of planets in the constellation of Aquarius. Finally, Aristotelian physicians believed that the positions of the planets affect the balance of humors in the body, and thus human health.

Of these three topics, only two survived in the Cartesian and Newtonian mosaics. Thus, both Cartesians and Newtonians accepted that the position of the Moon plays an important role in the ebb and flow of the tides. While they would not agree as to what actual mechanism causes this ebb and flow, they would all accept that a “celestial” influence upon the tides exists. Similarly, they would both agree that the weather on Earth might be affected by the Sun. However, the question of the celestial influence upon the human body was rejected in the Cartesian and Newtonian worldviews.

In short, while some traditional topics of natural astrology survived in the Cartesian and Newtonian worldviews, they didn’t do so under the label of natural astrology. Instead they were absorbed by other fields of science.

Newtonian Method

How did Newtonians evaluate the theories that would become a part of their worldview? If we were to ask this question to Newtonians themselves, especially in the eighteenth century, their *explicit* answer would be in accord with *the empiricist-inductivist methodology* of Locke and Newton. As mentioned in chapter 3, the empiricist-inductivist methodology prescribed that a theory is acceptable if it merely *inductively generalizes* from empirical results without postulating any *hypothetical* entities. However, the actual expectations – their *method* not methodology – of the Newtonian community was different. When we study the actual eighteenth-century transitions in accepted theories, it becomes apparent that the scientists of the time were willing to accept theories that postulated unobservable entities. Recall, for instance, the fluid theory of electricity that postulated the existence of an *electric fluid*, the theory of preformation that postulated invisibly small *homunculi* in men’s semen, or Newton’s theory that postulated the existence of absolute space, absolute time, and the force of gravity. In other words, the actual expectations of the community of the time, i.e. their methods, were different from their explicitly proclaimed methodological rules.

So how did Newtonians *actually* evaluate their theories? In fact, the same way as Cartesians: all theories in the Newtonian mosaic had to satisfy the requirements of *the hypothetico-deductive (HD) method* in order to be accepted. Indeed, the employment of this new method also had to follow *the third law of scientific change*: just as in the case of the Cartesian mosaic, the HD method was a logical consequence of some of the key metaphysical principles underlying the Newtonian worldview. These metaphysical principles would be the same in both mosaics, but they would be arrived at differently in Cartesian and Newtonian mosaics.

In previous chapters, we explained how the HD method became employed in the Cartesian mosaic because it followed from their belief that the principle attribute of matter is extension. Newtonians, on the other hand, had a slightly different understanding of matter. They believed in a *dynamical* conception of matter, that matter is an extended substance interacting through forces. We can draw two conclusions from the Newtonian belief in dynamic matter. First, the secondary qualities of matter, like taste, smell, and colour, result from the combination and dynamic interaction of material parts. Since these secondary qualities were taken as the products of a more fundamental inner mechanism, albeit one that allows for the influence of forces, Newtonians accepted the principle of *complexity*. Second, any phenomenon can be produced by an infinite number of different combinations of particles

interacting through collisions and forces. Accordingly, it is possible for many different, equally precise explanations to be given for any phenomenon after the fact. Thus, Newtonians also accepted the principle that *post-hoc explanations* should be distrusted, and novel otherwise unexpected predictions should be valued. If these conclusions seem similar to the conclusions Cartesians drew from their belief that matter is extension, it's because they are. The fact that Newtonians also accepted forces doesn't seem to have influenced their employment of the HD method. As per the third law of scientific change, the HD method becomes employed because it is a deductive consequence of Newtonians' belief in complexity and their mistrust for post-hoc explanations.

Summary

Let's summarize the many metaphysical conceptions we've uncovered in this chapter. While Aristotelians believed in the heterogeneity of the laws of nature and that the universe was finite, Cartesians and Newtonians alike believed the laws of nature to be *homogenous* and the universe to, in fact, be *infinite*. Newtonians also shared the Cartesian belief in *dualism*, that there are two substances in the world: mind and matter. At the same time, the metaphysical assumptions of the Newtonian worldview contrasted in many ways with both those of the Aristotelian-Medieval and the Cartesian worldviews. Newtonians replaced the idea of plenism with that of *vacuism* – that there can be empty space; they expanded their understanding of change and motion from action by contact to allowing for the possibility of *action at a distance*; and they modified the conception of matter from being mechanical and inert to being *active* and *dynamic*.

Aristotelian-Medieval		Cartesian		Newtonian
Finite Universe	#	Infinite Universe	—	Infinite Universe
Heterogeneity	#	Homogeneity	—	Homogeneity
Plenism	—	Plenism	#	Vacuism
Pluralism	#	Dualism	—	Dualism
Teleology	#	Action by Contact	#	Action at a Distance
Hylomorphism	#	Mechanicism	#	Dynamism
Aristotelian method	#	HD method	—	HD method

One point that we want to emphasize about all of these metaphysical elements insofar as the Newtonian worldview is concerned is that they weren't always explicitly discussed or taught in the eighteenth and nineteenth centuries. Rather, some of these assumptions are *implicit* elements of the worldview; they are ideas that, had we had a conversation with a Newtonian, we would expect them to agree with, for they all follow from their accepted theories. For instance, the reason Newtonians would accept the conception of dynamism is that their explicit acceptance of the law of gravity *implies* that matter can interact through forces.

Before concluding this chapter, there's another important point to re-emphasize about these historical chapters in general: all mosaics change in a *piecemeal* fashion. What this means is that the Newtonian worldview of, say, 1760, looked vastly different from the Newtonian worldview of 1900. While we've tried to describe the theories accepted by Newtonians and the metaphysical principles that characterized their worldview around the second half of the eighteenth century, this may create a false impression that the Newtonian mosaic did not change much after that. It actually did.

One notable change *within* the Newtonian worldview was a shift in the belief over how to conceive of matter, a shift from *dynamism* to the belief in *particles and waves*. Up until around the 1820s, Newtonians conceived of matter merely in terms of particles. These particles could interact via collisions and forces – hence the *dynamic* conception of matter – but they were always understood in a literally corpuscular sense. After the 1820s, however, Newtonians had experimentally observed matter behaving in a non-corpuscular way. Some matter had been observed acting more like a wave in a fluid medium than like a particle. The idea of dynamic matter consisting exclusively of particles interacting through collisions and forces came to be replaced by the idea of dynamic matter consisting of both particles *and waves* interacting through collisions and forces. It is possible that some Newtonians sought to replace the idea of a force with that of a wave, so that all forces pushing and pulling dynamic matter about

could be interpreted in terms of waves in a subtle fluid medium sometimes called the luminiferous ether. But not all forces at the time of the Newtonian worldview were explained in terms of waves, so the most that we can say about them is that they accepted that both particles and waves interacted through collisions and forces. Recall the discussion of Fresnel's wave theory of light from chapter 3.

Did the Newtonian conception of matter in terms of particles and waves persist in the twentieth century? That will be a topic of our next chapter, on the Contemporary worldview.

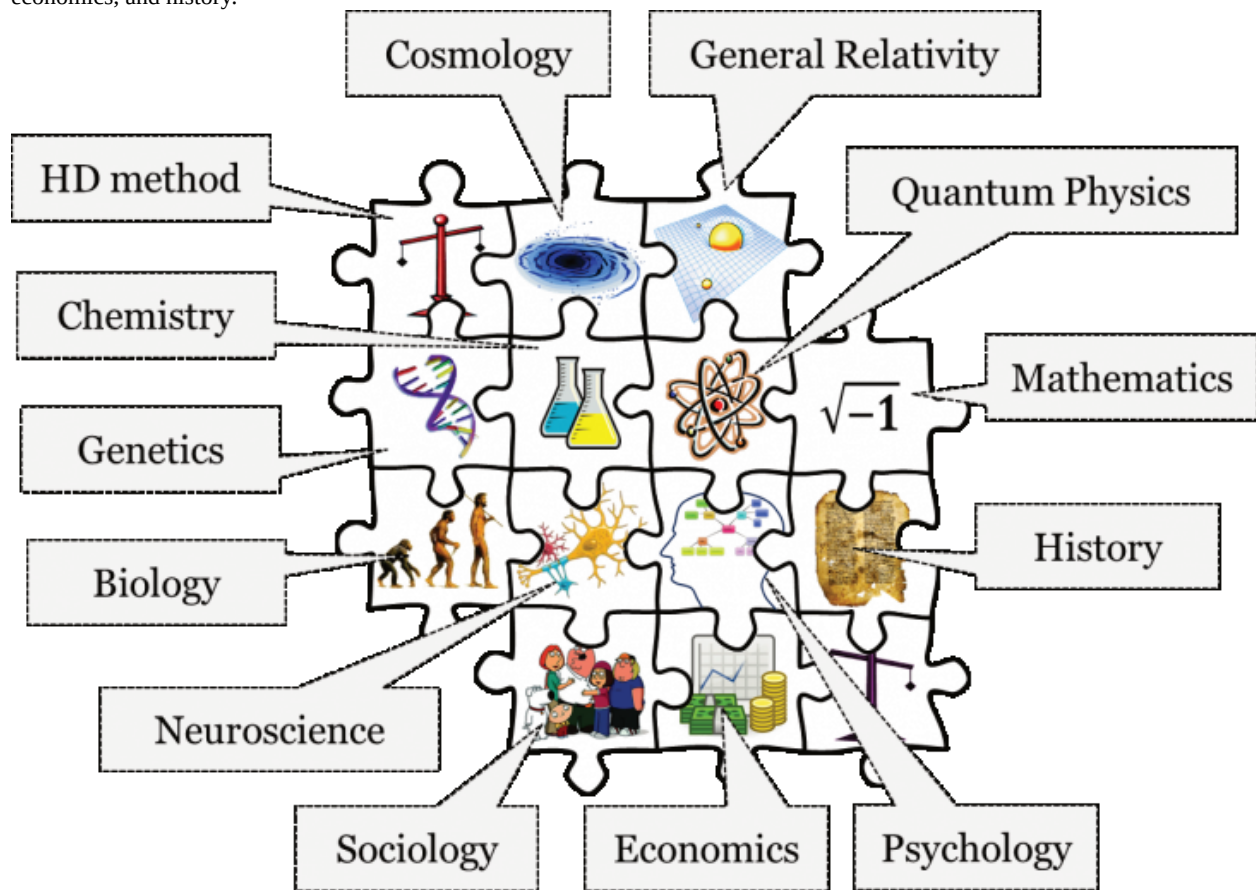
10.

Contemporary Worldview -- Introduction to History and Philosophy of Science

Chapter 10: Contemporary Worldview

Intro

The Contemporary mosaic will probably look more familiar to us than its predecessors. It includes, after all, the theories and methods taught at the very universities you're attending. We have borne the Contemporary mosaic since about the 1920s. It includes accepted theories like neuroscience, quantum mechanics, special and general relativity, and cosmology, which we will consider in more depth below, in addition to the theories of evolutionary biology, genetics, chemistry, psychology, sociology, economics, and history.



It is important to remember that, as with all previous mosaics, the Contemporary mosaic is in a continuous state of flux. In other words, the mosaic of 1920 is not identical to the mosaic of today. Furthermore, as we stressed in chapter 2, our current set of accepted theories and employed methods are in no way absolutely true. We do think that they're the best available descriptions of their respective objects, but we are also prepared to replace our best descriptions with even better ones. Consider, for instance, the acceptance of the Higgs boson, the mapping of the human genome, or the discovery of new pyramids in Egypt. So, although we sometimes treat our Contemporary worldview as unchangeable and something already crystallized, it is very much modifiable as we learn more about the world.


Let's start our focused discussion of the Contemporary worldview with the field of cognitive neuroscience and the metaphysical principles implicit in this discipline.

Cognitive Neuroscience

Neuroscience is the study of the nervous system and aims to better understand the functions of the brain. Cognitive science is the study of the mind and its processes. Cognitive neuroscience melds these two fields and seeks to understand the processes of the

mind as functions of the brain. The central metaphysical question here is whether the mind can exist independently of the brain or anything material, or whether the mind requires the existence of a material substrate to produce it. In the previous historical chapters, we've seen that Cartesians and Newtonians alike accepted the metaphysical conception of *dualism*, which holds that mind and matter are independent substances. Recall, for instance, how Cartesian mechanistic physiologists couldn't figure out an acceptable way for the mind to interact with matter. This problem grew more acute in the nineteenth century, when the law of conservation of energy became accepted. The law implies that the physical world is causally closed, and that it is therefore impossible for a non-physical mind to influence a physical brain. Given this constraint, and neuroscientific theories accepted within the Contemporary mosaic, substance dualism is no longer regarded as the best available description of the relationship between mind and matter.

Dualism is a response to a question about the number of distinct independent substances that actually exist. It answers that there are two independent substances. *Pluralism* – the metaphysical principle implicit in the Aristotelian-Medieval worldview – answers that there are many independent substances. Another possible response is that of *monism*: that there is only one single independent substance.

 How many distinct <i>substances</i> are there?		
One	Two	Many
Monism: There is only one substance. All things are essentially of one kind.	Dualism: There are two distinct substances. All things are essentially of two kinds.	Pluralism: There are many distinct substances, i.e. many distinct kinds of things.

Different varieties of monism have a different understanding of the nature of this single independent substance. Let's consider the various types of monism.

One type of monism is called *idealism*. Idealism tells us that there is only one independent substance, the mind. In other words, all material things in the world are simply the products or manifestations of the mind. Idealism has two major subtypes: subjective and objective idealism. *Subjective idealism* suggests that the existence of material objects depends on the perception of those objects by an individual mind, i.e. to be is to be perceived.

Essentially, subjective idealists claim that objects can be said to exist only as perceptions within an individual mind. To bolster this claim, they ask how one imagines an *unperceived* material object. That is, they ask what something independent of the mind (i.e. something unperceived) looks like. Answering this question, however, is impossible, because as soon as you try to imagine something, you're relying on a mind. Hence, according to subjective idealism, material objects cannot exist independently of the mind. Subjective idealism is most commonly tied to the Irish philosopher George Berkeley.

Subjective idealism is problematic because it's not quite clear why different individual minds perceive the same things. For instance, if two people are looking at the same object, they will usually have similar perceptions of that object. Therefore, there must exist some sort of mechanism that guarantees the coherence of perceptions among different individual minds.

To solve this problem, some philosophers postulated the existence of a universal mind, such as that of God, which guarantees that the perceptions of different individual minds will cohere. This step, however, takes us to *objective idealism*, which holds that everything depends on and exists within the mind of a single, objective perceiver.

Importantly, objective idealism doesn't deny that material objects exist outside of individual minds. But it does deny that the material world can exist outside of this single, universal mind. This subtype of idealism is associated with German philosopher George Wilhelm Friedrich Hegel. Often, objective idealists understand the material world as a manifestation of this objective, universal mind. In our modern understanding, this is similar to a computer simulation. Within this simulation, a couple of things happen. One, anyone's perception of the (simulated) material world would be similar to everyone else's. Two, anyone's experience of this material world depends on the program that runs the simulation, and not the individual minds within the simulation.



To be clear, idealism does not imply that the material world does *not* exist. Instead, idealism states that the material world cannot exist *without* the mind – either subjective or objective. In other words, mind is the only independent substance.

In contrast with idealism, there is another monist position known as *materialism*. Materialism also tells us that there is only one independent substance, but that this substance is *matter* – not mind. For materialists, all mental states, thoughts, emotions, memories, and perceptions are the products of some material process. Before modern times, it was very difficult to understand how this might be possible.

A third variety of monism is called *neutral monism*. According to neutral monism, there is only one independent substance, which is neither purely material nor purely ideal, but a neutral substance which is both material and ideal at the same time.

According to neutral monists, everything in the universe is both matter and mind: there is nothing purely material or purely ideal. In other words, everything material is said to have some mental capacities, albeit very minimal. Similarly, everything ideal is said to presuppose one material process or another. In this view, matter and mind are simply two sides of the same coin; nature itself is, in a sense, neutral. One traditional variety of neutral monism is the idea that nature (matter) and God (mind) are essentially one and the same: the divine mind doesn't exist outside of nature, and nature doesn't exist independently of the divine mind. One of the most notable champions of this version of neutral monism is the philosopher Baruch Spinoza. A similar view was endorsed by the physicist Albert Einstein. An even more contemporary version of neutral monism suggests that every bit of matter contains within itself information regarding its current state, as well as its past states and, probably, its possible future states. But information, in this view, is something ideal – it is the data “stored” in nature and “processed” by nature. All of nature, in this view, becomes a huge computer that processes vast amounts of information every second to decide what happens to every bit of matter.

So, dualism, idealism, materialism and neutral monism are our options regarding the relationship between mind and matter. The four conceptions are summed up in the following table:

		 Is <i>mind</i> a substance? Can it exist without anything <i>material</i>?	
		Yes	No
 Is <i>matter</i> a substance? Can it exist without <i>mind</i>?	Yes	Dualism: There are two substances – <i>matter</i> and <i>mind</i> .	Materialism: There is only one substance – <i>matter</i> .
	No	Idealism: There is only one substance – <i>mind</i> .	Neutral Monism: There is only one, <i>mental-physical</i> , substance.

Cartesians and Newtonians both rested within the dualism camp. This was not simply due to their acceptance of Christian theological beliefs about the immortality of the soul. For both Cartesians and Newtonians, material systems were in a sense machine-like, but it seemed inconceivable that our mind could be a machine. Descartes noted that human reason was a ‘universal instrument’ which could be flexibly used in all sorts of situations, whereas all the machines he knew of were rigidly disposed to perform one particular function.

What view of the relationship between mind and matter is part of the Contemporary worldview? What have we learned about the mind in the last century that might shed light on this issue?

Neuroscience began in the nineteenth century when major improvements in microscopes made it possible to understand the microscale structure of living things. The Spanish neuroanatomist Santiago Ramón y Cajal showed that the nervous system, like other bodily systems, was composed of distinct living cells, which he called *neurons*. Ramón y Cajal supposed that neurons, with their complex tree-like branches, signalled to one another and formed the working units of the brain. By a current estimate, the human brain contains 86 billion neurons. The substance of the brain clearly possessed the organized complexity that one might suppose a thinking machine would need. But how could a machine think?

In the early twentieth century many neuroscientists worked to understand the electrical and chemical signalling mechanisms of neurons. But this work, by itself, provided little insight into how or whether the brain could be the physical basis of the mind. The needed insight came from another field. In 1936, the mathematician Alan Turing proved that it was theoretically possible for a simple machine, now called a *Turing machine*, to perform any mathematical computation whatsoever, so long as it could be clearly specified. Turing’s finding, in effect, refuted Descartes claim that a machine lacked the flexibility to be a universal instrument.

Although Turing's work was originally intended to address a problem in the philosophical foundations of mathematics, he and others soon realized it also had practical significance. It led directly to the development of the digital electronic computer. By the early twenty-first century, this machine, and a rudimentary form of machine 'thought', have become an ever present feature of daily life.

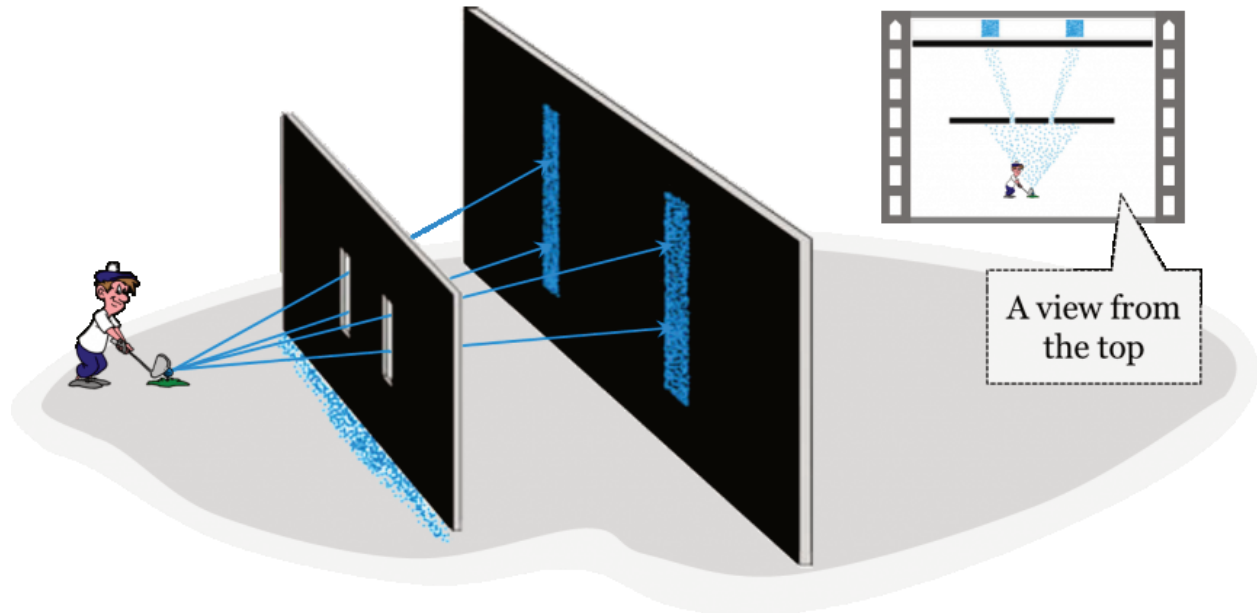
The ability to perform, in seconds, mathematical computations that would take years or centuries for an unaided human being had a profound impact on many areas of science, including, especially the attempt to understand the mind. Early cognitive scientists supposed that the mind literally worked like current computers and functioned by manipulating symbols according to logical rules. This approach met with some successes, as computers could be programmed to perform tasks that would take a great deal of intellectual effort if performed by human beings, like beat grand masters at chess or prove mathematical theorems. But the researchers soon realized that the brain's style of computation was very different from that of current computers and that understanding the mind would require understanding how the neurons of the brain interacted with one another to create it. By the end of the twentieth century, cognitive science and neuroscience had merged as cognitive neuroscience. Being universal instruments, digital computers could be programmed to simulate idealized networks of interacting neurons. These *artificial neural networks* have the capacity to learn and proved capable of flexibly learning to perform tasks like recognizing and classifying patterns, tasks which are thought to be critical to their biological counterparts. Rather than simply reacting to its inputs, the brain has been found to be in a state of constant, internally driven activity, forming a dynamical system that constantly anticipates and predicts its inputs.

More evidence that the brain is the physical substrate of the mind comes from new technologies, such as functional magnetic resonance imaging (fMRI) for directly imaging functional processes in living human brains. In essence, fMRI measures blood flow within the brain. An increase of blood flow to a certain area of the brain indicates neural activity and cognitive use of that area of the brain. fMRI enables neuroscientists to image different parts of the brain activated in response to various stimuli. During a behavioural experiment, a neuroscientist might ask a test subject to perform a simple task or recall a memory that elicits certain emotions. Having mapped different brain processes onto different areas of the brain, neuroscientists are able to successfully predict and measure which parts of the subject's brain will activate in response to their instructions. Thus, they reason that mental tasks, like feeling an emotion or reliving a memory are the result of neural activity in the brain.

Cognitive neuroscientists accept that the brain is the physical substrate of the mind and pursue a theoretical account that seeks to explain cognitive processes like perception, reason, emotion, and decision making on that basis. The claim that mental states are produced by physical processes is incompatible with substance dualism. In fact, if mental states are produced by a physical process, this is a strong indication that the mind does not exist without underlying material processes. In other words, this suggests the view *materialism*, which states that matter is the only independent substance.

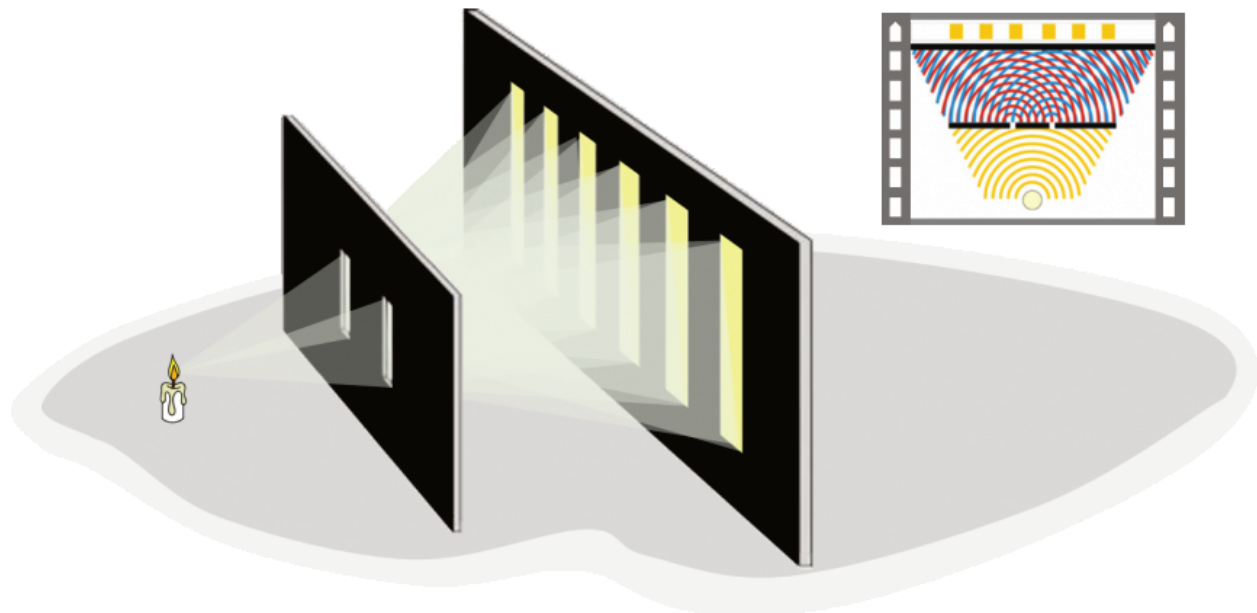
Quantum Mechanics

To appreciate that matter is the only substance is one thing, to have a clear understanding of what matter *is* and what properties it has is a different thing. It is safe to say that the very notion of matter changed through time. To illustrate these changes, we will consider different versions of the famous *double-slit experiment*. Imagine a wall with a slit in front of another solid wall. Now, consider a golfer hitting balls covered in paint in the direction of the two walls. While some of the balls will likely hit the first wall, others will pass through the slit and leave marks on the back wall. Gradually, a linear shape will emerge on the back wall. Now, let's add another slit to the front wall. What sort of pattern would we expect in this set-up? We would expect two painted lines on the back wall, like so:



This picture is very much in tune with the predictions of Newtonian mechanics, where, by Newton's first law, balls are supposed to travel along straight lines if unaffected by any other force.

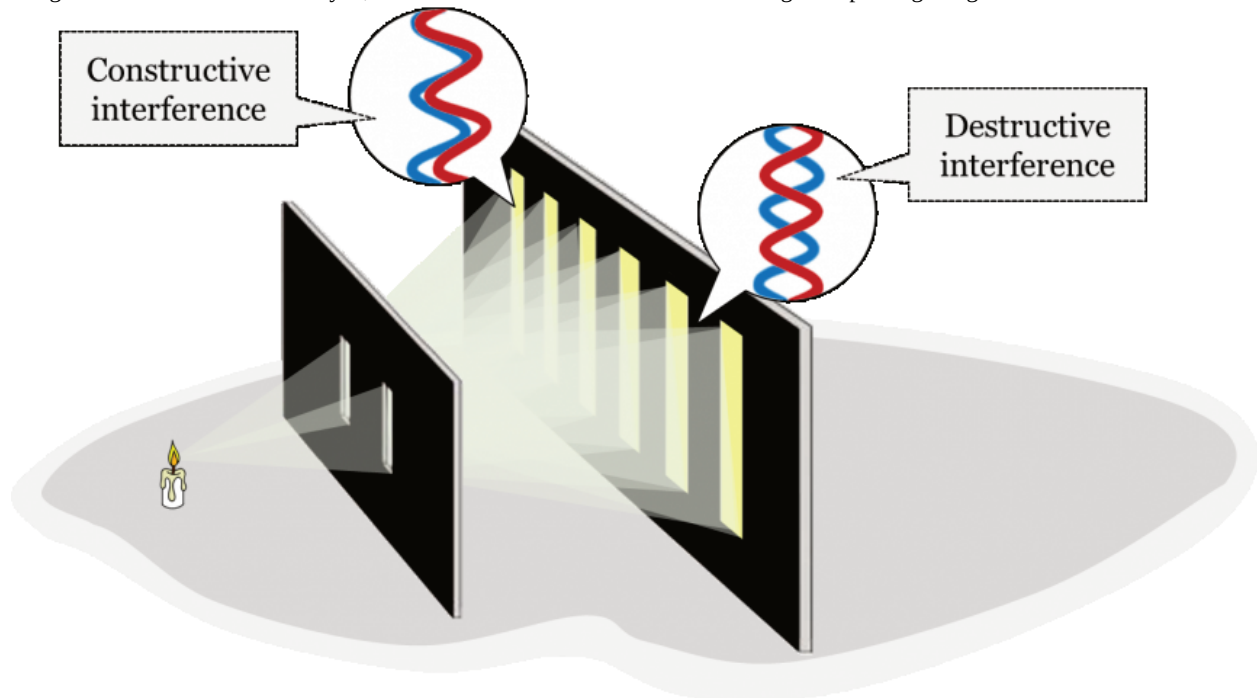
Recall, however, the change in the Newtonian conception of matter discussed at the end of chapter 9. After the acceptance of the wave theory of light in 1819, the notion of matter was widened to include not only corpuscles, but also waves in a fluid medium called *luminiferous ether*. Imagine we replaced the golfer with a light source. If light consisted of corpuscles, then, as for the golf balls, one would expect an image of two straight lines on the back screen. However, if the slits are narrow enough, and close enough together, the image caused by light shining through the two slits and depicted on the back screen would *not* be of two lines. Instead it there would be *multiple* lines:



This strange phenomenon was nicely explained by the wave theory of light. According to this theory, light waves undergo a process called *diffraction*. Diffraction is the bending of a wave around an object or after passing through a hole that is narrow in relation to the wavelength of the light wave. So, rather than continuing in a straight line after passing through a single slit, the wave begins to spread out. In the experiment, however, we have two slits. As the light waves pass through both slits and begin to diffract on the other side of the first screen, they begin to interfere with one another. It's similar to dropping two stones into a pool of calm water at the same time. The ripples will eventually hit one another and create bigger and smaller waves. This phenomenon is known as *interference*, which also happens with waves of light.

There are two types of interference. In *destructive interference*, the crests and troughs of the two waves are out-of-sync and

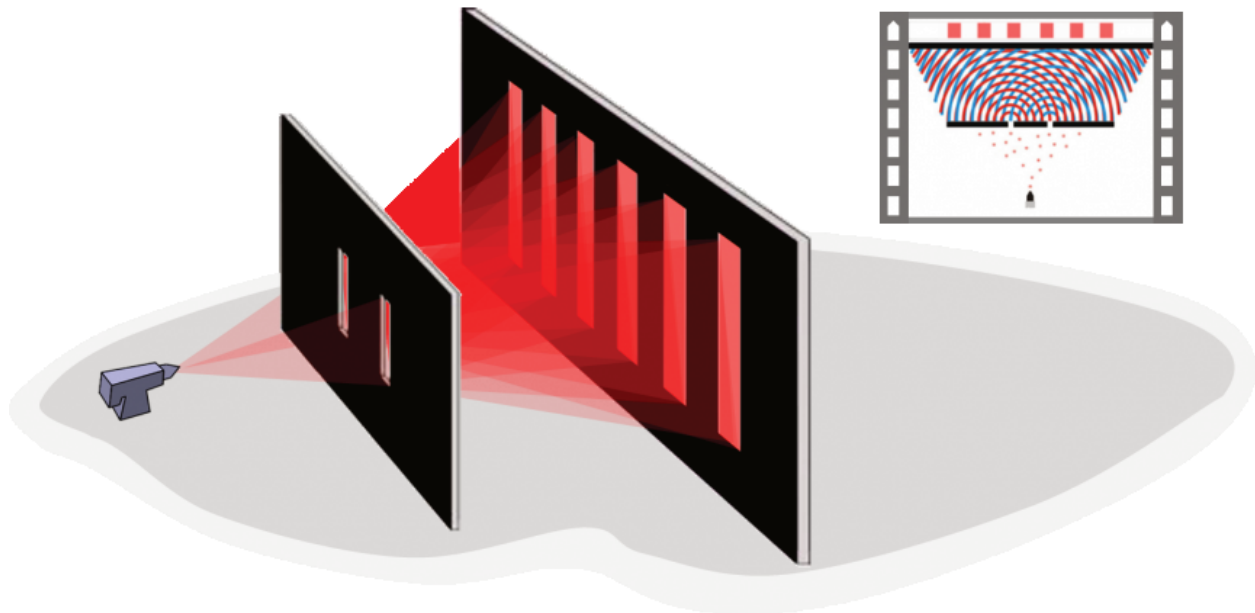
cancel each other out forming darker regions on the screen behind the slits. *Constructive interference* happens when the crests and troughs of the two waves are in-sync, and the waves reinforce each other forming multiple bright regions on the screen.



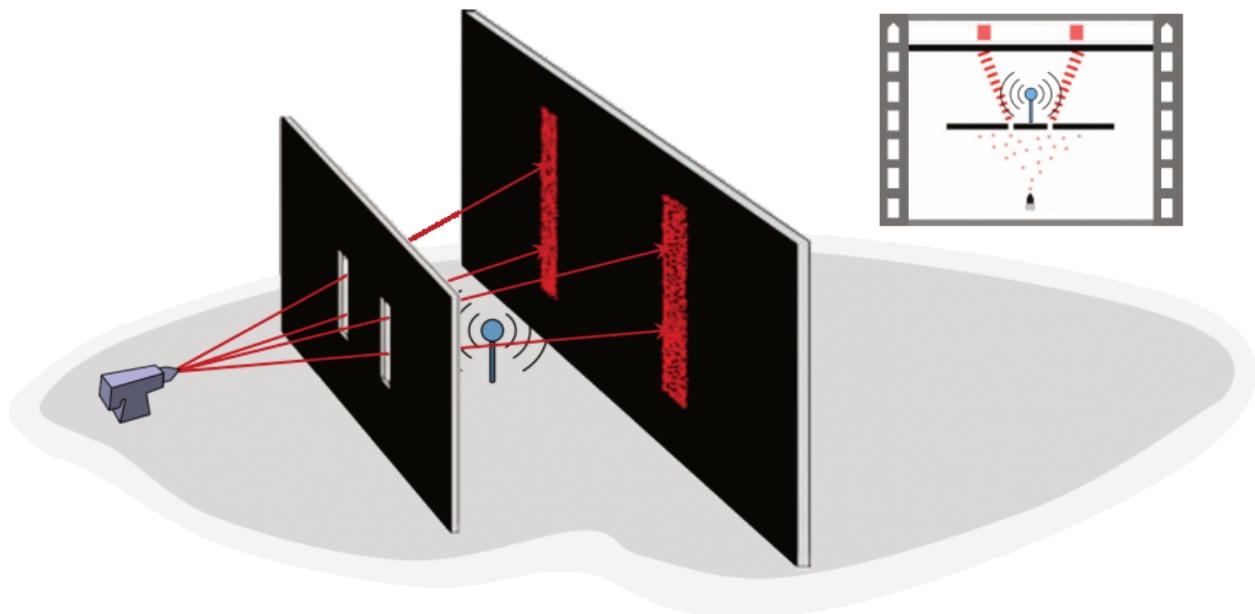
The wave theory of light was accepted in the Newtonian worldview up until about the 1880s. After that, Newtonians accepted classical electrodynamics, which was a theory that gave a unified account of both electrical and magnetic phenomena. In this theory, light is still treated as a wave, but this time as electromagnetic radiation, i.e. a wave in an electromagnetic field that propagates through empty space without the need for a medium like the luminiferous ether.

By the 1920s, however, the Newtonian idea that fundamental physical entities were either *particles or waves* came to be replaced with another idea – the idea of *wave-particle duality*. To better understand this transition, we need to delve into an element of the Contemporary mosaic: *quantum mechanics*.

Quantum mechanics is the physical study of the quantum world – the physics of *very* tiny things. It tries to explain how matter behaves and interacts with energy at the scale of atoms and subatomic particles. The double-slit experiment was conducted again in the 1920s to gain new insight into the nature of quantum phenomena. The task was to find out whether elementary particles such as electrons are corpuscles or waves. The double-slit experiment, in this context, was slightly modified. Instead of a light source, a particle gun was used to fire elementary particles such as electrons. If electrons were corpuscular in nature, then one would expect two straight lines on the back screen. If, on the other hand, electrons were waves, one would expect diffraction patterns on the back screen. A diffraction pattern emerged on the back screen, suggesting that electrons are waves.



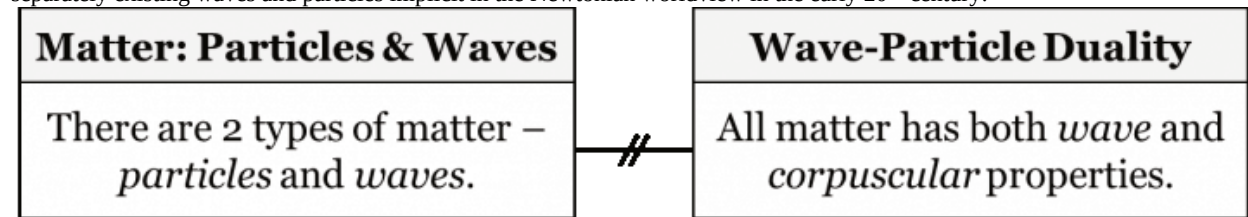
Even if electrons were fired at the screen one at a time, the pattern would still appear. This meant that each electron had to somehow pass through both slits, then diffract and interfere with itself in order to produce the diffraction pattern on the screen. To confirm that each electron passed through both slits, experimenters added a detector next to the slits that would tell us which of the two slits an electron passed through: the detector would indicate to the experimenters whether an electron passed through the left slit, the right slit, both slits, or neither slit. Because of the diffraction patterns they observed, physicists expected the detector to indicate that an electron passed through *both* slits at once. However, the moment the detector was added to the setup, the diffraction patterns mysteriously disappeared; it was now two straight lines!



Each individual electron was detected to pass through only one of the two slits, but never through both. This was a surprising result, since electrons were expected to pass through both slits. It was a behaviour expected of corpuscles, not waves! What the experimenters realized, though, was that the electrons that were emitted behaved differently depending on the presence or absence of the detector. Whenever the detector was present, the electrons fired behaved like particles: *only one* detector would go off to indicate something had passed through its slit and no diffraction pattern would emerge. In contrast, whenever the detector was absent, the electrons fired behaved like waves and diffraction patterns would emerge.

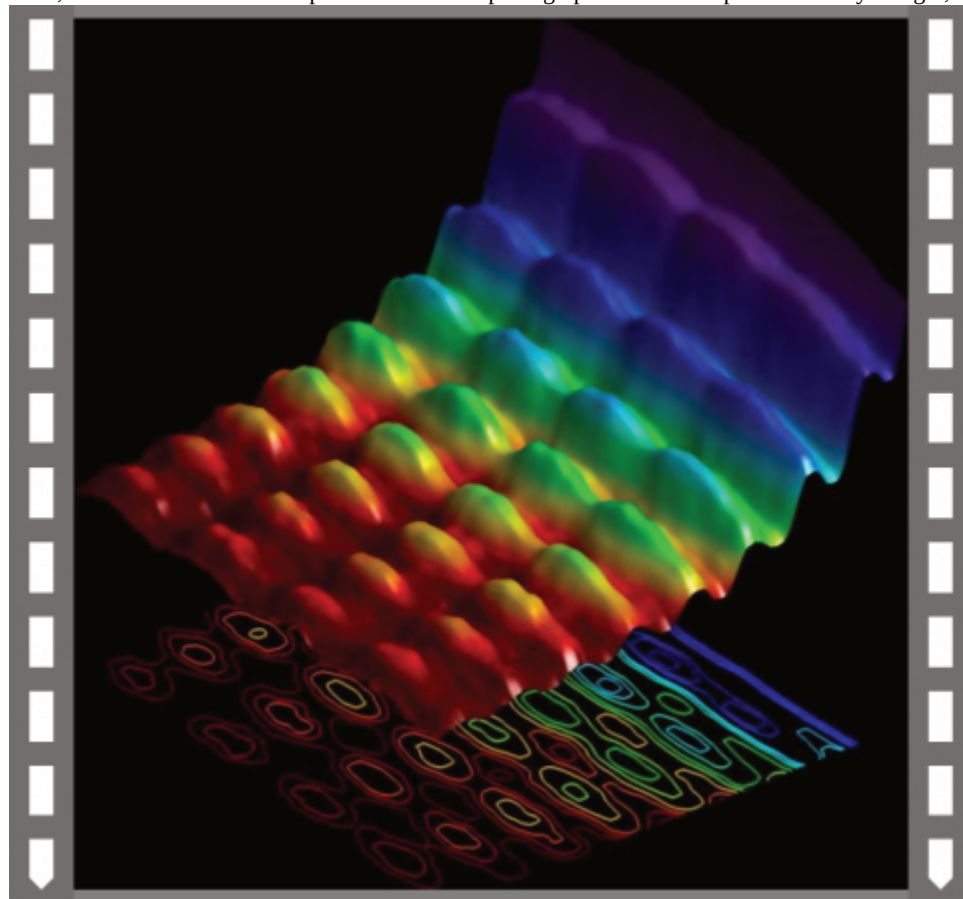
Physicists concluded from this experiment that all matter has both wave and corpuscular (i.e. particle-like) properties. All particles in the *Standard Model of particle physics* – the currently accepted theory of what basic building blocks of matter exist and how they interact – exhibit this dual behaviour. This includes electrons, photons (“light particles”), and even the (relatively)

recently discovered Higgs boson. We consider this dual-behaviour of all matter a metaphysical principle of the Contemporary worldview, and call it *wave-particle duality*. The metaphysical principle of wave-particle duality came to replace the idea of separately existing waves and particles implicit in the Newtonian worldview in the early 20th century:



As another example of wave-particle duality, consider an experiment concerning the so-called *photoelectric effect*. The photoelectric effect is a phenomenon whereby electrons are ejected from a metal surface after light is shone onto that surface. The question Einstein sought to answer was whether waves or particles of light were responsible for the ejection of the electrons. To test for this, one would measure both the frequency and intensity of the light. Einstein hypothesized that light's capacity to eject electrons with a certain energy depended solely on its *frequency*: no matter how high the light's intensity, if its frequency didn't pass a certain threshold, then it would never eject any electrons from the metal surface. This fact was unaccounted for in the wave theory of light, which predicted that increasing the intensity of the light would also eject electrons. Although Newtonians had shown that light behaved like waves, Einstein concluded that *particles* of light had displaced electrons on the metal surface. Einstein's theories about the photoelectric effect did not do away with earlier evidence supporting the wave theory of light. Light still exhibited the properties of a wave. Only now, it also exhibited the properties of a particle.

It is accepted in Contemporary quantum physics that light behaves sometimes as a particle and sometimes as a wave. In fact, in 2015, *Nature Communications* published the first photograph of the wave-particle duality of light, seen below.



The duality of matter is also expressed in Heisenberg's uncertainty principle, one of the fundamental principles of quantum physics. If formulated for a particle's position and momentum, the uncertainty principle states:

The more precise the position (σ_x), the more uncertain the momentum (σ_p), and vice versa:

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}$$


The principle states that when we try to measure the position of an elementary particle, the uncertainty in the particle's momentum increases, and conversely, when we try to measure the particle's momentum, the uncertainty in its position increases. The two uncertainties are inversely related. Importantly, the principle is not about our inability to measure things precisely due to mere technological limitations. It has to do with the fact that elementary particles really are fuzzy entities that only become particle-like when they interact with an external system. This fundamental fuzziness is a consequence of the dual wave-particle nature of matter.

In quantum mechanics, the evolution of a quantum system is described by Schrödinger's wave equation. According to Schrödinger's wave equation, elementary particles are fuzzy, wavelike entities that nonetheless exhibit a definite particle-like state when they are observed. But, even in identical circumstances, this state is not always the same state. The new theories accepted as part of quantum mechanics meant that scientists of the Contemporary worldview adopted a new perspective on cause and effect relationships, modifying their earlier belief in *determinism*. So, let's consider what determinism is all about and briefly highlight what scientists of previous worldviews had to say about the concept, before further delving into what the Contemporary view is on this issue.

Quantum mechanics affected more than just our views on the nature of matter, but also our understanding of *causation*. The question of whether the future is completely determined by the past, or whether it is, in some sense, open, has been one of the central questions of metaphysics since the days of Aristotle. Different stances on this issue affect many aspects of one's worldview, including one's belief in free will, fate, and predetermination. The question also has serious moral repercussions. For instance, if everything in the universe turns out to be strictly determined by the past course of events, then how can we blame a criminal for their crimes? If their actions were indeed completely determined by the whole past of the universe, then, by some philosophical accounts, it is the universe that's to be blamed, not the criminal! Just as with other metaphysical questions, the question of causation has had several different answers. The two basic views on causation are determinism and indeterminism.




Determinism (also known as *causal determinism*, or *strict determinism*) is essentially the idea that every event is determined by preceding events. More crudely, it means that for every effect, there is always a cause that bring it about. Thus, if we could know the present state of the universe (e.g. the positions, masses, and velocities of all the material particles in the universe), we could, in principle, calculate the future states of the universe. Of course, since we are not omniscient and since we don't have unlimited measuring and calculating capacities, we may never *actually* be in a position to make such predictions and know the future with absolute precision. But this is beside the point; what is important here is that, according to strict determinists, nature itself "knows" where it is going: its future is determined by its past.

The opposite of determinism is *indeterminism*, which suggests that there can be uncaused events, i.e. events that are completely random and don't follow from the preceding course of events. Note that indeterminists don't necessarily claim that *all* events are uncaused. It's sufficient to accept the existence of *some* uncaused events to qualify as an indeterminist.

 Can there be uncaused events?	
Yes	No
Indeterminism: There can be <i>uncaused</i> events. In some situations, there is an infinite number of possible outcomes.	Strict Determinism: All events have their causes and the same initial conditions <i>always</i> produce the same effects.

Determinism and indeterminism don't exhaust the spectrum of views on the issue. One historically popular conception is that of *dualistic determinism*, which holds that while events in the material world are strictly determined, the human mind has free will and is capable of acting spontaneously. It was this view that was implicit in several worldviews.

In the Aristotelian-Medieval worldview, nearly all events were considered strictly determined, except for those events affected by human free will or divine acts. The Aristotelian-Medieval community believed that all celestial phenomena are strictly deterministic, as they knew that the future positions of stars and planets were predictable. They extended this deterministic view to terrestrial phenomena as well, because the celestial realm exerted influence on the natural processes of the terrestrial realm. That is, medieval scholars held that *terrestrial* processes unaffected by human free will were also deterministic. However, they also believed that a benevolent God had granted humans free will – an ability to act spontaneously and make decisions that do not always follow from the preceding courses of events. They accepted that many human actions and decisions were *uncaused*. In other words, the conception of *dualistic determinism* was implicit in the Aristotelian-Medieval worldview: in the material world all events are deterministic, but the mind is free to act spontaneously. Philosophically speaking, the conception of dualistic determinism differs from both strict determinism and indeterminism in that it assumes that causation and spontaneity can exist in different worlds:

 Can causation and spontaneity exist in different worlds?		
Yes	 No	
Dualistic Determinism: While in the <i>material</i> world, all events are strictly deterministic, the <i>mind</i> is free to act spontaneously.	 Can there be uncaused events?	
	Yes	No
	Indeterminism: There can be <i>uncaused</i> events. In some situations, there is an infinite number of possible outcomes.	Strict Determinism: All events have their causes and the same initial conditions <i>always</i> produce the same effects.

While disagreeing with Aristotelian-Medieval scholars on many metaphysical issues, both Cartesians and Newtonians held a very similar perspective on the issues of causation: both communities accepted dualistic determinism, albeit for their own reasons. Their deterministic stances regarding the material world actually stemmed from the universality of their laws of physics. For instance, in Newton's physics, if one knows the current arrangement of planets in the solar system, one can *at least in principle* predict the positions of planets for any given moment in the future. The same goes for any material system, no matter how complex. While it may be virtually impossible for humans to do the actual calculations and make these predictions, according to the laws of Newtonian physics the future of any material system strictly follows from its past. The same holds for Cartesian physics: the same initial conditions *always* produce the same effects. That said, Cartesians and Newtonians also accepted that the mind has free will to make uncaused decisions and, thus, disrupt the otherwise deterministic course of events. As a result, both Cartesians and Newtonians believed that the future of the world wasn't strictly deterministic as long as it was affected by creatures with free will. Thus, dualistic determinism was implicit in not only the Aristotelian-Medieval worldview, but in the Cartesian and Newtonian worldviews as well.

Now we can consider how the theories of quantum mechanics and the metaphysical principle of wave-particle duality in the Contemporary worldview affect current scientists' views on determinism. A quick hint: they're *not* dualistic determinists.

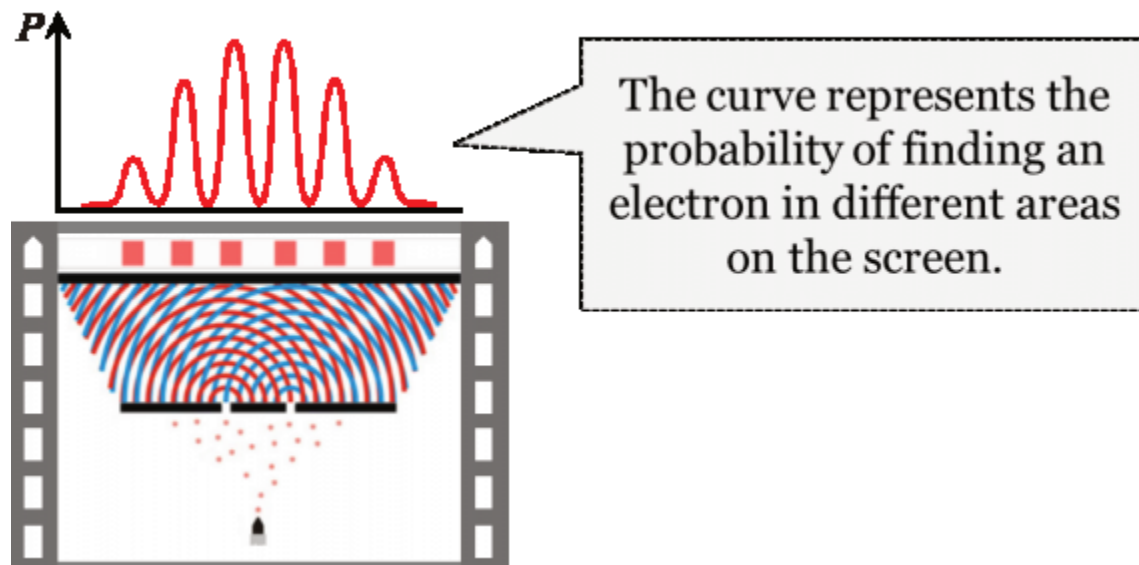
To better understand the stance of the Contemporary worldview on the question of causation, let's consider the process of radioactive decay. Nowadays, physicists accept that there is a 50% probability that an atom of radium will decay into radon and helium after 1600 years. This period of time is known as the *half-life* of radium. Say we could isolate one thousand atoms of radium, place them into a sealed container, and open the container again after 1600 years. Statistically speaking, we're most likely to find about 500 atoms of radium left, while the other 500 atoms will have decayed into radon and helium. What we observe here is that all of the atoms of radium were placed into the same container, yet only half of them decayed while the rest remained intact. The fact that the same initial conditions can lead to two distinct outcomes suggests that the fundamental processes of the world are probably not *strictly* deterministic. Instead, quantum physics seems to suggest that fundamental processes are determined *probabilistically*. This is the view of probabilistic determinism, which is implicit in the Contemporary worldview.

According to *probabilistic determinism*, all events have their causes, but the same initial conditions may produce different effects.

<p>• Can causation be such that the same initial conditions produce different effects?</p>			
Yes	No		
Probabilistic Determinism	<p>• Can causation and spontaneity exist in different worlds?</p>		
	Yes	No	
	Dualistic Determinism	<p>• Can there be uncaused events?</p>	
		Yes	Indeterminism

All of these effects are probabilistically determined, in the sense that there is a certain likelihood that a certain course of events will occur, but this course of events is not strictly determined by the past course of events. In our example above, there is a statistical likelihood for decay, i.e. there is a 50% chance of radium decaying after 1600 years. Yet, the theory doesn't indicate which of the atoms will or will not decay in 1600 years. Probabilistic determinism interprets this as suggesting that future events are determined by past events but not strictly: nature is constrained to a limited number of options.

Consider another example, the double-slit experiment in quantum physics. Essentially, the theory allows us to predict the probability with which an electron will strike a certain region on the wall. That probability will be greater in lighter regions of the diffraction pattern, and smaller in darker regions.



It is impossible to say with certainty where on the back screen a given electron will strike after passing through the slits. But there is a high *probability* that it will strike in one of the bright bands. According to probabilistic determinism, this is another indication that natural processes are caused probabilistically rather than strictly.

We should be careful not to confuse probabilistic determinism with *indeterminism*. While both concepts admit to a multiplicity of outcomes following the same initial conditions, there is a clear limit to that number of outcomes under probabilistic determinism. An indeterminist would, technically speaking, be open to the possibility that after 1600 years radium could transform into a bar of gold, or a pile of dirt, or Slimer the slimy green ghost from *Ghostbusters*. However, a probabilistic determinist places a limitation on the number of potential effects that could follow the same cause and suggests that not everything can happen: nature 'chooses' from a limited set of options. In other words, according to probabilistic determinism, there is not an infinite number of possible outcomes given a certain set of initial conditions.

Now, how would a strict determinist react to this? One natural reaction is to suspect that there is something deficient in our knowledge of the initial conditions. Indeed, when quantum mechanics was created in the 1920s, this was precisely the reaction of some famous physicists, including Albert Einstein. While Einstein appreciated that quantum mechanics provided a great improvement in our understanding of the world of elementary particles, he also argued that the theory is deficient as it fails to provide precise predictions of quantum phenomena. According to Einstein and other strict determinists, if half of the atoms of radium end up decaying while the other half do not, then perhaps this is because they didn't start from the same initial conditions; perhaps there was some unknown cause that leads to only some of the atoms of radium decaying. This is the interpretation of the strict determinist, who believes that all events are strictly caused by past events and the probabilistic nature of our predictions is a result of the lack of knowledge on our part. According to strict determinists, there must be a hidden cause that explains why those particular 500 atoms ended up decaying. Our theories may be incapable of telling us what that hidden cause is, but that is exclusively our problem; nature itself "knows" where it's going. In this view, future research may reveal that hidden cause and restore the strictly deterministic picture of the world. Strict determinists would compare the case of radioactive decay with that of a coin toss. If we were to toss a coin one thousand times, we would probably observe it landing on heads approximately half the time. Colloquially, we might say that there is a 50/50 chance of a flipped coin landing on heads or tails. However, this is clearly a result of our lack of knowledge of initial conditions. If we were to measure the initial position of the coin before each toss as well as the force applied to the coin, then we would be able to predict precisely whether it will land heads or tails. If we wanted to, we could even construct a contraption that could toss a coin in such a way that it would consistently land on heads. According to strict determinists, the situation with radioactive decay is similar, except that we are not in a position to measure the initial conditions precisely enough to be able to predict the outcome. However, nature "knows" the outcome of the process of decay, just as it "knows" whether a coin will land on heads or tails. In short, strict determinists would deny the idea of something being caused probabilistically; in their view, everything has a cause and follows from that cause in a strict fashion.

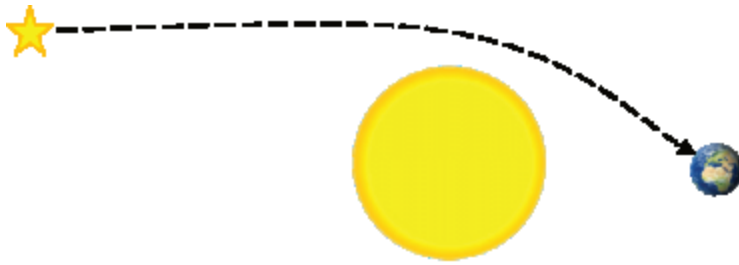
How would a probabilistic determinist respond to this? Probabilistic determinists would readily agree that our theories are not perfect descriptions of the world; after all, we are all fallibilists and understand that no empirical theory is perfect. But is this reason enough to ignore what our current theories tell us and instead speculate what our future theories will be like? In the absence of an absolutely correct description of the world, our best bet is to study carefully what our currently accepted theories tell us about the world. Is it conceivable that our future quantum theories will be strictly deterministic and will provide precise predictions of all quantum phenomena? If we are fallibilists, our answer is "yes": such a scenario is conceivable. But it is equally conceivable that our future theories will be probabilistic – we have no way of forecasting this. Our best option is to stop guessing what our future theories may or may not bring us and focus instead on what our currently accepted theories have to say on the matter. Thus, according to probabilistic determinists, the implications of our current quantum theory should be taken seriously, and probabilistic determinism is one of these implications. On this view, the probabilistic predictions of quantum theory have nothing to do with our lack of knowledge or our failure to predict the outcome: the process is itself probabilistic. Since the theory tells us that radioactive decay is not a strictly deterministic process, then that's how we should view it. Everything else is a speculation on what the future may have in store.

General Relativity and Cosmology

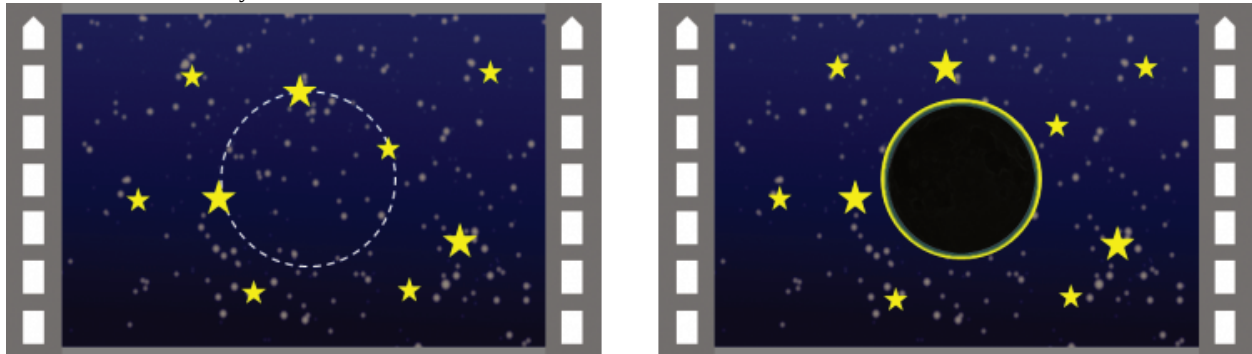
So, we understand a bit about how physics works at the quantum scale and its implications for the Contemporary worldview, but what about physics at the large cosmological scale? For this, we should refer to general theory of relativity proposed by Einstein in 1915. As we learned in chapter 1, Einstein's theory of general relativity posits that objects with mass curve the space around them. This applies to *all* objects with mass but becomes especially apparent in the case of more massive objects, such as the Earth, the Sun, and the Milky Way galaxy, or extremely compressed objects like black holes. The curvature of space around an object depends on its mass and its density. This is why the curvature of space around a massive object compressed to extreme density, like the singularity of a black hole, is such that even rays of light cannot escape. Consequently, according to general relativity, there is no force of gravity and, therefore, objects don't really attract each other, but merely *appear* to be attracting each other as a result of moving inertially in a curved space.

Similarly, general relativity predicts that material objects affect time, by making it run slower relative to a clock in a less strongly curved region of space-time. Suppose we have two synchronized clocks; we keep one on the surface of the Earth, and we take the other to the International Space Station, where space-time is less strongly curved than at the Earth's surface. According to general relativity, the clock in the Space Station will run slightly faster than the one on Earth. This is because time on the Earth runs slightly slower due to the space-time curvature of Earth's gravitational field. As a result, when the Space Station clock is brought back to the Earth, it will show a time that is slightly ahead of that shown by the one on the Earth. This difference between the elapsed times is known as *time dilation*. In general, the stronger the gravitational field of the object, the greater the time dilation. For instance, if the Earth were either compressed to a greater density or increased in mass, space-time would be more strongly curved at its surface. The difference between the two clocks would therefore be greater.

Both the effects of curved space and the effect of time dilation have been experimentally confirmed. It follows from Einstein's general relativity that even light itself will bend noticeably around very massive objects. Consider a light ray coming from a distant star. According to the theory, it should slightly bend while passing in the vicinity of massive objects. It follows, therefore, that the same star should appear at a different location in space when the light coming from that star passes near the Sun.



To observe this effect, we first take a picture of that specific region of space at night and then take another picture of that same fragment of space when the Sun is in the vicinity. To take the latter picture we will need to wait for a solar eclipse to make sure that the light rays of the Sun don't obstruct the light rays coming from the distant star. According to general relativity, the stars in the vicinity of the Sun will appear slightly dispersed from the Sun due to the fact that the light rays coming from these stars will bend in the Sun's vicinity.



The phenomenon of light-bending was one of the novel predictions of general relativity. Thus, when the amount of light bending predicted by Einstein's theory was first observed by Arthur Eddington in 1919, it was considered a confirmation of general relativity. Since the method of the time was the hypothetico-deductive method, this led to the acceptance of general relativity ca. 1920 and rejection of the Newtonian theory of gravity. Since then several other novel predictions of general relativity have been confirmed. For instance, the phenomenon of time dilation has been confirmed by slight differences in times shown by a pair of precise clocks on the Earth and the Space Station. The phenomenon was also confirmed by the fact that the clocks on GPS satellites run slightly faster than those on the Earth.

As with any fundamental theory, general relativity has an effect on accepted metaphysical principles. So how does general relativity shape our *contemporary* metaphysical principles? Explicating the metaphysical principles implicit in the Contemporary worldview is not always an easy task, since our currently accepted theories sometimes seem to suggest conflicting metaphysical principles. The point here is not that contemporary scientists are not in tune with their own theories. It's instead that it's difficult to know what principles to extract from our currently accepted theories. In what follows we will outline a number of metaphysical questions to which we won't provide the contemporary answers. Instead, we will show why the extraction of the answers to these questions is not a simple task.

Let's start with the question of empty space: according to the Contemporary worldview, can there be empty space? Physicists sometimes speak of the 'vacuum of outer space' to emphasize the overall emptiness of the universe. However, when pressed, they might admit that a few hydrogen atoms exist for every cubic meter of outer space, which might be interpreted in tune with the conception of plenism. Yet, between the hydrogen atoms would be nothingness, meaning a very close approximation of a vacuum seems to be possible. On the other hand, observations of cosmic microwave background radiation also suggest that photons leftover from the Big Bang permeate all seeming empty space of the universe. This seems to suggest that a *perfect* vacuum in nature is impossible after all. Then again, general relativity tells us that space and time lack an independent existence and instead constitute a unified four-dimensional *space-time*, which is affected by both matter and energy in the universe. The dependence of space-time on matter and energy seems to suggest that there can be *no* empty space. Quantum mechanics also implies that there is no empty space. In addition to the position and momentum uncertainty relationship mentioned earlier, Heisenberg's uncertainty principle involves a time and energy uncertainty relationship, which means that the shorter the time interval one considers, the greater the uncertainty regarding the amount of energy in a region of space. In effect, in any region of space, so-called *virtual particles* are constantly popping into existence and vanishing from existence. This phenomenon is referred to as *quantum foam*. This surprising effect is a well-tested phenomenon, supported by many experiments. Perhaps, then, plenism is implicit in the Contemporary worldview. Thus, the question is: which metaphysical view – vacuism or plenism – is implicit in our Contemporary worldview?

What about *motion* in the universe: do objects move exclusively as a result of action by contact, or is action at a distance possible in the Contemporary worldview? On the one hand, we accept nowadays that the speed of light is the ultimate speed in the universe. Traveling faster than the speed of light would violate our current laws of physics. Physics therefore rules out special cases of action at a distance faster than the speed of light, most clearly cases of *instantaneous* action at a distance. According to general relativity

even changes in the curvature of space-time propagate at the speed of light. To appreciate this, consider that the Sun curves the space-time around it. It takes the light rays of the Sun approximately 8 minutes to reach the Earth. If, by any miracle, the Sun were to disappear, it would take approximately 8 minutes for the changes in the curvature of the space-time to propagate and reach the Earth and change the Earth's orbital motion. The Earth wouldn't start hurtling away into space immediately. If instantaneous action at a distance is theoretically impossible, then perhaps all matter moves and interacts through actual contact limited by the speed of light.

On the other hand, physicists working on quantum mechanics tell us of *quantum entanglement*, which – at least at face value – seems to allow for instantaneous action at a distance. Quantum entanglement is a phenomenon whereby a pair of particles at an arbitrary distance apart is entangled such that measuring and manipulating the properties of one particle in the pair causes the other, at any distance away, to instantaneously change as a result. This phenomenon was first experimentally confirmed in 1982 by Alain Aspect. Since then, numerous experiments were conducted with different types of particles, with different properties of particles (e.g. momentum, spin, location), as well as with varied distances between the entangled particles. Every time, manipulating one particle instantaneously changes the other. These experiments with quantum entanglement seem to suggest that instantaneous action at a distance is possible, despite the limitations imposed by the speed of light. This is clearly in conflict with the idea of action by contact that follows from general relativity. Now the question is: do contemporary physicists accept the possibility of action at a distance, or do they manage to reconcile the seeming action at a distance of entangled particles with the idea of action by contact?

Our final question of this section is whether the universe is finite or infinite. As far as the related question of physical boundaries of the universe is concerned, Contemporary science accepts that the universe is *boundless*, i.e. there are no physical boundaries to the universe. But a universe without boundaries can still be finite if the universe loops back into itself, like the four-dimensional equivalent of a sphere. This kind of a universe will be boundless but finite. Alternatively, the universe might be boundless and infinitely stretching outwards. So, do we accept in the Contemporary worldview that the universe is finite or infinite? Answering this question actually depends on the actual curvature of the space of the universe. Alexander Friedmann suggested three models of the universe, each with its distinct curvature:

Open universe: An open universe has a *negative curvature*, meaning that space-time curves in two different directions from a single point. While the curvature of a three-dimensional space can be precisely described mathematically, it is difficult (if not impossible) to imagine. It becomes easier if we think in terms of a two-dimensional analogue, in which case a three-dimensional space becomes a two-dimensional surface. In the case of an open universe this surface is curved into a shape like a saddle, or a Pringle chip. Since space-time would never curve back onto itself in this type of universe, the open universe will also be an *infinite universe*.

Closed universe: A closed universe has a *positive curvature*, meaning that space-time curves away in the same direction from any single point in the universe. Again, using our two-dimensional analogy, we can think of a closed universe as a surface shaped like a sphere, curving in the same direction from any point. It would also be a *finite universe*, for the continuous curve would eventually circle back around to itself. If we travelled far enough out into space on a spaceship, we would eventually reach Earth!

Flat universe: A flat universe has, overall, *zero curvature*, meaning that space-time extends in straight lines in all directions away from any single point. We can think of a flat universe as being shaped like a flat piece of paper, or a flat bedsheet. Such a universe would still have occasional bends here and there around massive objects; however, it would be flat overall. A flat universe would be an *infinite universe*, as space-time is never curving back onto itself.

All three models of the universe are compatible with general relativity. Yet these three models of the universe propose different geometries of the universe. Now, which of these models best describes the space of our universe? The answer to this question depends on how much mass there is in the universe. If the mass of the universe is greater than a certain threshold, then the universe is curved onto itself and is finite; otherwise it is open and infinite.

One takeaway here is that it's not always easy to extract the metaphysical principles implicit in a worldview. Often it requires a meticulous analysis of the theories accepted in that mosaic and the consequences that follow from these theories. Thus, the stance of a community towards a metaphysical question may or may not be easy to unearth. The question of empty space, the question of action at a distance, and the question of finite/infinite universe are some of the most challenging metaphysical questions in the context of the Contemporary worldview.

The Big Bang theory

What about the topic of religion in the Contemporary worldview? It's a strange topic to bring up while discussing the physics of special and general relativity. But in a sense, physicists traversed into the territory of theologians in the twentieth century when they attempted to look at the origins of the universe. The currently accepted theory on the origins of the universe is called the *Big Bang theory*.

In the late 1800s and early 1900s, astronomers began using a new instrument to observe celestial objects called a *spectroscope*. Spectroscopes essentially measure the distribution of wavelengths of light being emitted from an astronomical object, allowing astronomers to determine certain features of a star, galaxy, or other celestial body. Because each chemical element has its own characteristic spectral "fingerprint", the spectroscope can be used to determine the chemical composition of an astronomical object. The spectroscope can also be used to determine whether a light-emitting object is approaching or moving away from the point of

may require the unification of the two theories in a theory of quantum gravity, currently a pursued goal of theoretical physics. Some physicists speculate that the Big Bang may have been a random quantum event, like the appearance of virtual particles in the quantum foam. Others argue for models in which the extreme conditions of the singularity are avoided, and the Big Bang was caused by the collapse of a prior universe. So, the same objections that David Hume raised to the natural theology of his time (see chapter 9) apply with equal force to the idea that God caused the Big Bang or “fine-tuned” the physical properties of the universe to make them compatible with our form of life.

There is another reason why modern physicists and astronomers may be deeply reluctant to accept theological conclusions regarding the Big Bang. The search for signs of other intelligent beings in the cosmos is a pursued goal of modern astronomy. Any claim of such intelligence must meet the criteria of the contemporary employed method summarized by astronomer Carl Sagan’s maxim that “extraordinary claims demand extraordinary evidence”. This search has occasioned a series of false alarms – each one an illustration of David Hume’s worry that unfamiliar natural phenomena can easily be confused with the designed artefacts of a powerful intelligence. By the method currently employed to evaluate hypotheses concerning extra-terrestrial intelligence, no claim that a cosmic phenomenon was engineered by a non-human intelligence would be accepted unless all plausible explanations involving non-intelligent natural phenomena were first ruled out. In the case of theological claims about the Big Bang, the discussion above illustrates why physicists and astronomers, as a community, are unlikely to believe that the standards of this method have been met.

Questions of what happened before the Big Bang are, at best, regarded as highly speculative by contemporary scientists. A strong indication of this reticence among scientists towards religious questions would be to look at successful grant winners in the hard sciences and recognize the types of questions they are trying to answer have nothing to do with God and God’s role in the universe. While individual scientists hold varied religious views, the accepted view of the scientific community appears to be one of agnosticism about the existence of God. God may or not have played a role in the origins of the universe, but this is a question unanswerable by our current mosaic, so we simply state that we do not know.

Answering exactly how and when theology was exiled from the mosaic is quite difficult. Was it with the observation of redshift and subsequent acceptance of the Big Bang theory? Was it the consequence of developments in evolutionary biology that similarly eliminated the role of God in the creation of humanity? Was David Hume’s critique of natural theology the critical factor? At this point, we simply don’t know. Suffice it to say, while monotheism was implicit in the Aristotelian-Medieval, Cartesian, and Newtonian worldviews, agnosticism is implicit in the Contemporary worldview.

Contemporary Methods

What are the methods employed in the Contemporary worldview? As far as the fundamental requirements of empirical science are concerned, not much has changed since the Newtonian worldview. Contemporary scientists working in the empirical sciences continue to employ the *hypothetico-deductive (HD) method*. So, in order to become accepted, an empirical theory is usually expected to either provide a confirmed novel prediction (if it happens to introduce changes into the accepted ontology) or be sufficiently accurate and precise (in all other cases).

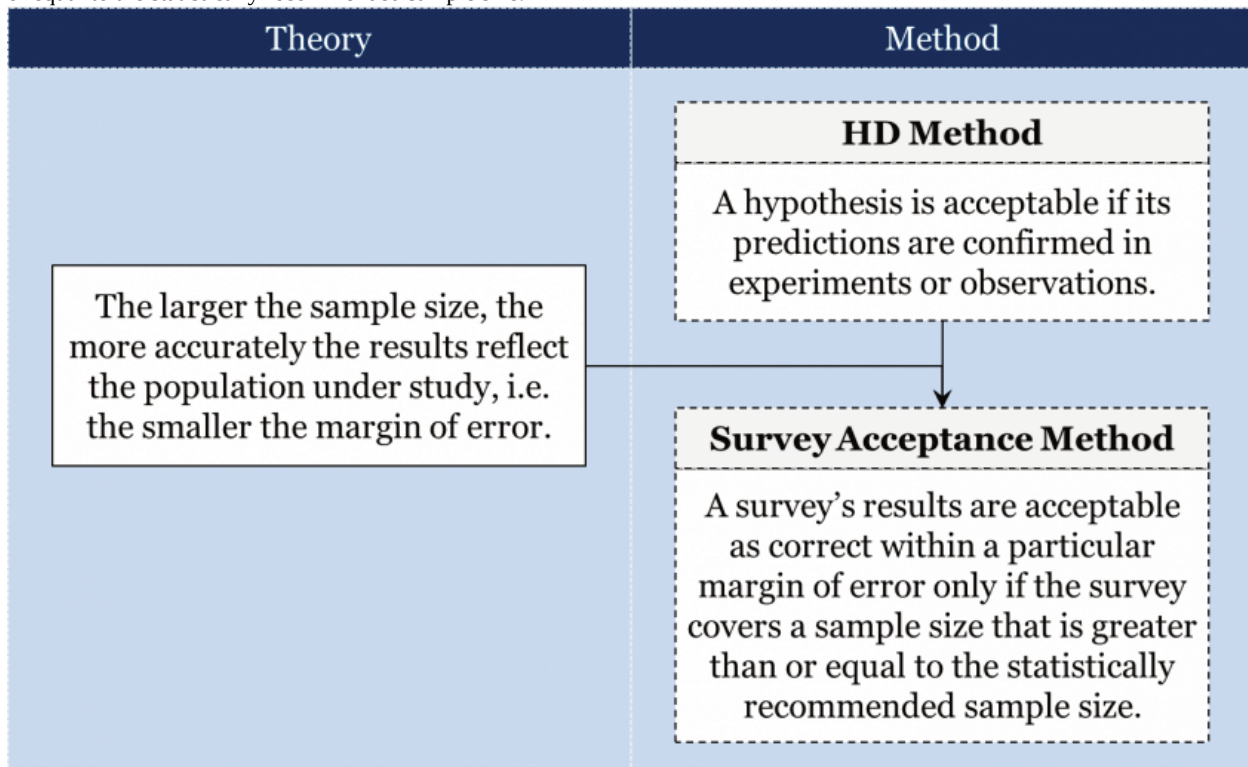
It is important to note that this doesn’t mean that there have been no changes in methods since the 18th century. In fact, many more *concrete* methods have become employed since then; each of these methods are based on the general requirements of the HD method. One obvious example is the different drug-testing methods that we considered in chapter 4: all of them were more specific implementations of the requirements of the HD method. For example, the double-blind trial method has the same requirements as the HD method plus a few additional requirements: that the drug’s effect should be confirmed in a double-blind trial. Generally speaking, there are many concrete methods, that impose additional requirements on top of the ones stipulated by the HD method.

Consider, for instance, our contemporary requirements concerning the acceptability of new elementary particles. Suppose there is a new theory that hypothesizes the existence of a new type of particle. Under what circumstances would we be prepared to accept the existence of that hypothetical particle? We would probably expect the hypothesis to be tested just as the HD method stipulates. Yet, we would likely expect something even more specific, i.e. we would expect the particle to be observed in a series of experiments in a properly functioning particle accelerator under predicted conditions. Thus, we could say that our *particle existence method* can be explicated along these lines:

Particle Existence Method

An elementary particle can be said to exist if its existence has been confirmed in a series of experiments in a properly functioning particle accelerator under predicted conditions.

Consider another example. Suppose there is a sociological hypothesis that proposes a certain correlation between family size and media consumption: say, that larger families devote less time to consuming media content than smaller ones do. Note that this is not proposing a causal relationship between family size and media consumption time, but merely suggesting a correlation. Now, how would we evaluate this hypothesis? We would probably expect it to be confirmed by a suitable sociological survey of a properly selected population sample. It is clear that we wouldn't expect 100% of the population to be surveyed; that's not feasible. Instead, we would expect the survey to cover a certain smaller sample of the population. But how small can the sample size be without rendering the results of the survey unacceptable? This is where our knowledge of probability theory and statistics comes into play. According to the theory of probability, if a sample is selected randomly, then a greater sample size normally results in a more accurate survey. Thus, every such survey will necessarily have a certain margin of error – the range of values within which the figure can be said to correctly reflect the actual state of affairs. The greater the sample size, the smaller the margin of error. There are statistical equations that allow one to calculate the margin of error given the population size and sample size. For example, if we were to study the relationship between family size and media consumption across Canada with its population of about 37 million people, then a sample size of 1,500 would have a margin of error of 2.53%. In short, we would accept the results of a sociological survey as correct within a particular margin of error only if the survey covered a sample size that was greater than or equal to the statistically recommended sample size.



In addition to the HD method, contemporary science also seems to employ the *axiomatic-deductive method*. According to this method, a proposition is acceptable either if it is an axiom of a theory or if it is a theorem that logically follows from the accepted axioms. This is the method currently employed in formal sciences, such as mathematics or logic. Each mathematical theory postulates a set of axioms which define the mathematical structure under study and then proceed with deducing all sorts of theorems about these structures. Recall, for instance, the axioms of different geometries from chapter 2 which provided different definitions of “triangle”. In each of these different geometries, all the theorems about triangles logically follow from the respective axioms.

This method sounds similar to the *intuitive-deductive* method, and indeed it is: after all both methods require the theorems to be deductive consequences of the accepted axioms. But a key difference is that the method of intuition requires the axioms to be intuitively true, while the axiomatic-deductive method doesn't require that. The axiomatic-deductive method considers the axioms to be merely defining the formal object under study regardless of whether they appear intuitively true or counterintuitive.

In brief, it is important to keep in mind that there might be many different methods in different fields of inquiry. Some methods might be very general and applicable to a wide range of phenomena (e.g. the HD method, the axiomatic-deductive method), while other methods may be applicable only to a very specific set of phenomena (e.g. drug-testing methods). In any event, these methods will be the deductive consequences of our accepted theories.

Summary

The following table summarizes some of the metaphysical principles of the Contemporary worldview:

Cartesian	Newtonian	Contemporary
Dualism	Dualism	Materialism
Mechanicism	Particles & Waves	Wave-Particle Duality
Dualistic Determinism	Dualistic Determinism	Probabilistic Determinism
Monotheism	Monotheism	Agnosticism
HD method	HD method	HD method

We've also left a few question marks here and there:

Cartesian	Newtonian	Contemporary
Action by Contact	Action at a Distance	?
Plenism	Vacuism	?
Infinite Universe	Infinite Universe	?

This concludes our discussion of our four snapshots from the history of science.

11.

Worldviews: Metaphysical Components -- Introduction to History and Philosophy of Science

Chapter 11: Worldviews: Metaphysical Components

Intro

Individual elements in a scientific mosaic are relatively stable. Recall that the first law of scientific change, *the law of scientific inertia*, states that once a theory or method becomes part of the mosaic, it remains there, unchanged, until it is replaced by a new element.

But what if we “zoom out” and consider an entire mosaic? Remember that a scientific mosaic is the set of *all* theories the scientific community accepts, and *all* methods that it employs. The mosaic of the contemporary scientific community, for instance, is extremely complex. It is made up of the countless elements accepted and employed in every particular field of empirical and formal science. While each element is relatively stable – as per the *first law* – ongoing scientific research ensures that the *mosaic* is in a state of near-constant change. *Any* change of *any* element, after all, counts as a change to the mosaic.

In 2017, for instance, a study was published showing that, on average, zoologists today confirm the existence of one new species of plant or animal in the Amazon rainforest *almost every other day!* Even though zoology is only one part of biology, and biology is only one part of our contemporary mosaic, these frequent changes in our knowledge of Amazonian rainforest biodiversity can lead us to conclude that the mosaic itself changes almost constantly. So, while individual elements within mosaics are relatively stable, mosaics themselves change every time the community accepts a new theory or employs a new method.

But if mosaics are in a state of near-constant change, what features of mosaics are stable enough to allow us to identify distinct, relatively stable “worldviews” within the flux of the process of scientific change? The Aristotelian-Medieval, Cartesian, and Newtonian mosaics all changed like our Contemporary mosaic does, but not every change in these mosaics led to a fundamental change in worldview, in the same way that the knowledge of new species in the Amazon doesn’t profoundly change our worldview today. Over the past four chapters we have defined each worldview by focusing on some central elements that are remarkably stable compared to others. These elements are the fundamental *assumptions* (conceptions, “isms”) you have become familiar with. Recall, for instance, hylomorphism, pluralism, teleology, plenism, monotheism, heterogeneity, or finite universe that were among the fundamental metaphysical assumptions of the Aristotelian-Medieval worldview. Or recall Cartesian mechanicism, dualism, action by contact, dualistic determinism, plenism, monotheism, homogeneity, and infinite universe. Similarly, Newtonian metaphysical assumptions included dynamism, dualism, action at a distance, dualistic determinism, vacuism, monotheism, homogeneity, and infinite universe among many others. Finally, among the contemporary metaphysical assumptions are wave-particle duality, probabilistic determinism, materialism, and agnosticism.

What all of these “isms” have in common, and what we argue accounts for their greater degree of stability, is the fact that these elements are the *metaphysical components* of each of these mosaics. But what is metaphysics?

The term *metaphysics* is notoriously difficult to define. For our purposes, however, a relatively simple definition will suffice: metaphysics is *a set of views about the nature of the world taken as a whole*.

Metaphysics ≡
≈ a set of views about the world taken as a <i>whole</i> .

That is, metaphysics grapples with questions about *reality in general*, and it seeks answers that are just as comprehensive. Good examples of such metaphysical questions are:

Can matter exist without mind? Can mind exist without matter?

What are the essential properties of matter?

Can empty space exist?

Does God exist?

Is the universe homogenous or heterogenous?

Is the future of the universe strictly determined?

Is there free will?

Does the universe have boundaries in space and time?

How a community answers questions like these establishes their overall *worldview*. A worldview is a community’s unique constellation of metaphysical components, and it characterizes how they understand the general features of the world taken as a

whole. You will notice that the answers to these questions are the *metaphysical components* that we have used to characterize each of the four worldviews we have studied. Moreover, each worldview's unique set of metaphysical components acts as something of a "signature". For instance, any community that accepts hylomorphism, pluralism, plenism, and teleology shares the same basic worldview as the Aristotelian-Medieval, which is substantially different from the Contemporary worldview.

Explicit and Implicit

How do we determine a community's answers to these metaphysical questions? Or, put another way, how would someone studying the history of science identify the metaphysical components of a community's mosaic?

The task can be relatively straightforward in communities who explicitly include metaphysics in their mosaics. That is, in some communities, metaphysics is recognized as a science alongside other sciences like physics or mathematics, and their attempt to articulate their own worldview is brought into direct conversation with these other sciences. Again, that is to say, in some communities their metaphysical commitments are *explicit*, or outwardly expressed and articulated, as part of the science of metaphysics.

The Aristotelian-Medieval and Cartesian mosaics are excellent examples of mosaics where the metaphysical components are *explicitly stated*. What evidence do we have for this? In the case of the Aristotelian-Medieval community, it is well-established that metaphysics was essential to the curricula of the typical medieval or early modern university. Since metaphysics was taken as a legitimate discipline that attempts to unearth the most general features of the world, and since its findings were accepted, metaphysics can be said to be part of their scientific mosaic. After progressing through three foundational disciplines of the so-called *trivium* – grammar, logic, and rhetoric – students working towards a degree would then progress from physics (or natural philosophy), to metaphysics, and then to moral philosophy. As we have also seen, such metaphysical components as *teleology* and *hylomorphism* were essential for understanding even the basics of Aristotelian-Medieval physics. Similarly, because most of the features of the Cartesian mosaic blossomed from the works of René Descartes himself, and his foundational texts explicitly establish such metaphysical views as *dualism* and the essential properties of matter, these metaphysical commitments were made explicit in texts of Cartesian science. Similar to medieval universities, universities which adopted the Cartesian mosaic also explicitly included metaphysics as part of their curriculum.

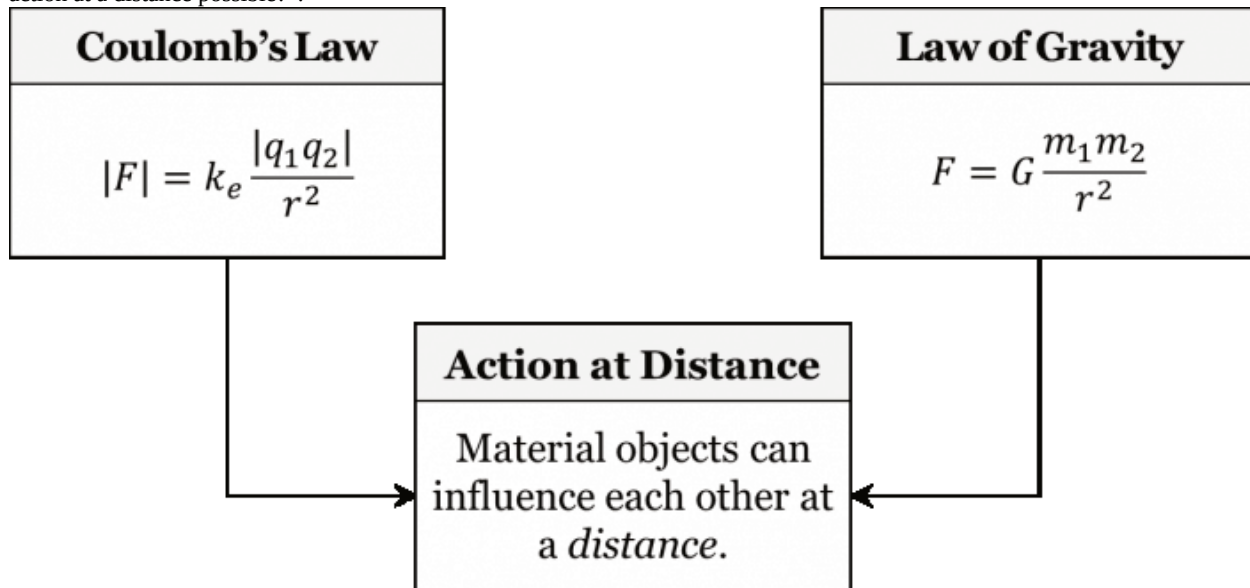
In cases where metaphysics is *explicit* in the mosaic, determining that mosaic's metaphysical components can be as simple as consulting textbooks and scholarly writings from that community. This being said, even in communities where metaphysics is explicitly part of the mosaic, there might still be some metaphysical components that are *not* explicitly formulated, and that we would need to reconstruct ourselves by analysing relevant materials. That is, sometimes a community's metaphysical views are *implicit* in their mosaic. In order to understand how a community would answer metaphysical questions which they themselves never openly asked or sought to answer, we need to do the work of explicating the metaphysical assumptions that are "folded into" their other accepted theories.

For some communities, *most* of their mosaic's metaphysical components are implicit, or simply assumed. In these cases, the task of uncovering the metaphysical foundations of their worldview falls to historians and philosophers of science like us. For example, by the early nineteenth century, the mosaics of most communities no longer contained explicit statements of their metaphysics. Thus, we can hardly find any explicit metaphysics in a typical Newtonian mosaic of the nineteenth century. The same goes for our contemporary mosaic: consult any of your science textbooks, and chances are you wouldn't find any explicit statement of such metaphysical components as probabilistic determinism or materialism. Again, this does not mean that these mosaics have no metaphysical components whatsoever. It simply means that these communities no longer consider the task of explicating those metaphysical components a part of science itself. This can be due to a number of factors, but the most straightforward is a change in their method's *demarcation criteria* (see chapter 6).

But as you've seen in chapters 9 and 10, we were still able to explicate a good number of metaphysical components from both the Newtonian and Contemporary communities' mosaics despite these challenges. In order to do so, we looked at sets of accepted scientific theories in each mosaic and determined what assumptions all those theories seemed to share. Remember, metaphysics (as we've defined it) is a set of views of the world taken as a whole. Therefore, our goal when looking at these theories is to see whether the theories themselves point to a single answer to one of those general metaphysical questions we asked above. In the case of the Newtonian community, for instance, we could ask: Based on the fact that the Newtonians accepted Coulomb's *law of electrostatics*, and Newton's *law of universal gravitation*, how would they answer the question "Is action at a distance possible?"

In the case of Coulomb's *law of electrostatics*, which describes the force between two particles charged with static electricity, the attractive or repulsive force ($|F|$) between the two particles (whose charges are represented by q_1 and q_2) at any distance (r) can be calculated accurately without formally representing *any mediating agent between them*. Indeed, while Coulomb's constant (k_e) might change its value depending on the medium between the point charges, the formula works just as well when *there is no medium between them* (i.e. a vacuum). This is similar in Newton's *law of gravity*, where F is the force of gravity between two objects with masses m_1 and m_2 at a certain distance r , and G is Newton's gravitational constant. The force due to gravity alone can be calculated *without accounting for any material between the objects*. These two major theories, accepted by a typical nineteenth-century Newtonian community, both postulate that the electrostatic and gravitational forces exist even if there is no medium between the objects in question. Since both of these theories share this assumption, we can deduce a general metaphysical

principle from it: that action at a distance is possible. Thus, we can conclude that Newtonians would answer yes to the question “Is action at a distance possible?”.



Sometimes making such generalizations presents considerable difficulties. Recall that in chapter 10 we noted the reasons why a number of metaphysical components of the Contemporary worldview are difficult to identify; it's unclear whether our scientific community accepts *vacuism*, *action at a distance*, or an *infinite universe*. While you can revisit chapter 10 to see precisely why we have trouble pinpointing the metaphysical components that are part of today's worldview, we can also frame this difficulty in terms of the topic of this chapter: without the benefit of a broader historical perspective, it is difficult to ascertain whether some theories today are *accepted*, and this subsequently makes it difficult to *explicate* the *implicit assumptions* folded into today's theories. Hopefully with more time and research we will soon be able to understand where today's scientific community stands with regard to these particular metaphysical questions.

In any event, we can conclude that regardless of whether the metaphysical components are made explicit or left implicit, they do shape the overall worldview of a community. This is all to say that metaphysical components – both implicit and explicit – seem to play an ineliminable role in the process of scientific change itself, whether we want to admit it or not!

Metaphysics and Empirical Science

So far, we've discussed what metaphysics is and what a mosaic's metaphysical components are, how they contribute to a community's worldview, and how we might uncover these components when they are explicit and implicit. In order to shed some light on why metaphysics might not be explicit in all mosaics, we will now briefly explore the possible relationships that exist between metaphysics and the rest of the empirical sciences. Let us start by asking: can there even be any legitimate metaphysics? In other words, can we ever know anything about the world taken as a whole?

Historically, there have been philosophers who explicitly denied that metaphysics is possible at all. For instance, by the turn of the 20th century, some scientists and philosophers – known as the logical positivists – argued that *only science* could be legitimately considered knowledge, since *only science is capable of empirical verification*. Scientific theories are verifiable because they refer to things that can, at least in principle, be observed or experimented with. This goes for all empirical sciences – both natural and social. As for metaphysics, according to logical positivists, it doesn't refer to anything that can be observed, even in principle. As such, all metaphysical “isms” are, in the logical positivist view, complete nonsense. The positivists therefore not only rejected metaphysics as a science, but they considered it a threat to knowledge itself. We can call this conception *anti-metaphysical science*. This idea can be traced back to the eighteenth century German philosopher Immanuel Kant, who famously argued that we can never know things *as they actually are* (what he called *noumena*) but could only have certain knowledge about the things *as they appear to us* (what he called *phenomena*). But knowing the general features of the mind-independent world is precisely the task of metaphysics. Thus, Kant concludes that there can be no metaphysics whatsoever.

But regardless of whether or not metaphysics can be said to be a legitimate source of knowledge, it is a historical fact that metaphysical assumptions often find their way into scientific mosaics. Recall the metaphysical components of any one of the four worldviews we have discussed. Since any major worldview that we consider contains plenty of metaphysical components, this suggests that metaphysics has a role to play in science. Now, the question is whether metaphysics and empirical science merely coexist in a mosaic without influencing one another, or whether they, in some way, shape one another:

Do metaphysics and empirical science depend on one another, or are they completely independent?

Some philosophers seem to have accepted that empirical science and metaphysics are utterly independent from one another, while others have maintained that they are intimately related. Let us consider both of these views in turn.

According to some philosophers, both empirical science and metaphysics have their own distinct domains and are independent of one another. On this view, scientific theories merely provide us with knowledge of this or that aspect of the world, without presupposing any single view on the world taken as a whole. In other words, physical, biological, or sociological theories do not bring any implicit metaphysics with them. For instance, Newton's law of universal gravitation, it is argued, merely captures a specific relation between the force of gravity, the masses of bodies and the distance between them. It says nothing about the possibility or impossibility of action at a distance. Similarly, according to the champions of this view, our contemporary quantum physics only tells us how subatomic particles behave, but it says nothing about any general features of matter such as its dual wave-particle nature. The motivation underlying this view is the fact that the same theory can receive many different metaphysical interpretations. Quantum mechanics has famously received a whole variety of different interpretations, each accepting the same set of equations, but differing drastically in their metaphysical views about the world and its general features. For instance, while some scientists and philosophers interpret the equations of quantum mechanics as suggesting the view of probabilistic determinism, others argue that the same equations can be interpreted in line with the view of strict determinism. Thus, it is argued, the question of the general features of the world are to be settled by metaphysicians alone without much reliance on the findings of empirical science, for the latter can often be interpreted in multiple ways and are therefore no guide in adjudicating between different metaphysical conceptions. In short, these philosophers champion the *independence of science and metaphysics*.

What this position seems to ignore is that often scientific theories already come with a certain metaphysical interpretation built into their very core. It is true that the same set of equations can be interpreted very differently depending on the meaning we assign to their variables. However, it would be wrong to argue that all scientists care about is their equations. In fact, empirical theories are not just their equations, but the equations *understood in a certain way*. For example, the idea of action at a distance was an essential part of the accepted Newtonian view of the force of gravity. Without such an interpretation, the equation would remain purely mathematical and would belong to *formal* rather than *empirical* science. Thus, in order for the law of universal gravitation to be an empirical statement about the world we inhabit, it has to have *some* interpretation. Can we, in principle, provide the law of universal gravitation with a different interpretation? Yes, we can. In this the proponents of the *independence of science and metaphysics* are correct. Yet, historically, empirical theories are always accepted with some interpretation, otherwise they would be pure equations and would remain *formal* theories, such as those of mathematics and logic. In addition, in many instances, separating a theory from its metaphysical assumptions is not only difficult but also ignores the context of the specific mosaic in which it was accepted historically. Thus, a typical Newtonian community would accept the law of universal gravitation as stating that objects with mass attract each other at a distance via the force of gravity. In other words, the law itself and its metaphysical interpretation were accepted as an inseparable package. Therefore, the question is not how a certain equation can be interpreted *in principle*, but how it *has* actually been interpreted and accepted by different communities at different times.

The idea that empirical science and metaphysics are intrinsically linked has had its champions since antiquity. Aristotle famously believed that natural philosophy (physics) and metaphysics are part of the same enterprise of understanding the world. According to Aristotle, any inquiry begins with the identification of the most common features of the object under study – the so-called *first principles*. Once these most common features are intuitively grasped and formulated as axioms, then one can proceed with deducing the more specific features of an object under study from these general axioms. This is a natural consequence of his method of intuition schooled by experience. It is for this reason that Aristotle used the term *first philosophy* to denote what we call *metaphysics*. It is the first, because that is where any scientific inquiry starts. For example, first we need to appreciate that every substance can be analytically decomposed into its form and matter; this is the idea of hylomorphism we discussed in chapter 8. Once hylomorphism is grasped and taken as an axiom, we can then deduce that humans can be thought of as having both a body (matter) and a soul (form). We can then apply the idea of hylomorphism to all living things. Finally, we can also explore the relations between form and matter in all sort of inorganic substances, such as rocks or chairs. This position can be titled *metaphysical science*, as it assumes that specific scientific theories can be somehow deduced from general metaphysical principles, such hylomorphism, plenism, or pluralism.

It is safe to say that the idea of metaphysical science was also accepted by Descartes and his followers. Recall how Descartes starts by identifying what he considered to be the most general intuitive truths, such as the idea that the doubting mind necessary exists (the cogito), that God exists, or that extension is the indispensable attribute of matter. Having established these and other fundamental principles as the axioms of his system, Descartes then attempts to deduce from them a whole system of the world, including his physics, optics, biology, psychology, and so on. Once again, this is yet another example of the conception we called *metaphysical science*.




Does the idea of metaphysical science hold water? Is it even possible to deduce specific scientific theories from general metaphysical assumptions? Suppose we were to somehow arrive at the metaphysical idea of action at a distance. Could we possibly deduce any specific physical law from this metaphysical principle? While the idea of action at a distance is implicit in, say, the law of universal gravitation, it is strictly speaking impossible to deduce the law itself merely from the idea of action at a distance. Indeed, there is nothing in the idea of action at a distance that suggests any specific relation between masses, distance, and the force of gravity. In fact, an infinite number of hypothetical physical laws are compatible with the idea of action at a distance. For example, if we were to replace the square of the distance in the original law with the cube of the distance, the law would still imply action at a distance. So how can any specific scientific law be deduced from a general metaphysical principle? The short answer is: it cannot.

This is the reason why most scientists and philosophers nowadays are very sceptical about the prospects of metaphysical science. The approach that seems to be accepted nowadays is the one that can be called *scientific metaphysics*. One good illustration of the idea of scientific metaphysics is the approach taken by this textbook: that the (general) metaphysical components of worldviews can be deduced from (particular) theories in a mosaic. In other words, metaphysical components are *consequences of*, and therefore depend upon, the empirical theories accepted in a mosaic. This constrains metaphysical speculation to within the realm of our accepted empirical theories and acknowledges that particular empirical theories always have broader ramifications for how a community conceives of the world taken as a whole.

Extracting the metaphysical assumptions implicit in our accepted theories is not always an easy task. It is also the case that sometimes accepted scientific theories do not yield one specific metaphysics but offer room for interpretation and discussion. This explains why there is so much debate on how exactly this or that aspect of quantum mechanics is to be understood and what it means for our metaphysics. In some cases, there is some leeway for metaphysical interpretations, and metaphysics is not always strictly dictated by the accepted scientific theories. Thus, metaphysics can at times be relatively independent from accepted scientific theories.

Nevertheless, often accepted theories come with very distinct metaphysical views implicit in them. In such cases, there is virtually no room for disagreement. Even in those cases, when accepted empirical theories do not strictly dictate one single metaphysics, they greatly constrain the set of possible metaphysical interpretations. Most of the metaphysical assumptions (the “isms”) of the four mosaics we have discussed in chapters 7-10, were in one way or another implicit in the empirical theories accepted in those four mosaics. However, we also noted that sometimes the accepted empirical theories do not come with any clear-cut answer to the fundamental metaphysical questions. Recall that we’ve had a hard time pinning down the contemporary attitudes towards the questions of plenism vs. vacuism, infinite vs. finite universe, and action by contact vs. action at a distance.

The differences between these four approaches concerning empirical science and metaphysics can be summarized in the following table:

 Can there be any metaphysics as a legitimate field of inquiry?			
No		Yes	
Anti-Metaphysics	 Do metaphysics and empirical science depend on one another in any way?		
	No		Yes
	 Are scientific theories deducible from metaphysical theories?		
	No		Yes
Independence of Science and Metaphysics		Scientific Metaphysics	Metaphysical Science

We believe that the view of scientific metaphysics squares better with the actual practice of science and philosophy. That’s why we adopted this approach in this textbook.

What can change a worldview?

Once again, a worldview is the unique constellation of metaphysical components accepted by a community. It is the particular way that one community would answer general questions about the world taken as a whole. We began this chapter by noting that, while mosaics seem to be in a state of near-constant change, worldviews have historically enjoyed a degree of stability. For instance, we noted that while *any* change of *any* theory or method constitutes a change of *the mosaic*, those same changes might have little-to-no effect on the worldview. So, what *can* change a worldview?

Perhaps changes in *ontology* are sufficient to change a worldview. Back in chapter 3 we defined ontology as *a set of views on the types of entities and interactions that populate the world*. Strictly speaking, a community’s ontology is part of its metaphysics, since ontology accounts for what entities and relations *actually exist in the world*. We know that communities employing the

HD method don't take changes to their accepted ontology lightly, only accepting such theories if they have met the stringent requirement of confirming their novel predictions.

In February 2013, the existence of the Higgs boson (the elementary physical particle which helps to explain why some other elementary particles have mass) had still not been experimentally confirmed. Thanks to experiments at the Large Hadron Collider near Geneva, Switzerland in March 2014, the relevant novel predictions were confirmed, leading to the acceptance of the Higgs boson. The result was a change to our contemporary ontology: the scientific community then began to accept that the Higgs boson exists alongside the other elementary particles, exhibiting the wave-particle duality typical of all such elementary particles (see chapter 10).

Now, since ontology is a part of metaphysics, technically any change to a community's ontology changes its metaphysics. As such, the addition of the Higgs boson to the set of elementary particles was, indeed, a change both to contemporary ontology *and* metaphysics. But, interestingly, the addition of the Higgs boson to the set of elementary particles did *not* lead to a change in the general metaphysical component of *wave-particle duality*. So even if changes to the mosaic have ontological or metaphysical implications, these changes might not have any significant effect on the mosaic's metaphysical components. For instance, if we were to someday discover that two of the elementary particles are actually one and the same, we will have effectively *lessened* the number of entities in our ontology. But – again – this would not affect *wave-particle duality*. That is, the question “What are the essential properties of matter?” would *still have the same answer*.

But what if some of the novel predictions of *superstring theory* were confirmed? This currently pursued theory – which promises to unite general relativity and quantum physics – postulates that all elementary “wavicles” are really the effects of vibrations of entities called *superstrings*. These superstrings vibrate in multiple dimensions, which accounts for many of the qualities of wavicles, including their wave-particle duality. If scientists were able to somehow devise a way to accurately measure at the Planck scale, or to experimentally detect certain predicted supersymmetries, superstring theory could very well become accepted. Unlike the acceptance of the existence of the Higgs boson, however, superstring theory would likely affect the scientific community's answer to the question “What are the essential properties of matter?”. With the inclusion of superstring theory in the mosaic, the metaphysical component *wave-particle duality* could be replaced by something like *multidimensional strings*.

A worldview will only change if its metaphysical components change, and metaphysical components only change if the mosaic contains theories that – taken together – give us a new answer to fundamental metaphysical questions. We therefore suggest that those fundamental metaphysical questions – like “What are the essential properties of matter?” or “Can empty space exist?” – are our best guides for tracking when metaphysical components change. When considering changes in a mosaic, we can ask ourselves: *does this change to the mosaic represent a new way the community answers a fundamental metaphysical question?* If the change to the mosaic does *not* lead to a new answer, then it hasn't replaced a metaphysical component, and therefore hasn't changed the worldview. If the change to the mosaic *does* lead the community to a new answer, then it likely changes a metaphysical component, and may contribute to a change in the worldview.

Clearly, it's hard to say beforehand what specific changes in our ontology will or will not actually end up affecting our worldview. Grand transitions in worldviews are often best seen at a certain distance, i.e. with some passage of time. Understandably, scientific textbooks and encyclopedias are not going to change overnight, even after a fundamental change in worldview. That being said, we can safely assume that any such change is going to involve some underlying change in the accepted metaphysical components.

To review, we identified the four worldviews focused on in this textbook by their unique constellations of metaphysical components. These metaphysical components are the scientific community's answers to questions about the world taken as a whole. We argued that despite the popularity of the idea of the *independence of science and metaphysics* as well as the idea of *metaphysical science*, a version of *scientific metaphysics* – with metaphysical components deriving from the theories in a community's mosaic – best accounts for the changes in the history of science. While many scientific changes might indeed constitute changes to scientific communities' metaphysics and ontologies, only changes which trigger a new answer to a fundamental metaphysical question signal a change in worldview.

Conclusion

The philosopher of science Imre Lakatos once said, “Philosophy of science without history of science is empty; history of science without philosophy of science is blind”, and we agree with him. We began this textbook with some of the most important topics in the philosophy of science, moved to an exploration of the dynamic evolution of four scientific worldviews, and in conclusion have brought this history back into explicit conversation with our philosophical investigations. Throughout the text we aimed to ensure that our approach to the philosophy of science was always informed by the history of science, and that our presentation of science's history remained clearly and explicitly guided by our theoretical commitments. Far from blind, our historical approach is lit by the lamp of philosophy. Far from empty, our philosophy has drawn from the deep well of science's history. Together, these introductions to the history and philosophy of science have equipped you to voyage further into the study of science, or simply to appreciate our scientific knowledge all the better.

But introductory textbooks, by definition, are unable to communicate all the nuance or detail that their topics demand. This is especially true of a textbook introducing a topic so broad as the history and philosophy of science! We therefore see it as our

responsibility to make some clarifications about, and indicate some limits of, this textbook's presentation of the material, and to point the way towards promising avenues of future research.

As we had an opportunity to reiterate in this chapter, individual mosaics are vastly diverse and in a state of near-constant change. The complexity of an actual scientific mosaic is therefore staggering, and the relatively simple historical snapshots that we were able to present of individual mosaics and their metaphysical components were only possible due to the incredible work done by professional historians of science. Their research involves synthesizing information from their subject's political, social, environmental, intellectual, and technological contexts – typically from *primary texts* – to produce a coherent, compelling picture of their historical subject. Our simplifications and generalizations were only possible due to other historians' labour and clarity. We limited our historical presentation principally to scientific mosaics and their changes with an emphasis on physical theories. While this narrow and unique focus was helpful and necessary, it only begins to scratch the surface of an immensely vast and limitlessly rich area of study.

In the history section of this textbook we concentrated on the Aristotelian-Medieval, Cartesian, Newtonian, and Contemporary worldviews. It was important that we focus on these four worldviews for a number of reasons. First, these worldviews and their transitions have been the central focus of the History and Philosophy of Science (HPS) *as a field of study, or discipline*. Those of you who choose to further explore this discipline will be expected to at least have a basic understanding of the theories and methods of these communities, as well as the major figures (like Aristotle, Descartes, Newton, or Einstein) who played important roles in those communities. Second, because these worldviews have been a central focus of HPS, they are arguably also the most thoroughly researched mosaics in science's history. Drawing from this research ensured that even our historical abstractions and generalizations were grounded in deep, nuanced traditions of scholarship. Third, because these worldviews have been so thoroughly researched, we were also able to learn from the mistakes of previous generations of scholars and to correct certain myths or narratives which we may have received in popular culture or from a more cursory understanding of science's history. As a simple example: many introductions to science's history fail to recognize the immense geographic reach and intellectual impact of the Cartesian worldview. We have placed the Cartesians squarely between the Aristotelian-Medieval and Newtonian worldviews, where they belong. Fourth, these worldviews are distinct from one another, transitioned clearly from one to the other, and were ultimately precursors to our Contemporary, global scientific worldview. All of these factors were important for understanding both the *mechanism* by which science changes, and for tracing and appreciating the genealogy of our current mosaic and worldview.

But the textbook's focus on these four worldviews might lead to undesirable misinterpretations or misunderstandings. For instance, we have *not* focused on these four worldviews because they are the *only* scientific worldviews that have existed. It is obvious that there have been numerous other scientific communities – with *different worldviews* – spread over the globe, and over humanity's long history. Specifically, there have been scientific communities in other parts of the world, including Africa, Asia, Oceania, and the Americas. The mosaics produced and developed by communities in these geographic regions have been as diverse as the cultures from which they arise: accepting different theories, employing different methods, and very often having different metaphysical components. Some of these scientific communities likely developed on their own, others are historical branches of older communities and their mosaics, often contributing to lesser and greater degrees of similarity, respectively. We should therefore appreciate that the mosaics in different geographic regions and different historical periods can have drastically different accepted theories and employed methods. Consider for instance, a typical Medieval-Arabic scientific mosaic. While it would share many of its elements with that of the Aristotelian-Medieval mosaic accepted in many Christian regions, it would also considerably differ from the latter specifically in its theology. A case can be made that accepted Islamic theology shaped the respective methods of theory acceptance employed in the Medieval-Arabic mosaic, which were in many respects different from those employed in many regions of Christendom. The aforementioned importance of these four worldviews notwithstanding, they are chosen in this textbook to illustrate the *laws of scientific change* in action and serve as the entryway into a more globalized perspective of HPS no longer bound to a Eurocentric narrative.

Rather than let you wander off into the frontiers of the history and philosophy of science, we think it is responsible to gesture to a few of the most prevalent fields exploring topics touched upon in this text and give you a taste of what to expect in each of them. The fields/disciplines are: the history of science; science & technology studies (STS); philosophy of science; and integrated HPS.

Thanks to an ever-widening global perspective and the power of recent research and networking technologies, *the history of science* offers the opportunity to dig deeply into the stories of individuals, practices, institutions, technologies, and mosaics of an ever-growing list of epistemic communities. If you ever wondered about particular sciences that we were not able to explore in great detail – like chemistry, evolutionary biology, or medicine – chances are that excellent research is being done on the topic.

Science & technology studies, often abbreviated STS, approaches science with research methods and theory grounded in anthropology and sociology. STS focuses on the social, political, and technological factors which shape science, and the ways that science, in turn, is used to shape those societies, political landscapes, and technologies. If you were ever wondering about the origins or implications of the science we covered, it might be worth looking into some STS scholarship.

In addition to the main topics we introduced in this textbook, *the philosophy of science* is engaged in the ongoing project of seeking clarity and greater understanding of our scientific terminology and practices. There are many different ways of doing philosophy, including typical focuses like metaphysics, epistemology, and ethics. But we guarantee that if you pick any one of those focuses and combine it with your favourite science (say, the epistemology of biology, or the metaphysics of economics), the field will probably exist, and will fascinate you.

Echoing Lakatos' quote from above, we maintain that great philosophy of science is historical, and great history of science is philosophical, even if you are predisposed or drawn to one or the other. *The integrated history and philosophy of science*

consciously attempts to tie these approaches together in a way similar to how we have in this text, allowing the best history of science to shape our philosophy and vice-versa. The growing community of *scientonomy*, for instance, works on developing and refining the laws of scientific change, and on reconstructing the mosaics of diverse communities through history. By uniting the history and philosophy of science in this way, scientonomists hope to reconstruct a historical “tree of knowledge” (akin to the evolutionary “tree of life”) embodying the global vision of HPS described earlier.

In only a few minutes you won’t be reading this textbook anymore. Perhaps soon you will have an opportunity to *actually* lay on your back in some soft grass, far from city lights, and stare into the vast, dark night sky. We hope that as you marvel at the beauty of the universe, you will now also be able to appreciate the community of wonderers who came before you: women and men whose eyes beheld the same stars, and whose wonder inspired the creation of magnificent mosaics to which they contributed. In the same way that those stargazers and astronomers from ages past were driven by curiosity to understand the dynamisms and nature of the heavens, we hope that you have become curious about science’s history and philosophy and are driven to understand it better. We sincerely encourage you to follow that curiosity.