

A System of Logic Ratiocinative and Inductive

Presenting a Connected View of the Principles of Evidence and the Methods of Scientific Investigation

John Stuart Mill

Copyright ©2010–2015 All rights reserved. Jonathan Bennett

[Brackets] enclose editorial explanations. Small ·dots· enclose material that has been added, but can be read as though it were part of the original text. Occasional •bullets, and also indenting of passages that are not quotations, are meant as aids to grasping the structure of a sentence or a thought. Every four-point ellipsis indicates the omission of a brief passage that seems to present more difficulty than it is worth. Longer omissions are reported between brackets in normal-sized type.

First launched: September 2102.

Contents

Book III; Induction	139
Chapter 1. Preliminary remarks about induction in general	139
Chapter 2. Inductions improperly so called	141
Chapter 3. The ground of induction	151
Chapter 4. Laws of nature	155
Chapter 5. The law of universal causation	160
Chapter 6. The composition of causes	182
Chapter 7. Observation and experiment	186
Chapter 8. The four methods of experimental inquiry	191
Chapter 9. Examples of the four methods	202
Chapter 10. Plurality of causes, and the intermixture of effects	214

Chapter 11. The deductive method	225
Chapter 12. Explaining laws of nature	231
Chapter 13. Examples of the explanation of laws of nature	236
Chapter 14. The limits to the explanation of laws of nature. Hypotheses	241
Chapter 15. Progressive effects. The continued action of causes	253
Chapter 16. Empirical laws	258
Chapter 17. Chance and its elimination	263
Chapter 18. The calculation of chances	269
Chapter 19. Extending derivative laws to adjacent cases	278
Chapter 20. Analogy	282
Chapter 21. Evidence for the law of universal causation	286
Chapter 22. Uniformities of coexistence that don't depend on causation	296
Chapter 23. Approximate generalisations. Probable evidence	303
Chapter 24. The remaining laws of nature	310
Chapter 25. The grounds of disbelief	322

Glossary

circumstances: In Mill's usage, the 'circumstances' of a given experiment are all the details of what is the case when the experiment is performed—not only in the environment but also in the experiment itself.

coextensive: 'Law L is coextensive with field F' means not merely that nothing in F is a counter-example to L but that everything in F is an example of L.

coincidence: In Mill's usage, the coincidence of two events is simply their occurring at the same time (and usually, perhaps, in the same place). What you and I call a 'coincidence' is the occurring together of two events that have no causal relation to one another; in Mill's terminology that is a 'casual coincidence'. (**Be alert to the difference between 'casual' and 'causal', both of which occur often in this work.**) On page 328 he introduces a different sense of 'coincidence', which he explains there.

collocation: Arrangement in space; structure. When in the footnote on page 167 Mill explains that by 'the constitution of things' he means 'ultimate laws of nature' and not 'collocations', what he is rejecting is the use of 'constitution' to mean 'how things are arranged, structured, in space'. On page 231 we learn that items entering into a 'collocation' can include powers = forces as well as physical things.

concomitant: 'Of a quality, circumstance, etc.: occurring along with something else, accompanying' (OED).

concurrence: The concurrence of several events is their occurring together, usually meaning at the same time and in roughly the same place. From Latin meaning 'run together'.

connote: To say that word W connotes attribute A is to say that the meaning of W is such that it can't apply to anything

that doesn't have A. For example, 'man' connotes humanity.

corpuscule: An extremely small bit of matter—far too small for us to be able to pick it out visually. Adjective **corpuscular**.

cultivation: Carefully developing (a skill or habit), analogous to cultivating roses or cabbages. On page 178 the two are linked metaphorically.

data: Until about the middle of the 20th century 'data' was the plural of 'datum'. Since then it has become a singular mass term, like 'soup'.

deus ex machina: Latin literally meaning 'a god out of a machine', referring to the use of theatrical machinery to float a god onto the stage to make everything come right at a crucial point in a drama. Nearly always the phrase is used metaphorically, to refer to some problem-'solving' item that a theorist introduces in a suspiciously convenient way and without good reasons. On page 179 Mill uses the phrase in both ways at once: the suspiciously convenient item that he refers to is literally God.

efficient cause: This is an Aristotelian technical term. The •formal cause of a coin is its design, the plan according to which it was made; its •material cause is the stuff it is made of; its •final cause is its purpose, namely to be used in commerce; and its •**efficient cause** is the action of the die in stamping the coin out of a metal sheet. So the efficient cause is what you and I would call, simply, 'the cause'. But on page 162 and thereafter Mill is clearly using 'efficient cause' to mean something like: real, metaphysically deep, empirically inaccessible causes, as distinct from the mere

orderly event-followings that are the only causes Mill believes in.

fact: In Mill's usage a 'fact' can be a state of affairs or an event or a proposition (not necessarily true) asserting the existence of a state of affairs or event. In the present version, no attempt is made to sort all this out.

***inductio per enumerationem simplicem*:** Latin meaning 'induction by simple enumeration'. This comes from Bacon, who used it meaning something like 'reaching a generalisation by simply looking at positive instances and naively failing to look for counter-instances or complications'. Mill seems to be using it that way too.

irritability: Proneness to respond to physical stimuli.

luminiferous ether: The ether was a supposed finely divided or gaseous matter pervading the whole universe; 'luminiferous' means 'light-bearing': it was thought that light consisted of some kind of disturbance of the ether.

material: The 'material circumstances' are the circumstances or details that matter. A 'material change' is a change that makes a significant difference.

***mutatis mutandis*:** A Latin phrase that is still in current use. It means '(mutatis) with changes made (mutandis) in the things that need to be changed'.

natural theology: This is theology based on facts about the natural world, e.g. empirical evidence about what the 'purposes' are of parts of organisms etc. In this context, 'natural' is the antonym of 'revealed'.

numeral: A name of a number, usually confined to names like '7' and not like 'seven'. Mill doesn't use the word here, but this version uses it instead of 'name' in some contexts where the topic is obviously names of numbers.

occult: It means 'hidden', but in the early modern period it always carried the extra sense of 'mysterious, out of reach of ordinary understanding' or the like. The statement that gravity is an 'occult force' meant that the ultimate truth about gravity, whatever it is, won't be a part of ordinary physics.

original: Sometimes Mill uses this to mean 'basic' or 'foundational'. An 'original natural agent' (page 170) is a natural cause that wasn't caused by anything we know about. Mill also uses 'primeval' and 'primitive' with the same meaning.

patient: The same Latin words lie behind three contrasts:

- adjectives: 'active' and 'passive'
- abstract nouns: 'action' and 'passion'
- concrete nouns: 'agent' and 'patient'

We don't now use 'passion' to refer to any undergoing or being-acted-on, or 'patient' to refer to anything that is acted on; but until the end of the 19th century both of those uses were current.

***petitio principii*:** A Latin phrase referring to the procedure of offering a 'proof' of P from premises that include P. The English name for this used to be 'begging the question', but that phrase has recently come to mean 'raising the question' ('That begs the question of what he was doing on the roof in the first place.')

popular: It means 'of the people'; in early modern times it usually doesn't mean 'liked by the people'.

precession of the equinoxes: The slow, steady change in the earth's axis of rotation.

principle: In the passage by Whewell on page 145, the phrase 'principle of connection' may mean 'something that *physically connects* them', thus using 'principle' in a sense—now obsolete but extremely common in the early modern

period—in which it means ‘source’, ‘cause’, ‘driver’, ‘energizer’, or the like. It is certainly used in that sense by Mill on page 182 and page 188 and by Powell on page 289.

putrefy: rot; and the rotten state is **putrefaction**.

quadrature of the cycloid: A cycloid is the curve traced by a point on the rim of a circular wheel rolling on a plane surface. That curve and the line of the surface enclose an area; its quadrature is the process of discovering the size of that area.

resolve: To resolve x into y and z is to analyse x in terms of y and z, to show that all there is to x is y and z, or the like. Mill explains this on page 231. The noun is **resolution**.

rigor mortis: Latin for ‘stiffness (or rigidity) of death’. Mill calls it ‘cadaveric rigidity’, but these days the Latin phrase is also the colloquial English one.

sagacity: Here it means something like ‘alert intelligence’.

sui generis: Latin for ‘of its own kind’—not significantly like anything else.

synchronous: Occurring at the same time.

type: ‘the real type of scientific induction’ (page 158) means ‘the central defining paradigm of scientific induction’. Similarly with ‘the type of uncertainty and caprice’ on page 293 and ‘the type of a deductive science’ on page 316.

vera causa: Latin meaning ‘true cause’. A technical term of Newton’s. To say that x is a *vera causa* of y is to say that x is already known about independently of its causing of y, or perhaps (see page 247) that x *could be* known about independently etc.

virtue: power, causal capacity, or the like.

vortex: Descartes’s term for a rapidly rotating collection of fine particles. The plural is **vortices**.

Book III; Induction

Chapter 1. Preliminary remarks about induction in general

§1. We are now approaching what can be regarded as the principal topic in this work—because it is more intricate than any of the others, and because it concerns a process that I have shown in Book II to be the one that the investigation of nature essentially consists in. I showed that all inference, and consequently all proof and all discovery of truths that aren't self-evident, consists of **inductions and the interpretation of inductions**—i.e. that all our knowledge that isn't intuitive comes from that source. So it has to be accepted that the main question of the science of logic—the question that includes all others—is

What is induction? and what conditions make it legitimate?

Yet professed writers on logic have almost entirely ignored this question. Metaphysicians haven't altogether neglected its broad outlines. But they haven't known enough about the processes by which science has actually succeeded in establishing general truths, so that their analysis of the inductive operation, even when perfectly correct, hasn't been specific enough to be made the foundation of practical rules that could serve •induction itself in the way the rules of the syllogism serve •the interpretation of induction. As for those who have brought physical science to its present state of improvement, never until very recently have they tried seriously to philosophise on the subject; they haven't regarded their way of arriving at their conclusions—as distinct from the conclusions themselves—as worth studying. •It's a pity, because •all they needed to do to get a complete theory of the process was to focus on the methods that *they*

had been using, and to generalise these and adapt them to all sorts of problems,

§2. For the purposes of the present inquiry, **induction** can be defined as **the operation of discovering and proving general propositions**. As I have already shown, the process of indirectly ascertaining individual facts... is a form of the very same process, because (a) general facts are merely collections of particular facts, definite in kind but indefinite in number; and (b) whenever the empirical evidence justifies us in drawing a conclusion about even one unknown case, it would also justify us in drawing a similar conclusion regarding a whole class of cases. The inference either •doesn't hold at all or •holds in all cases of a certain description—all cases which, in certain definable respects, resemble those we have observed.

If I'm right in maintaining that the principles and rules of inference are the same whether we are inferring general propositions or individual facts, then a complete logic of •the sciences would also be a complete logic of •practical affairs and common life. An analysis of the process by which general truths are arrived at is virtually an analysis of all induction whatever. Why? Because in *any* legitimate inference from experience, the conclusion could legitimately be a general proposition. Whether we're inquiring into a scientific principle or an individual fact, and whether we proceed by experiment or by ratiocination, every link in the chain of inferences is essentially inductive, and the legitimacy of the induction depends in both cases on the same conditions.

When a practical inquirer (e.g. an advocate or judge) is trying to ascertain facts for the purposes not of science but of practical affairs, the principles of induction won't help him with his chief difficulty. It lies not in making his inductions, but in the selection of them—choosing from among all general propositions ascertained to be true the ones that provide marks by which he can trace whether the given subject of study does or doesn't have the predicate in question. When an advocate is arguing a doubtful question of fact before a jury, the general propositions he appeals to are mostly in themselves pretty trite, and are assented to as soon as stated; his skill lies in bringing his case under those propositions; in calling to mind any known or accepted maxims of probability that can be applied to the case in hand, and selecting from among them those that are most favourable to the case he is trying to make. His success will depend on his natural or acquired sagacity [see Glossary], aided by his knowledge of the particular subject and of subjects allied with it. Invention, though it can be cultivated [see Glossary], can't be reduced to rule; there's no science to enable a man to bring to mind what he needs when he needs it.

But when he *has* thought of something *x*, science can tell him whether *x* will suit his purpose or not. When the inquirer or arguer is selecting the inductions out of which he will construct his argument, his only guide is his own knowledge and sagacity. But the validity of the argument he constructs depends on principles, and must be subjected to tests that are the same for all kinds of inquiries—whether the result is •to give someone an estate or •to enrich science with a new general truth. Either way,

- (1) The individual facts must be decided on the basis of the senses, or testimony;
- (2) The rules of the syllogism will determine whether the

case really falls within the formulae of the inductions under which it has been successively brought; and finally

- (3) The legitimacy of the inductions themselves must be decided by other rules. . .

. . . and these rules are what I intend now to investigate. In many everyday practical contexts this third part of the operation is its least difficult part; but we've seen that this is also the case in some big scientific fields. I'm referring to the sciences that are principally deductive, especially mathematics, where •the inductions are few in number and so obvious and elementary that they seem not to need any backing from experience, whereas •combining them so as to prove a given theorem or solve a problem may require the utmost powers of invention and contrivance that our species is gifted with.

If you want further confirmation of my claim that the logical processes that •prove particular facts are the very ones that •establish general scientific truths, consider this: In many branches of science there's a need to prove single facts; they're as completely individual as any that are debated in a court of justice, but are proved in the same way as the other truths of the science—without lessening in the slightest the homogeneity of its method. Astronomy provides remarkable example of this. Most of the individual facts on which that science bases its most important deductions—

- the sizes of the bodies of the solar system,
- their distances from one another,
- the shape of the earth, and
- the earth's rotation

—can't be established by direct observation; they are proved indirectly, using inductions based on other facts that we can more easily reach. [Mill cites the example of the discovery of the moon's distance from the earth; two direct observations

(of the moon's relation to two widely separated points on the earth's surface), followed by sheer trigonometry.]

The process by which that individual astronomical fact was ascertained is exactly like those by which astronomy establishes its general truths; and indeed (as I have shown for all legitimate reasoning) a general proposition could have been concluded instead of a single fact. Strictly speaking, indeed, the result of the reasoning is a general proposition; it's a theorem about the distance from the earth of *any* inaccessible object, showing how that distance relates to certain other quantities. The moon is almost the only body whose distance from the earth can really be ascertained in this way, but that's a mere upshot of facts about the other heavenly bodies that make them incapable of providing such data as the application of the theorem requires. The theorem

itself is as true of them as it is of the moon. [Mill has here a footnote responding to criticisms by Whewell of Mill's use of 'induction'. He says that Whewell's preferred sense of the word isn't justifiable by any philosophical arguments] or supported by usage, at least from the time of Reid and Stewart, who are the principal legislators (as far as the English language is concerned) of modern metaphysical terminology.

So we shan't fall into error if in treating of induction we limit our attention to the establishment of general propositions. The principles and rules of •induction as directed to this end are the principles and rules of •all induction; and the logic of science is the universal logic, applicable to all inquiries in which man can engage.

Chapter 2. Inductions improperly so called

§1. Induction, then, is the mental operation by which we infer that what we know to be true in a particular case or cases will be true in all cases that resemble the former in certain assignable respects. In other words, induction is the process by which we conclude that what is true of certain individuals in a class is true of the whole class, or that what is true at certain times will be true in similar circumstances at all times.

This definition excludes from the meaning of 'induction' various logical operations that are quite often regarded as examples of 'induction'.

Induction, as I have defined it, is a process of *inference* from the known to the unknown; so it excludes any process in which the apparent conclusion is no wider than the

premises it is drawn from. Yet the common books of logic present something of this latter kind as the most perfect—indeed the *only* entirely perfect—form of induction! In those books, every process that sets out from a less general and terminates in a more general expression—which admits of being stated in the form 'This and that A are B, therefore every A is B'—is called an induction, whether or not anything is really concluded in it. And the induction is said not to be perfect unless every single individual of the class A is included in the premise, i.e. unless what we affirm of the class has already been ascertained to be true of every individual in it, so that the supposed conclusion is really a mere re-assertion of the premises. If we say 'All the planets shine by the sun's light' because we have observed that

Mercury, Venus, etc. shine by the sun's light; or that 'All the Apostles were Jews' because we know this regarding Peter, Paul, John, and every other apostle—these and their like are called perfect (and the only perfect) inductions. But this is totally different in kind from my kind of induction; it's not an inference from known facts to unknown facts, but a mere short-hand record of known facts. Their 'conclusions' are not really general propositions. In a general proposition the predicate is affirmed or denied of an unlimited number of individuals, namely all that have the properties connoted by the subject of the proposition—*all*, existing or possible, whether few or many. 'All men are mortal' doesn't mean •all now living but •all men past, present and future. When the word's signification is limited so as to make it a name only for each of a number n of individuals, designated as such and (as it were) counted off individually, the proposition, despite its general language ('All the planets...', 'All the Apostles...') is not a general proposition but merely n singular propositions, written in an abridged form. The operation may be useful, as most forms of abridged notation are; but it's not a part of the investigation of truth, though it often has an important role in preparing the materials for that investigation.

Just as we can sum up n singular propositions in one proposition that will be apparently—but not really—general, so also we can sum up n general propositions in one proposition that will be apparently—but not really—more general. Suppose that for each distinct species of animals we establish by induction that every animal of that species has a nervous system, and on that basis assert that *all species of animals have a nervous system*. This looks like a generalisation, but in fact it merely affirms of •all what has already been affirmed of •each, so it tells us nothing that we didn't already know. This 'conclusion' means the

same as 'All *known* species of animals have a nervous system'. Don't confuse this case with the following quite distinct one. Our observations of the various species of animals have revealed to us a law of animal nature, putting us in a condition to say that a nervous system will be found even in species of animals that haven't yet been discovered. This is indeed an induction, in which the conclusion is a general proposition containing more than the sum of the special propositions from which it is inferred. The difference between these two is further marked by the fact that the latter of them—the genuine induction—could be legitimate even if we hadn't examined every single known species of animals. . . . Returning to the earlier example, think about the difference between these;

'All *the* planets shine by reflected light.'

'All *planets* shine by reflected light.'

The latter is an induction; the former is not. . . .

§2. Several mathematical processes should be distinguished from induction, because they are often (wrongly) called by that name, and share something important with genuine inductions, namely leading to conclusions that really are general propositions. For example, when we have proved that a straight line can't meet a •circle at more than two points, and then successively prove the same thing of the •ellipse, the •parabola, and the •hyperbola, we can lay it down as a universal property of *all conic sections*. The distinction drawn in the two previous examples has no place here because there's no difference between 'all *known* conic sections' and 'all conic sections', as a cone demonstrably can't be intersected by a plane except in one of these •four lines. So we can hardly deny that the proposition arrived at is a generalisation, because there's no room for any generalisation beyond it. But there's no induction because

there's no inference; the conclusion is a mere summing up of the content of the four propositions from which it is drawn. The proof of a geometrical theorem by means of a diagram (on paper or in the imagination) is a bit like that though not entirely so. As I said earlier, such a demonstration doesn't directly prove the general theorem; all it proves is that the general conclusion asserted in the theorem is true of the particular triangle or circle exhibited in the diagram. But we can see that we could prove it of *any* circle in the same way that we have proved it of that one; so we gather up all the singular propositions that could be thus proved, and embody them in a universal proposition. Having shown that the three angles of the triangle ABC are together equal to two right angles, we conclude that this is true of every other triangle, not •because it is true of ABC but •for the same reason that proved it to be true of ABC. The term 'induction' isn't really right for this, because although its conclusion is really general it isn't believed on the evidence of particular instances. We don't conclude that all triangles have that property because some triangles have; rather, we accept the conclusion on the evidence that was the basis for our conviction in the particular instances.

In some mathematical arguments—so-called 'inductions'—the conclusion does look like a generalisation based on some of the particular cases covered by it. When a mathematician has calculated a sufficient number of the terms of an algebraic or arithmetical series to have ascertained what is called the *law* of the series, he doesn't hesitate to supply any number of the succeeding terms without repeating the calculations. But I take it that he does this only when it is apparent from *a priori* considerations (which could be exhibited in the form of demonstration) that the way each subsequent term is formed from its immediate predecessor is the same as the way each previous term was formed from

its predecessor. There are instances on record of wrong results' being reached when a series was continued without the backing of such general considerations.

Newton is said to have discovered the binomial theorem by induction, specifically calculating that

$$(a + b)^2 = a^2 + 2ab + b^2$$

and that

$$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

and so on, and comparing all those results until he detected the general relation that the general binomial theorem expresses concerning the general form $(a + b)^n$ for all values of *n*. It's likely enough that he did; but a mathematician like Newton, who seemed to *leap* to principles and conclusions that ordinary mathematicians reached only by a succession of steps, certainly couldn't have performed the comparison in question without being led by it to the *a priori* ground of the law; since anyone who understands multiplication well enough to venture on multiplying several lines of symbols at one operation can't help seeing that in raising a binomial to a power the coefficients must depend on the laws of permutation and combination; and as soon as that is recognised the theorem is demonstrated. Indeed, once it was seen that the law prevailed in a few of the lower powers, its identity with the law of permutation would at once suggest the considerations that prove it to hold universally. So even cases like this are only examples of what I have called 'induction by parity of reasoning', i.e. not really induction because it doesn't involve inference of a general proposition from particular instances.

§3. It is really important to clear up a third improper use of the term 'induction', because •the theory of induction has been greatly confused by it, and •the confusion shows up in the most recent and elaborate treatise on the inductive philosophy that exists in our language. The error in question

is that of failing to distinguish •an induction from a set of observed phenomena from •a mere description of them in general terms.

Take a phenomenon consisting of parts that can only be observed separately, as it were piecemeal. When the observations have been made, there's a convenience (amounting for many purposes to a necessity) in getting a representation of the phenomenon as a whole by piecing these detached fragments together. A navigator sailing the ocean meets land; he can't by any one observation determine whether it's a continent or an island; but he coasts along it, and after a week sees that he has sailed completely round it, and then declares it to be an island. There was no particular time or place of observation at which he could see that this land was entirely surrounded by water; he learned this fact by a succession of partial observations, and then chose a general expression—'It's an island.'—which summed up in three words the whole of what he observed during that week. Is there anything in the nature of an *induction* in this process? Did he infer something that hadn't been observed from something that had? Certainly not. He had observed the whole of what the proposition asserts. That this land is an island isn't an inference from the partial facts that the navigator saw in the course of his circumnavigation; it is

- the facts themselves,
- a summary of those facts,
- the description of a complex fact to which those simpler ones are as the parts of a whole.

I don't think there is any difference in kind between this simple operation and the one by which Kepler ascertained the nature of the planetary orbits; and Kepler's operation—or anyway all that was characteristic in it—was no more *inductive* than the navigator's.

Kepler aimed to determine the real path followed by each

of the planets. (Let's take *Mars*, because that was the subject of the two of his three laws that didn't require a comparison of planets.) The only way to do this was by direct observation; and all that observation could do was to ascertain many of the successive places—or rather, apparent places—of the planet. The unaided senses could establish this much:

- The planet successively occupied all these positions, or anyway positions that produced the same impressions on the eye, and
- It passed from one of these to another insensibly, with no apparent break in the continuity.

What Kepler did beyond this was to find what sort of a curve these different points would make if they were all joined together. He expressed the whole series of the observed places of Mars by the general conception of *an ellipse*. This operation was much harder than that of the navigator who expressed the series of his observations on successive points of the coast by the general conception of *an island*. But it's the very same sort of operation; and if the navigator's operation is not an induction but a description, this must also be true of Kepler's.

The only real induction consisted in inferring that because the observed places of Mars were correctly represented by points in an imaginary ellipse, therefore Mars would continue to revolve in that same ellipse; and in concluding that the positions of the planet between two observations must have coincided with the intermediate points of the curve. These were *inferences from* the observations—facts inferred, not facts seen—so they involved genuine induction. But these inferences, far from being a part of what Kepler did, had been conducted long before he was born. Astronomers had long known that the planets periodically returned to the same places. With this established, there was no induction left for Kepler to make; he merely applied his new conception

to the inferred facts as well as to the observed facts. When he found that an ellipse correctly represented the past path, he knew that it would represent the future path; in finding a compendious expression for the one set of facts, he found one for the other. But that's all he found—the •expression only, not the •inference—and this didn't add anything to the power of prediction already possessed.

§4. Whewell has given an apt name, the 'colligation of facts', to the descriptive operation that enables a number of details to be summed up in a single proposition. I fully agree with most of what he says about that mental process, and would gladly transfer all that part of his book into my own pages. But I think he makes one mistake, namely treating this kind of operation as the central, primary kind of induction, presenting the principles of mere colligation as principles of 'induction'. In fact, colligation is not 'induction' at all in the old and accepted meaning of the word.

Whewell maintains that the general proposition that binds together the particular facts and makes them into one fact is not the mere sum of those facts but something more, because it introduces a mental conception that didn't exist in the facts themselves. He writes:

'The particular facts are not merely brought together, but a new element is added to them by the very act of thought by which they are combined. . . . When the Greeks, after long observing the motions of the planets, saw that these motions could be considered as produced by the motion of one wheel revolving inside another wheel, these wheels were creations of their minds added to the facts they perceived by sense. Even if the wheels were no longer supposed to be material, but were reduced to mere geometrical spheres or circles, they were still products of the mind

alone—something additional to the facts observed. The same is the case in all other discoveries. The facts are known, but they are insulated and unconnected until the discoverer supplies from his own store a principle [see Glossary] of connection. The pearls are there, but they won't hang together until someone provides the string.'

In this passage Whewell indiscriminately blends together examples of both the processes that I am trying to distinguish. When the Greeks abandoned the supposition that the planetary motions were produced by the turning of material wheels, and fell back on the idea of 'mere geometrical spheres or circles', more was going on than the mere substitution of an ideal curve for a physical one. There was the abandonment of a •theory, and the replacement of it by a mere •description. No-one would call the doctrine of material wheels a mere description! That doctrine was an attempt to identify the force by which the planets were acted on and compelled to move in their orbits. But when the •materiality of the wheels was discarded and only their •geometrical forms retained, this was a great step in philosophy in which the attempt to *account for* the motions was given up and what was left of the theory was a mere *description* of the orbits. The proposition that

the planets were carried round by wheels revolving inside other wheels

was replaced by the proposition that

the planets moved in the lines that *would be* followed by bodies carried by wheels within wheels.

This was a mere way of representing the sum of the observed facts; and Kepler's was another (and better) way of representing the same observations.

It's true that for these merely descriptive operations, as well as for the erroneous inductive one, a mental conception

was required. The conception of an ellipse had to present itself to Kepler's mind before he could identify the planetary orbits with it. Whewell sees the conception as something added to the facts. He implies that Kepler put something into the facts by his way of conceiving them. But Kepler did no such thing. The ellipse was in the facts before Kepler recognised it; just as the island was an island before it had been sailed around. Kepler didn't *put* what he had conceived into the facts, but *saw* it in them. A conception implies, and corresponds to, something conceived; and though the conception itself is not in the facts but in our mind, if it is to convey any knowledge about •the facts it must be a conception of something that really is in •them, some property that they actually have and that they would show to our senses if our senses were able to take cognisance of it. Suppose that Mars left behind it a visible track, and that an observer was in a fixed position that let him see the whole plane of the orbit at once; he would see it to be an ellipse. . . . I don't think anyone would deny that in this case identifying the planet's path with an ellipse is *describing* it; and I can't see why it makes any difference that in fact that path of Mars is not directly an object of sense, given that every point in it is as exactly ascertained as if it were so.

. . . .I don't think that the role of conceptions in the operation of studying facts has ever been overlooked or undervalued. No-one ever disputed that in order to reason about something we must have a conception of it; or that when we include a multitude of things under a general expression the expression implies a conception of something common to those things. But it by no means follows that the conception must be pre-existent, or that the mind constructs it out of its own materials. If the facts are rightly classed under the conception, that's because the facts themselves contain something x of which the conception is a copy;

and if we can't directly perceive x, that's because of the limited power of our organs and not because x isn't there. The conception itself is often obtained by abstraction from the very facts which. . . .it is afterward called in to connect. Whewell admits this himself when he observes. . . .what a great service it would be to the science of physiology if a philosopher were to 'establish a precise, tenable, and consistent conception of life'. Such a conception has to be abstracted from the phenomena of life itself—from the very facts that it is required to connect. In other cases, instead of •collecting the conception from the phenomena we are trying to colligate, we •select it from among the conceptions that have already been collected by abstraction from other facts. That's what happened with Kepler's laws. The facts were out of the reach of any observation that could enable the senses to identify the path of the planet directly, so the required conception couldn't be collected by abstraction from the observations themselves; the mind had to supply hypothetically, from among the conceptions it had obtained from other portions of its experience, some one conception that would correctly represent the series of the observed facts. It had to form a supposition regarding the general course of the phenomenon, and ask itself 'If this is the general description, what will the details be?', and then compare these with the details actually observed. If they agreed, the hypothesis would serve as a description of the phenomenon; if not, it had to be abandoned and another tried. It's this sort of case that gives rise to the ·false· doctrine that the mind in forming the descriptions adds something of its own that it doesn't find in the facts.

Mars *does* follow an ellipse; that is a fact, surely, and one that we could *see* if we had adequate visual organs and a suitable position. Lacking these advantages but possessing the conception of an ellipse. . . ., Kepler looked to see whether

the observed places of the planet were consistent with such a path. He found they were so; which led him to assert as a fact that the planet moves in an ellipse. But this fact, which Kepler didn't add to the motions of the planet but found in them, was the very fact whose separate parts had been separately observed; it was the sum of the different observations.

Having stated this basic difference between my opinion and Whewell's, I must add that his account of how a conception is selected that is suitable to express the facts seems to me absolutely right. The process is tentative: it consists of a series of guesses of which many are rejected until eventually one is found that is fit to be chosen. We know from Kepler himself that before hitting upon the 'conception' of an ellipse he tried nineteen other imaginary paths which he had to reject because they didn't fit the observations. Whewell is right in saying that the successful hypothesis, though it is a guess, is usually not a lucky guess but a skillful one. The guesses that give mental unity and wholeness to a chaos of scattered particulars seldom occur except in minds abounding in knowledge and disciplined in intellectual combinations.

The tentative method is indispensable to the •colligation of facts for purposes of description. How far can it be applied to induction itself? and what functions does it have in that department? I'll discuss this in chapter 14 of this Book. Right now my main task is to distinguish colligation from induction properly so-called; and to make that distinction clearer I'll discuss a curious and interesting remark •of Whewell's• which is as strikingly true of colligation as it is false of induction, or so it seems to me.

[The 'remark' in question is something Whewell says about the successive accounts of the movements of (for example) the planet Mars—that it moves

- in a circle with the earth as centre;
- in a circle with the earth inside the circle but not at its centre;
- in epicycles, i.e. little circles whose centres move in a circle around the earth;
- in an ellipse.

Whewell says that each of these was correct as far as it went. In Mill's words: 'They all served the purpose of colligation; they all enabled the mind to represent to itself easily and all at once the whole body of facts •about Mars• that had been established up to then.' Mill also quotes Comte as saying the same thing, and goes on to express his own agreement and then to draw a line:] Whewell's remark, therefore, is philosophically correct. Successive expressions for the colligation of observed facts—i.e. successive **descriptions** of a whole phenomenon that has been observed only in parts—can all be correct as far as they go, although they conflict with one another. But it would surely be absurd to assert this of conflicting **inductions**.

A scientific study of facts may be undertaken purposes of **(a)** the simple description of the facts, **(b)** the explanation of the facts, or **(c)** the prediction of similar facts. Of these, **(a)** is not, while **(b)** and **(c)** are, properly called 'induction'. Whewell's remark is true of **(a)**: . . . The elliptical theory, as a mere description, was simpler and more easily usable than its predecessors, but it wouldn't really be more true than they were. So different **(a)** descriptions can all be true, but surely not different **(b)** explanations, such as these explanations of the movements of the planets:

- (i)** They are moved by a 'virtue' [see Glossary] inherent in their celestial nature;
- (ii)** They are moved by impact (which led to the hypothesis of vortices [see Glossary] as the only pushing force capable of whirling bodies in circles);

(iii) They are moved by the composition of a centripetal force with an original projectile force (Newton).

These are explanations collected by real induction from supposed parallel cases; and each had its time of being accepted. Can it be said of *these*. . . that they are all true as far as they go? Isn't it clear that at most one of them is true—that only one can be true in any degree, and the other two must be altogether false? Now consider two **(c)** predictions:

- Eclipses will occur when one planet or satellite casts its shadow on another;
- Eclipses will occur when some great calamity is impending over mankind.

Do these two doctrines differ only in expressing real facts with different degrees of accuracy? Assuredly one of them is true and the other absolutely false.

·START OF A LONG FOOTNOTE·

[This footnote reports and responds to two replies that Whewell made to the content of this section up to here. According to Whewell, the three explanations of planetary motion that Mill cites could all be true. His defence of this involves construing each of **(i)** and **(ii)** as being abstract and formal in such a way that **(iii)** can be seen as a factually contentful cashing in of it:]

'If **(i)** had been maintained in such a way as to agree with the facts, the inherent virtue would have had its laws determined; and then it would have been found that the virtue related to the central body; and so, the "inherent virtue" would have coincided in its effect with **(iii)** the Newtonian force and the two explanations would agree—except in regard to the word "inherent". This word indicates a part of theory **(i)** that was found to be untenable, so it was of course rejected in the transition to later and more exact theories.'

[Mill replies:] Whewell says that the theory of an inherent virtue agrees with Newton's when the word 'inherent' is left out, which of course it would be (he says) if 'found to be untenable'. But leave that out and where's the theory? The word 'inherent' is the theory. When it is omitted, all that remains is the statement that the heavenly bodies move 'by a virtue'—i.e. by a power of some sort—or by virtue of their celestial nature, which directly contradicts **(iii)** the doctrine that terrestrial bodies fall by the same law.

[Whewell again:] 'The doctrine **(ii)** that the heavenly bodies were moved by vortices was successfully modified so that it came to coincide in its results with the doctrine **(iii)** of an inverse-quadratic centripetal force. . . . When this point was reached, the vortex was merely a machinery. . . . for producing such a centripetal force, and therefore didn't contradict the doctrine of a centripetal force. . . .'

[Mill replies:] If the doctrine **(ii)** of vortices had meant not that vortices existed but only that the planets moved as though they were whirled by vortices; if the hypothesis had been merely a way of representing the facts and not an attempt to account for them; if (in short) it had been only a description; no doubt it *would* have been reconcilable with **(iii)** the Newtonian theory. But the vortices were *not* a mere aid to conceiving the motions of the planets, but a supposed physical agent actively pushing them. . . . and according to Newton's theory this was not true. Whewell seems to think of Newton's theory as stating only the •directions of the forces and not •their nature, and therefore as not conflicting with any hypothesis about how they are produced. Well, **(iii)** the Newtonian theory regarded as a mere •description of the planetary motions doesn't conflict with **(ii)**; but the Newtonian theory as an •explanation of them *does*. The explanation consists in

ascribing those motions to a general law that holds between all particles of matter, and identifying this with the law by which bodies fall to the ground.

If the planets are kept in their orbits by a force that draws the particles composing them toward every other particle of matter in the solar system, then they are *not* kept in those orbits by the impulsive force of certain streams of matter that whirl them around. One explanation absolutely excludes the other. . . . Denying this is like saying that there's no contradiction between 'That man died because somebody killed him' and 'That man died a natural death'.

If Whewell is not yet satisfied, any other subject will serve equally well to test his doctrine. He will hardly say that there's no contradiction between the members of each of these pairs:

- Light is a stream of particles
- Light is a series of waves.

- Higher organic forms arose by development from the lower.
- The different organic forms came from separate and successive acts of creation.

- Volcanoes are fed from a central fire.
- Volcanoes come from chemical action at a comparatively small depth below the earth's surface.

If different explanations of the same fact can't both be true, still less can different predictions. Whewell quarrels (never mind why) with my choice of example on this point, and thinks that a theory is sufficiently answered by an objection to an illustration of it. Well, examples not liable to his objection are easily found, if the production of many

examples is really needed to support the thesis that conflicting predictions can't both be true! [Mill then gives some examples.]

Whewell sees no distinction between •holding contradictory opinions on a question of fact, and •merely employing different analogies to help the conception of the same fact. Different inductions belongs to the former class, different descriptions to the latter.

•END OF LONG FOOTNOTE•

. . . .But induction is *connected* with colligation in two ways. •Induction is always colligation. The assertion that the planets move in ellipses was only a colligation; whereas the assertion that the planets are drawn (or tend) toward the sun was the statement of a new fact, inferred by induction. But it *also* served as a colligation: it brought the facts which Kepler had connected by his conception of an ellipse under the additional conception of bodies acted upon by a central force. . . . •The descriptions reached by colligation are a necessary preparation for induction. . . . Without the previous colligation of detached observations by means of one general conception we could never have obtained any basis for an induction except in the case of very limited phenomena. We couldn't affirm any predicates of a subject that we could observe only piecemeal, let alone extending those predicates by induction to other similar subjects. . . .

§5. . . .Whewell has replied to all this, re-stating his case but not strengthening it, as far as I can see. Since my arguments have not had the good fortune to make any impression upon him, I will add a few remarks to show more clearly what we are differing about and in some measure to account for the difference.

Nearly all writers of authority define *induction* as drawing inferences from known cases to unknown:

- applying to a class a predicate that has been found true of some members of the class;
- inferring from the fact that things have a certain property that other things resembling them have the same property;
- inferring from the fact that a thing had a property at a certain time that it does and will have that property at other times.

Kepler's operation clearly wasn't an induction in this sense of the term! The statement that Mars moves in an elliptical orbit wasn't any kind of extension from facts to further facts. . . . Kepler didn't extend an observed truth to cases other than those in which it had been observed; he didn't widen the **subject** of the proposition expressing the observed facts. The alteration he made was in the **predicate**. Instead of saying 'The successive places of Mars are so-and-so' he summed them up in the statement 'The successive places of Mars are points in an ellipse'. Whewell says that this statement wasn't the sum of the observations *merely*, and I agree; it was the sum of the observations *seen under a new point of view*. But it wasn't the sum of more than the observations, as a real induction is. It covered only cases that had been actually observed or could have been inferred from the observations before the new point of view presented itself. There was not the transition from known cases to unknown ones that constitutes 'induction' in the original and acknowledged meaning of the word.

Old definitions can't prevail against new knowledge: scientific language ought to adapt itself to the true relations that hold between the things it is used to designate. If the Keplerian operation really is identical—considered as a logical process—with what happens in acknowledged induction, the definition of 'induction' should be widened so as to take it in. This is where I take issue with Whewell. He does think

that the operations are identical. He holds that the only logical process in any induction is one that also occurs in Kepler's case, namely *guessing until a guess is found that squares with the facts*. That leads him to reject all canons of induction, because it's not by means of them that we guess. Whewell's theory of the logic of science would be very perfect if it didn't pass over altogether the question of *proof*. But I think there *is* such a thing as proof, and inductions relate to it quite differently from how descriptions do. Induction is proof; it is inferring something unobserved from something observed; it requires, therefore, an appropriate test of proof; and to provide that test is the special purpose of inductive logic. [That sentence is verbatim from Mill.] When on the other hand we merely collate known observations and (as Whewell puts it) connect them by means of a new conception, if the conception does serve to connect the observations we have all we want. The proposition containing it claims only to have truth that it may share with many other ways of representing the same facts. So all it requires is to be consistent with the facts; it can't be proved and doesn't need to be. It may serve to prove *other* things: it places facts into a mental connection with other facts that hadn't previously been seen to resemble them, and thereby assimilates the case to another class of phenomena concerning which real inductions have already been made. Thus Kepler's so-called 'law' brought the orbit of Mars into the class *ellipse*, thereby proving all the properties of an ellipse to be true of the orbit; but in this proof Kepler's 'law' supplied the minor premise and not (as with real inductions) the major.

[Explaining that last clause: Mill is thinking in terms of syllogisms (dealt with in II.2) of the form:

- (1) All elliptical orbits have the property F.
- (2) Mars's orbit is an elliptical orbit.

Therefore

- (3) Mars's orbit has the property F.

In this syllogisms **(2)**—Kepler's 'law'—is the minor premise (meaning that it contains the subject of the conclusion). If that syllogism expressed a real induction, Mill thinks, it would be the inductive leap from **(1)** to **(3)**.]

Whewell calls something an induction if, and only if, it introduces a new mental conception; but this is running together two very different things, •invention and •proof. Introducing a new conception belongs to invention; this may be required in any operation, but it isn't the essence of any. . . . Most inductions require no conception except what was present in each of the particular instances on which the induction is based. That *all men are mortal* is surely an inductive conclusion, but it doesn't introduce any new conception; if you know that *some man has died* you have all the conceptions involved in the inductive generalisation. Whewell, however, considers the process of invention—i.e. forming a new conception consistent with the facts—to be not merely a necessary •part of all induction but the •whole of it.

The mental operation that extracts from a number of detached observations certain general characters in which the observed phenomena resemble one another, or resemble other known facts, is what Bacon, Locke, and most subsequent metaphysicians have understood by 'abstraction'. I think it is strictly logically correct to call

a general expression obtained by abstraction, connecting known facts by means of shared characteristics but without inferring further facts from them

a 'description'; and I don't know how else anything could be described! But I don't rely on the use of that particular word; I'm quite content to use Whewell's term 'colligation', or the more general 'mode of representing or of expressing phenomena'; provided it is clearly seen that the process is not induction but something radically different.

[Mill says that he will return to these matters, and remove difficulties that the reader may have had with the present chapter, in Book IV.]

Chapter 3. The ground of induction

§1. Thus, induction *properly so-called* can be briefly defined as **generalisation from experience**. It consists in inferring from some individual instances in which a phenomenon is observed to occur that it occurs in all instances that resemble the former in what are regarded as the material [see Glossary] circumstances.

How are we to tell material [see Glossary] circumstances from ones that are immaterial? Why are some circumstances material and others not so? I'll come to those questions in due course. I must first point out that there's a principle implied in the very statement of what induction is—an

assumption about the course of nature and the order of the universe—namely that there are such things in nature as parallel cases; that what happens once will happen again when the circumstances are sufficiently alike, and not only *again* but *as often as the same circumstances recur*. This assumption is involved in every case of induction; and looking at the actual course of nature we see that the assumption is justified. The universe, as far as we know it, is constituted in such a way that whatever is true in any one case is true in all cases of a certain description; the only difficulty is to find what description.

This universal fact, which is our warrant for all inferences from experience, has been stated variously by different philosophers: 'The course of nature is uniform', 'The universe is governed by general laws', and the like. Metaphysicians of the school of Reid and Stewart have popularised one of the most inadequate of these formulations. The human mind's disposition to generalise from experience—a propensity these philosophers regard as an instinct of our nature—they usually describe as something like 'our intuitive conviction that the future will resemble the past'. Now, Bailey rightly said that *time* doesn't come into it. We believe that fire will burn tomorrow because it burned today and yesterday; but we believe on precisely the same grounds that it burned before we were born and that it burns today in China. We don't infer from the •past to the •future as such; we infer from the •known to the •unknown, from observed facts to unobserved facts. . . . This second category includes the whole region of the future; but it also includes nearly the whole of the present and of the past.

Express it how you will, the proposition that *the course of nature is uniform* is the basic principle or general axiom of induction. But it doesn't *explain* the inductive process. On the contrary, it is itself an instance of induction, and induction that is by no means obvious. Far from being the first induction we make, it's one of the last—or anyway one of the last that we make in a philosophically accurate form. In fact it has hardly entered into the minds of any but philosophers, and we'll see that even *they* have haven't always had a sound conception of its extent and limits. This great generalisation is, in fact, itself based on prior generalisations. The obscurer laws of nature were discovered by means of it, but the more obvious ones must have been understood and assented to as general truths before it was ever heard of. We would never have thought of saying that all

phenomena conform to general laws if we hadn't first arrived at some knowledge of many of the laws themselves—which had to be done by induction. In what sense, then, can a principle that is so far from being our earliest induction be our warrant for all the others? In the only sense in which the general propositions that we place at the head of our syllogisms ever really contribute to their validity. (I explained what this is in II.3.) Whately remarks that every induction is a syllogism with the major premise suppressed; a better formulation would be that every induction can be put into syllogistic form by supplying a major premise. If this is actually done, the principle of the uniformity of the course of nature will appear as the ultimate major premise of all inductions; so it will relate to all inductions in the way the major premise of *every* syllogism relates to its conclusion. And what relation is that? It doesn't contribute to proving the conclusion; but it's a necessary condition of its being proved, because no conclusion is proved unless there's a true major premise.

You may want me to explain this claim that the uniformity of the course of nature is the ultimate major premise in all inductions. It certainly isn't the *immediate* major premise in every inductive argument. . . . The induction, 'John, Peter, etc. are mortal, therefore all mankind are mortal' can be put into syllogistic form by prefixing the major premise that what is true of John, Peter, etc. is true of all mankind. How did we get this major premise? It isn't self-evident. . . ., so we must have arrived at it by induction. [Mill says 'by induction or ratiocination', but he drops ratiocination without comment.] If by induction, this process like all other inductive arguments can be put into syllogistic form; and we need to construct this previous syllogism. There is in the long run only one possible construction: the real proof that what is true of John, Peter, etc. is true of all mankind can only be that a

different supposition would be inconsistent with the known uniformity in the course of nature. Whether there actually *would* be this inconsistency may be a matter of long and delicate inquiry; but if there wouldn't, we have no sufficient ground for the major premise of the inductive syllogism. It seems, then, that if we put the whole course of any inductive argument into a series of syllogisms, we'll eventually arrive at an ultimate syllogism whose major premise is the principle or axiom of the uniformity of the course of nature.¹

Why should this axiom be accepted as true? It wasn't to be expected that thinkers would unanimously give one answer to this question, any more than with other axioms. I have already said that I think it is itself a generalisation from experience. Others think we're compelled by the constitution of our thinking faculty to assume it as true in advance of any verification by experience. Having in II.5–6 fought at such length against a similar doctrine regarding the axioms of mathematics, using arguments that largely apply also to the present case, I'll postpone going into more detail about it until chapter 21. At present it matters more to understand the import of the axiom itself. The proposition 'The course of nature is uniform' has the brevity suitable for ordinary talk rather than the precision required in philosophical language; if it's to be accepted as true, its terms need to be explained, and given a stricter signification than they have in ordinary speech.

§2. Everyone knows that he doesn't always expect uniformity in the course of events; he doesn't always believe that the unknown resembles the known, that the future will resemble the past. Nobody believes that the pattern of rain and sunshine will always be the same as it is this year. Nobody expects to have the same dreams every night; indeed when the course of nature *is* constant in these things, everyone mentions it as something extraordinary. To look for constancy where it's not to be expected—e.g. to expect that a date that once brought good fortune will always be a fortunate date—is rightly regarded as superstition.

The course of nature is not only uniform, it's infinitely various. Some phenomena always recur in the combinations they had when we first met with them; others seem altogether capricious; and some get us used to experiencing them in one particular combination and then unexpectedly break that pattern. The experience of an inhabitant of Central Africa fifty years ago supported 'All human beings are black' as well as it supported anything. To Europeans until recently 'All swans are white' seemed to be an equally straightforward example of uniformity in the course of nature. Each group had to wait fifty centuries for the experience that showed them to be wrong. During that time, mankind believed in a uniformity in the course of nature where no such uniformity really existed.

¹ It needn't be uniformity that pervades all of nature. It's enough that it pervades the particular class of phenomena to which the induction relates. An induction about planetary motion wouldn't be spoiled if we thought that wind and weather are the sport of chance, provided we are assuming that astronomical phenomena are governed by general laws. Otherwise the early experience of mankind would have rested on a very weak foundation, because in the infancy of science it couldn't be known that *all* phenomena are regular in their course.

The major premise doesn't have to be known in advance; it's enough if we *can* now know it. . . . The conclusion 'The Duke of Wellington is mortal', inferred from the instances A, B, and C, implies either that we have concluded all men to be mortal or are now entitled to do so from the same evidence. A vast amount of confusion and bad logic regarding the grounds of induction would be dispelled by keeping these simple considerations in view.

According to the ancients' notion of induction, the 'Black person' and 'White swan' conclusions involved inferences that were as legitimate as any inductions whatever. Because each conclusion was false, the ground of inference must have been insufficient, but still there was as much ground for it as this conception of induction admitted of. The induction of the ancients has been well described by Bacon under the label 'induction by simple enumeration. . . .'. It consists in ascribing the character of general truths to all propositions that are true in every instance that we happen to know of. This is the kind of induction that is natural to a mind that isn't used to scientific methods. The tendency (some say 'instinct', others say 'association') to infer the future from the past, the known from the unknown, is simply a habit of expecting that what has been found true once or several times, and never yet found false, will be found true again. It makes no difference whether the instances are few or many, conclusive or inconclusive; these considerations occur only on reflection. The unprompted unreflective tendency of the mind is to generalise its experience, provided it all points in one direction and no conflicting experience comes •unsought. The notion of •seeking it, •experimenting for it, •interrogating nature (Bacon's phrase) is a much later development. When uncultivated intellects observe nature, they are passive; they accept the facts that present themselves, without actively searching for more. Only a superior mind asks itself 'What facts do I need to come to a safe conclusion?' and then looks out for these.

But though we always tend to generalise from unvarying experience, we aren't always justified in doing so. We aren't entitled to conclude that something is universally true because we have never known a counter-instance to it unless we have reason to believe that if there were any counter-instances in nature we would have known of them. When we

do have this assurance—this reason to believe. . . etc.—this may enable induction by simple enumeration to amount practically to proof. But cases where that is how things stand are very •remarkable—I'll discuss them in chapters 21–22 below. No such assurance can be had on any of the •ordinary subjects of scientific inquiry. Popular [see Glossary] notions are usually based on induction by simple enumeration; in science that doesn't take us far. We're forced to begin with it; we often have to rely on it provisionally, in the absence of anything better; but for the accurate study of nature we need a surer and a more potent instrument.

Bacon's usual title of 'Founder of the Inductive Philosophy' is one he deserved primarily for pointing out the insufficiency of this rough and loose conception of induction. The value of his own contributions to a more philosophical theory of the subject has certainly been exaggerated. Although his writings contain. . . more or less fully developed statements of several important principles of the inductive method, physical science has now far outgrown the Baconian conception of induction. Moral and political inquiry still haven't caught up with that conception. The approved modes of reasoning on these subjects are still. . . the very induction by simple enumeration that he condemns; and the 'experience' that we hear so confidently appealed to by all sects, parties, and interests is still, in Bacon's own emphatic words, *mera palpatio*. [Latin for 'mere feeling'. The kind of feeling he is referring to can be gathered from this: 'Those who steer by simple experience are like men in the dark, patting the walls as they go along hoping to find their way, when they'd have done much better to wait for daylight, or light a candle, and *then* set off.']

§3. For a better understanding of the problem the logician must solve if he's to establish a scientific theory of induction, let us compare some incorrect inductions with others that

are acknowledged to be legitimate. . . . That all swans are white can't have been a good induction, because the conclusion turned out to be false. But. . . from the earliest records, the testimony of the inhabitants of the known world was unanimous on the point. So the uniform experience of the inhabitants of the known world, agreeing in a common result with no known counter-examples, isn't always sufficient to establish a general conclusion.

Now take an instance that is apparently not very dissimilar to this. Mankind were wrong in concluding that all swans are white; are we also wrong when we conclude that all men's heads grow above their shoulders, and never below, in spite of the conflicting testimony of the naturalist Pliny? As there were black swans, though civilised people had existed for 3000 years on the earth without meeting with them, may there not also be 'men whose heads do grow beneath their shoulders', despite a rather less perfect unanimity of negative testimony from observers? [The quoted phrase is from *Othello*.] Most people would answer No: it's more credible that a bird should vary in its colour than that men should vary in the relative position of their principal organs. Of course they would be right—but *why* are they right? We can't answer that without going into the true theory of induction more deeply than is usually done.

. . . .When a chemist announces the existence and properties of a newly-discovered substance, if we trust his accuracy we feel assured that his conclusions will hold universally, although the induction is based on a single instance. We don't withhold our assent until the experiment is repeated; or if we do, it's because we aren't sure that the one experiment was properly done; we are sure that if it was properly done it was conclusive. So here we have a general law of nature, inferred without hesitation from a single instance; a universal proposition from a singular one. Now contrast that with another case. All the instances that have been observed since the beginning of the world in support of 'All crows are black' wouldn't outweigh the testimony of one apparently reliable witness who said that in a region of the earth not fully explored he had caught and examined a crow and found it to be gray.

Why is it that in some cases a single instance is sufficient for a complete induction, while in others myriads of concurring instances without a single known or presumed exception go such a very little way toward establishing a universal proposition? Whoever can answer this question knows more of the philosophy of logic than the wisest of the ancients, and has solved the problem of induction.

Chapter 4. Laws of nature

§1. . . . The uniformity of the course of nature is really *uniformities*. . . . The course of nature in general is constant because the course of each of the various phenomena that compose it is so. A certain fact invariably occurs whenever certain circumstances are present, and doesn't occur when

they are absent; the same is true of another fact; and so on. From these separate threads of connection between parts of the great whole that we call 'nature' a general tissue of connection unavoidably weaves itself, by which the whole is held together. If

A is always accompanied by D,
 B by E and
 C by F,

it follows that

A B is accompanied by D E,
 A C by D F,
 B C by E F, and finally
 A B C by D E F.

That is how the general character of regularity is produced—a regularity which, along with and in the midst of infinite diversity, pervades all nature.

So the uniformity of the course of nature is itself a complex fact, compounded of all the separate uniformities that exist in respect of single phenomena. The ordinary name of these various uniformities, when they are established by what we regard as a sufficient induction, is 'laws of nature'. In scientific usage we use 'law of nature' in a more restricted sense, to refer to the uniformities when reduced to their simplest expression. In the above illustration, already seven uniformities. . . would be called 'laws of nature' in the everyday loose sense of that phrase; but only the first three of the seven are properly distinct and independent; and when they are presupposed, the others automatically follow. So the first three are 'laws of nature' in the narrower sense; the other four are not, because they are mere cases of the first three, virtually included in them, and are therefore said to result from them. Anyone who affirms those three has already affirmed all the rest.

Here are three uniformities, or call them 'laws of nature':

- (1) Air has weight,
- (2) Pressure on a fluid spreads equally in all directions.
- (3) Pressure in one direction, not opposed by equal pressure in the opposite direction, produces motion that doesn't cease until equilibrium is restored.

These three uniformities should enable us to predict another uniformity, namely the rise of the mercury in the Torricellian tube [= a barometer]. This isn't a 'law of nature' in the narrower sense of the phrase; it's a *result of* laws of nature. It's a case of each of the three laws; and it couldn't be different without infringing at least one of them. If the mercury were not held up in the barometer at a height such that the column of mercury had the same weight as a column of the atmosphere of the same diameter, this would be a case of

- (1) the air not pressing on the surface of the mercury with the force that is called its 'weight', or of
- (2) the downward pressure on the mercury not being passed on equally in an upward direction, or of
- (3) a body pressed in one direction and either not moving in that direction or stopping before reaching equilibrium.

So if we knew the three simple laws but had never tried the Torricellian experiment, we could have *deduced* its result from those laws. The known weight of the air combined with the position of the apparatus would bring the mercury within the range of the first of the three inductions; the first induction would bring it within the second, and the second within the third—all in the way I described in my account of ratiocination. We would be coming to know the more complex uniformity independently of specific experience, through our knowledge of the simpler ones from which it results; though in due course we'll see reasons why *verification* by specific experience would still be desirable and might even be indispensable.

[Mill makes remarks about broad and narrow senses of 'law of nature', suggesting that the narrow sense favoured by scientists is explained by a 'tacit reference to the original sense of "law", namely "the expression of the will of a superior".' He then offers two ways of restating the narrow-sense

version of the question 'What are the laws of nature?:']

- What are the fewest and simplest assumptions, which being granted the whole existing order of nature would result?
- What are the fewest general propositions from which all the uniformities which exist in the universe might be deductively inferred?

Every great advance that marks an epoch in the progress of science has been a step towards solving this problem. Even a simple colligation of inductions already made, with no fresh extension of the inductive inference, is already an advance in that direction. When Kepler expressed the regularity in the observed motions of the planets by the three general propositions called his 'laws', he was showing three simple suppositions that would suffice to construct the whole scheme of **planetary motion** so far as it was then known. A still greater step was made when these laws... were discovered to be cases of the three laws of **motion** generally. . . . After this great discovery, Kepler's three propositions, though still called 'laws', wouldn't be called 'laws of nature' by anyone accustomed to using language with precision; that phrase would be reserved for the simpler and more general laws into which Newton is said to have resolved [see Glossary] them.

Every well-grounded inductive generalisation is either •a law of nature or •a result of laws of nature, something that could be predicted from them. And the problem of inductive logic can be summed up in two questions; how to ascertain the laws of nature; and how then to follow them into their results. But this isn't a real analysis of the problem; it's a mere verbal transformation of it. . . . Still, it is worth something to have reached the insight that the study of nature is the study of laws, not *a* law; of uniformities (plural); that the different natural phenomena have their separate

rules or ways of occurring, which—though intermixed and entangled with one another—can to some extent be studied separately; that the regularity in nature is a web composed of distinct threads, and only to be understood by tracing each thread separately, for which purpose it is often necessary to unravel some portion of the web and exhibit the fibres apart. The rules of experimental inquiry are the contrivances for unraveling the web.

§2. In trying in this way to ascertain the general order of nature by ascertaining the particular order of each one of the phenomena of nature, a scientific proceeding can't be more than an improved form of what the human understanding primitively did when not directed by science. When mankind first got the idea of studying phenomena by a method stricter and surer than the one they spontaneously started out with, they didn't start with the supposition that nothing had yet been ascertained (Descartes's well-meant but impracticable advice). Many of the uniformities in phenomena are so constant and so open that they force themselves on involuntary recognition. Some facts are so perpetually and familiarly accompanied by certain others that mankind learned—as children learn—to expect one where they found the other, long before they knew how to put this expectation into a proposition about there being a connection between those phenomena. No science was needed to teach that food nourishes, that water drowns or quenches thirst, that the sun gives light and heat, that bodies fall to the ground. The first scientific inquirers assumed these and their like, and set out from them to discover others that were unknown; and they weren't wrong to do this, though they later came to see that the initial spontaneous generalisations needed to be revised when the progress of knowledge showed limits to them, or showed their truth to depend on some detail

not originally attended to. You'll see later on that there's no logical fallacy in this procedure. Indeed, we can see that no alternative is workable, because it's impossible to develop any scientific method of induction, or test of the correctness of inductions, except on the hypothesis that some inductions deserving of reliance have already been made.

For an example, look back at one of our former illustrations. With exactly the same amount of evidence (negative and positive) in each case, why did we accept the assertion that **(i)** there are black swans yet refuse to believe any testimony saying that **(ii)** some men wear their heads underneath their shoulders? Well, **(i)** was more credible than **(ii)**; but why was it more credible? . . . Apparently because there is less constancy in the colours of animals than in the general structure of their anatomy. But how do we know this? From experience, of course. It appears, then, that we need experience to tell us how much we should rely on experience, and in what cases or sorts of cases. We have to consult experience to learn from it when and where arguments from it will be valid. We have no rock-bottom test to which we subject experience in general; we make experience its own test. Experience testifies that some of the uniformities it exhibits or seems to exhibit are more to be relied on than others. We have experienced uniformity U; how confident should we be that it holds in cases not yet observed? That depends on the extent to which U belongs to a class of uniformities that have hitherto been found to be uniform.

This way of correcting one generalisation by means of another—a narrower generalisation by a wider one that common sense suggests and adopts in practice—is the real type [see Glossary] of scientific induction. Skillfully contrived rules can give accuracy and precision to this process, and adapt it to all varieties of cases, but they can't make any essential difference to its principle.

To apply a test of this sort we must already have some general knowledge of the typical character of the uniformities existing throughout nature. So the indispensable foundation of a scientific formula of induction has to be a survey of the inductions mankind has been unscientifically led to, with the special purpose of ascertaining what *kinds* of uniformities have been found to be perfectly invariable, pervading all nature, and what kinds have been found to vary with difference of time, place, or other changeable details.

§3. The need for such a survey is confirmed by the fact that stronger inductions are the touchstone to which we always try to bring the weaker. If we find a way to deduce a weaker induction from stronger ones, it instantly acquires all their strength; and it even adds to that strength, because the experience on which the weaker induction previously rested becomes additional evidence for the truth of the stronger ones. Suppose we infer from historical evidence that the uncontrolled power of a monarch, of an aristocracy, or of the majority, will often be abused; we can rely on this generalisation with much greater confidence when it is shown to follow from facts that are even better established—the low degree of elevation of character ever yet attained by the average of mankind, and the poor success-rate of most known procedures for making reason and conscience predominate over the selfish propensities. And obviously these more general facts get more evidence from what history tells us about the effects of despotism. The strong induction becomes still stronger when a weaker one has been bound up with it.

On the other hand, if an induction conflicts with stronger inductions or with conclusions that follow from them, then the weaker one must give way—unless reconsideration shows that some of the stronger inductions have been expressed with greater universality than their evidence justifies. The

age-old opinion that a comet. . . was the precursor of calamities. . . ; the belief in the truthfulness of the oracles of Delphi or Dodona; the reliance on astrology or on the weather-prophecies in almanacs; these were doubtless inductions supposed to be based on experience; and it seems that faith in such delusions can hold out against many failures, as long as it's nourished by a reasonable number of casual coincidences between the prediction and the event.¹ What has really put an end to these insufficient inductions is their inconsistency with the stronger inductions subsequently obtained through scientific inquiry, concerning the causes that terrestrial events really depend on. In places where those scientific truths haven't yet penetrated, the same or similar delusions still prevail.

Here is a two-part general principle about *any* two inductions, whether strong or weak:

- (1) If they can be connected by ratiocination, they tend to confirm one another.
- (2) If they lead deductively to consequences that are incompatible, they become tests of each other, showing that one or other must be given up or at least limited in some way.

In case (1) the induction that becomes a conclusion from ratiocination becomes at least as certain as the weakest of those from which it is deduced; while in general all are more or less increased in certainty. Thus the mercury-in-tube experiment, though it's a mere case of three more general laws, doesn't just strengthen greatly the evidence on which those laws rested but raises the level of one of them (the weight of the atmosphere) from 'doubtful' to 'completely established'.

Thus, if among the uniformities that have been found to exist in nature there are some that may be considered quite certain and quite universal (so far as any human purpose requires certainty), then we may be able to use these to raise multitudes of other inductions to the same point on the certainty scale. For if we can show that either inductive inference I_2 is true or the certain and universal induction I_1 has an exception, then I_2 will attain the same certainty and security as I_1 It will be proved to be a law; and if it's not a result of other and simpler laws, it will be a law of nature.

There *are* such certain and universal inductions; and it's because there are that a logic of induction is possible.

¹ Whewell won't allow these and their like to count as 'inductions', because such superstitious fancies 'were not collected from the facts by seeking a law of their occurrence, but were suggested by an imagining of the anger of superior powers. . . .' But the question is not 'How were these notions first suggested?' but 'What evidence is. . . supposed to support them?' If the believers in these erroneous opinions had been challenged to defend them, they would have referred to experience: to the comet that preceded the death of Julius Caesar, or to oracles and other prophecies known to have been fulfilled. Analogous superstitions exist even today, and their hold on the believers' minds depends on the supposed evidence of experience—mostly consisting of casual coincidences [see Glossary]. I admit that the influence of such coincidences wouldn't be what it is if it weren't strengthened by an antecedent presumption; but this is not a special feature of superstitions; preconceived notions of probability help to explain many other cases of belief on insufficient evidence. The *a priori* prejudice improperly predisposes the believer's mind to interpret his experience in that way, but the believer still sincerely regards his belief as a legitimate conclusion from experience.

—My theme could easily be illustrated by cases where antecedent prejudice has no role. Whately writes: 'For many ages all farmers and gardeners were firmly convinced that the crops would never turn out good unless the seed was sown when the moon was waxing, and that they had learned this from experience.' This was induction, but bad induction; just as an invalid syllogism is reasoning, but bad reasoning.

Chapter 5. The law of universal causation

§1. The phenomena of nature exist in two relations to one another—simultaneity and succession. Every phenomenon is related in a uniform manner to •some phenomena that coexist with it and to •some that have preceded and will follow it.

Of the uniformities that exist among synchronous [see Glossary] phenomena, the most important in every way are the laws of number, closely followed by the laws of space, i.e. of extension and figure. The laws of number are common to synchronous and successive phenomena. That two and two make four is equally true whether the second two follow the first two or accompany them; it's as true of days and years as of feet and inches. In contrast with that, the laws of extension and figure (i.e. the theorems of geometry) are laws only of simultaneous phenomena. The various parts of space and of the objects that are said to 'fill' space *coexist*, and the unvarying laws that are the subject of the science of geometry express *how* they coexist. [Mill is here taking 'x coexists with y' to mean 'x exists at some time when y also exists'.]

To understand and prove these laws—i.e. these uniformities—you don't have to suppose any lapse of time, any variety of facts or events succeeding one another. The propositions of geometry are independent of the succession of events. All things that have extension, i.e. that fill space, are subject to geometrical laws. Having extension they must have figure; so they must •have some figure in particular and •have all the properties that geometry assigns to that figure. An example: If x is a sphere and y a cylinder with the same height and diameter, x's volume will be exactly two-thirds of y's, no matter what stuff x and y are made of. Another example: Each body and each point within a body

must occupy some place or position among other bodies; and the position of two bodies relatively to each other, whatever stuff they are made of, can be unerringly inferred from the position of each of them relatively to any third body.

In the laws of number, then, and in the laws of space, we clear cases of the rigorous universality that we're looking for. Those laws have always been the type [see Glossary] of certainty, the standard of comparison for all lower degrees of evidentness. They are so perfectly invariable that we can't even *conceive* any exception to them (and philosophers have been led—though wrongly—to think that what makes them evident is not experience but the basic constitution of the intellect). So if we could deduce from the laws of space and number any other kind of uniformities, that would be proof positive that those other uniformities had the same rigorous certainty. But we can't do this. From laws of space and number alone nothing can be deduced but laws of space and number.

For us the most valuable truths about phenomena are the ones concerning the order of their succession. Our knowledge of these truths is the basis for every reasonable anticipation of future facts, and for any power we have to influence those facts to our advantage. Even the laws of geometry—which don't involve succession—are chiefly of practical importance to us because they enter into premises from which we can infer the order of the succession of phenomena. The motion of bodies, the action of forces, and the propagation of influences of all sorts take place along certain lines and over definite distances; so the properties of those lines and distances are an important part of the laws to which those phenomena are subject. Similarly, •motions,

•forces or other influences, and •times are numerable quantities; so the properties of number are applicable to them as to everything else. But the laws of number and space can't *unaided* contribute to the discovery of uniformities of succession. They can be made to do that work only when we combine with them premises about uniformities of succession that we already know. Take for example the propositions:

- Bodies acted on by an instantaneous force move with uniform velocity in straight lines.
- Bodies acted on by a continuous force move with accelerated velocity in straight lines.
- Bodies acted on by two forces in different directions move in the diagonal of a parallelogram, whose sides represent the direction and quantity of those forces.

If we combine these truths with certain geometrical propositions (e.g. that a triangle is half a parallelogram of the same base and altitude), we can deduce another important uniformity of succession:

- A body moving around a centre of force marks off areas proportional to the times.

But we must have laws of succession in our premises if we are to reach truths of succession in our conclusions. . . .

The laws of space are only laws of simultaneous phenomena; and the laws of number, though true of successive phenomena, don't relate to their succession; so the rigorous •certainty and •universality of these laws don't carry through to laws of succession. We must try, then, to find some law of succession that has •those attributes, making it a fit basis for processes of discovering and testing all other uniformities of succession. This basic law must resemble the truths of geometry in their most remarkable special feature, namely that they are never *ever* defeated or suspended.

Of the uniformities in the succession of phenomena that

common observation brings to light very •few have even an *apparent* claim to this rigorous indefeasibility; and of those few only •one has been completely sustained in this claim. That one, however, is a law that is universal also in another sense; it is coextensive [see Glossary] with the entire field of successive phenomena, all instances of succession being examples of it. This law is the Law of Causation. The truth that *every fact that has a beginning has a cause* is coextensive with human experience.

You may think that this doesn't amount to much, because it only says 'It's a law that every event depends on some law' or 'It's a law that there's a law for everything'. But we shouldn't conclude that the principle's generality is merely verbal; when we look into it we'll find that far from being vague or meaningless it is a most important and really fundamental truth.

§2. The notion of *cause* is the root of the whole theory of induction; so we must at the outset of our inquiry get it fixed with as much precision as we can manage. There is an old and still-running battle among different schools of metaphysicians concerning the •origin and •analysis of our idea of causation; but—fortunately!—we don't need to settle *that* before starting our search for the true theory of induction. The science of *the investigation of truth by means of evidence* is •independent of many of the controversies that perplex the science of *the ultimate constitution of the human mind*, and •has no need to push the analysis of mental phenomenon to the extreme limit that a metaphysician ought to demand.

Thus, when in the course of this inquiry I speak of the 'cause' of any phenomenon, I don't mean a cause that isn't itself a phenomenon; I am not inquiring into the ultimate or ontological cause of anything. To adopt a distinction familiar

in the writings of Reid and other Scottish metaphysicians, the causes I'm concerned with are not **efficient causes** [see Glossary] but **physical causes**. They are 'causes' purely in the sense in which one physical fact is said to be the 'cause' of another. I'm not called upon to give an opinion about the 'efficient causes' of phenomena, or whether there are any. According to the schools of metaphysics that are most currently most fashionable,

The notion of causation implies a mysterious and most powerful tie of a kind that can't (or anyway doesn't) exist between two physical facts x and y such that x is always followed by y and is popularly caused y's 'cause'. So if we want to find the true cause, the cause that isn't only •followed by the effect but actually •produces it, we have to ascend higher [Mill's phrase; we might prefer to say 'dig deeper'] into the essences and inherent constitution of things.

I have no need to do that for the purposes of the present inquiry, and no such doctrine will be found in the following pages. The only notion of *cause* that the theory of induction needs is one that can be gained from experience. The Law of Causation, the recognition of which is the main pillar of inductive science, is merely the familiar truth that *invariability of succession* is empirically found to obtain between every fact in nature and some other fact that has preceded it—independently of any question about •the ultimate •or absolutely basic• cause of phenomena or about •the nature of 'things in themselves'.

So there's an invariable order of succession between phenomena existing at any instant and the phenomena that exist at the next instant. . . . Certain facts always are—and, we believe, always will be—followed by certain other facts. The invariable antecedent is termed the 'cause', and the invariable consequent the 'effect'. And the law

of causation holds universally because every consequent is connected in this way with some particular antecedent or set of antecedents. Every fact that has begun to exist was preceded by some fact(s) with which it is invariably connected. For every event E there's some combination of objects or events—some combination of circumstances, positive and negative—the occurrence of which is always followed by E. Even when we don't yet know what this combination is for a given E, we never doubt that there is one, and that it never occurs without E as its effect or consequence. If this truth weren't universal, we couldn't express the inductive process in rules. . . .

§3. This invariable sequence seldom if ever holds between a consequent and •a single antecedent. It's usually between a consequent and •the sum of several antecedents, the concurrence [see Glossary] of all of them being needed to produce—i.e. to be certain of being followed by—the consequent. People often single out one of the antecedents as the 'cause' and call the others merely 'conditions'. [Mill elaborates this in more detail than we need. Someone dies because of his eating some oysters. Many will say that his eating the oysters was 'the cause' of his death, but other things were also needed: his general physical constitution, his state of health at this moment, perhaps the room-temperature, etc. These plus the eating of the oysters made up the cause of his death: •the other causes were waiting to have the oyster-meal added to them so that the effect could be produced. •They tend to be left out of 'the cause' because they were relatively long-lasting states of affairs and not short-term *events* like the eating of the meal. Because the total cause was topped up by that one event, people get the impression that the event had 'a more immediate and close connection' with the death than did the other conditions; but it didn't.]

Even when we're aiming at accuracy we don't list all the conditions, but that's because some of them are understood without being expressed, or because our immediate purpose won't be harmed by omitting them. When we say that the cause of a man's death was that •his foot slipped when he was climbing a ladder, we omit •his weight as part of the clause (though it related to his death in the same way as his foot-slip did) because there was no need to mention it in this context. . . . When the decision of a legislative assembly is settled by the casting vote of the chairman, we sometimes say that *he* was the cause of everything that resulted from the enactment. We don't really think that his single vote contributed more to the result than any other affirmative vote; but for our present purpose, namely to insist on his individual responsibility, the part that anyone else had in the transaction is not material.

In all these examples, the fact that was picked out as 'the cause' was the condition that came into existence last. But don't think that in the use of 'the cause' we are strictly guided by this or any other rule. There's no scientific basis for the distinction between the •cause of a phenomenon and its •conditions; you can see this from how capriciously we select the condition of an event that we choose to call its 'cause'. However many conditions there are, almost any of them might count as 'the cause' because our immediate purpose can afford to neglect the others. Take a case where a stone thrown into water sinks to the bottom. What are the conditions of this event? In the first place there must be •a stone, and •water, and •the stone's being thrown into the water; but these suppositions are part of the statement of what the event *is*, and it's bad, a tautology, to include them among the 'conditions' of the event. This class of conditions have never been called 'cause' except by the Aristotelians, who called them 'the material cause'. [Actually, they'd have

said this about the stone and the water, but not about the stone's being thrown into the water.] The next condition is there being •an earth; and accordingly it's often said that the fall of a stone is caused by the earth, i.e. by

- a power of the earth,
- a property of the earth, or
- a force exerted by the earth,

all of which are merely roundabout ways of saying that it is caused by the earth. Or the fall may be said to be caused by

- the earth's attraction.

That's a technical way of saying that *the earth* causes the motion, with an extra special feature, namely that the motion is toward the earth—which is a feature of the effect, not of the cause. Now pass to another condition: it's not enough that the earth should exist; the body must be close enough to the earth for the earth's attraction to outweigh the attraction of any other body. So we can correctly say that the cause of the stone's falling is its being within the sphere of the earth's attraction. A further condition: because the stone is immersed in water, it can't reach the bottom unless its specific gravity exceeds that of the water, i.e. unless it weighs more than an equal volume of water. So it would be regarded as correct to say that the cause of the stone's going to the bottom was its exceeding in specific gravity the fluid in which it was immersed.

Thus we see that *each* condition of the phenomenon may be taken in its turn and spoken of as if it were the entire cause—with equal propriety in everyday speech and equal impropriety in scientific discourse. The particular condition picked out as 'the cause' is usually the one •whose share in the event is superficially the most conspicuous, or •whose status as required for the event we happen to be emphasising at the moment. This second consideration can even lead us to select as 'the cause' one of the negative conditions, as in '

The army was taken by surprise because the sentinel was off his post'. The sentinel's absence didn't create the enemy or put the soldiers to sleep, so how did it cause them to be surprised? All that is really meant is that the event wouldn't have happened if he had been at his duty. His being off his post was not

- a producing cause, but merely
- the absence of a preventing cause.

It was simply equivalent to his non-existence. No consequences can come from nothing, from a mere negation. All effects are connected by the law of causation with some set of *positive* conditions, though *negative* ones are almost always required in addition. In short: every fact or phenomenon that has a beginning arises when some certain combination of positive facts exists, provided certain other positive facts do not exist.

We tend to associate the idea of the cause of E_2 with the event E_1 that immediately precedes E_2 , rather than with any of the earlier states—i.e. permanent facts—that are also conditions of E_2 . (You can see this in the example of death caused by eating oysters.) The reason for this tendency is that E_1 begins to exist immediately before E_2 , whereas the other conditions may have pre-existed for an indefinite time. We see this tendency at work in the different logical fictions that even men of science resort to so as not to give the name 'cause' to anything that existed for an indeterminate length of time before the effect. Thus, rather than saying that *the earth* causes the fall of bodies, they ascribe it to *a force exerted by the earth* or *an attraction by the earth*; and they think of these abstractions as used up by each effort and therefore constituting at each successive instant a new fact that is simultaneous with the effect or immediately preceding it. . . .

·START OF A LONG FOOTNOTE·

An intelligent reviewer of this work in the *Prospective Review* [R. H. Hutton] disputes my thesis that *any* condition of a phenomenon may be—and on some occasions and for some purposes actually is—spoken of as 'the cause'. He says: 'We always apply the word "cause" to the element in the antecedents that exercises *force*. . . .' Also: 'everyone would feel' it to be wrong to say that the cause of a surprise was the sentinel's being off his post, but would feel that the 'allurement or force which *drew* him off his post might be so called. . . .'. I can't think that of these two:

- The event occurred because the sentinel was absent
- The event occurred because the sentinel was bribed to be absent

one is wrong and the other right. The only direct effect of the bribe was his absence, so the bribe could be called the remote cause of the surprise; but only because the absence was the immediate cause. I don't think anyone would accept one expression and reject the other unless he had a theory to support.

[The reviewer claimed that several statements implied by Mill's account—e.g. that a man's having bodily organs was part of the cause of his dying when he took poison—are things that no-one would say. Mill accepts this, and patiently repeats his explanation of why such things sound wrong though they are true. He continues:]

As for the assertion that nothing is called 'the cause' unless it exerts active force: I'll set aside the question of what 'active force' means, and will use the phrase in its popular sense. Well, then, of these two—

- He fell because his foot slipped in climbing a ladder
- He fell because of his weight

—which sounds better? The active force bringing about his fall was his weight, not the motion of his foot! [Mill gives other

examples in which the most intuitively-plausible candidate for the role of 'the cause' is *not* the force-exerting one, ending with:] The opening of flood-gates is said to be the cause of the flow of water; yet the active force is exerted by the water itself, and opening the flood-gates merely supplies a negative condition. The reviewer adds: 'Relations of space and time are absolutely passive conditions yet are absolutely necessary to physical phenomena; but no-one ever applies the word "cause" to these without being immediately stopped by those who hear him.' I have to disagree even with this. Few people would feel that it was wrong or strange to say that a secret became known because it was spoken of when X was within earshot (a condition of space), or that the cause why this tree is taller than that one is that it has been longer planted (a condition of time).

·END OF THE LONG FOOTNOTE·

Philosophically speaking, then, the cause is the sum total of the conditions, positive and negative—the whole of the contingencies of every sort from which the consequent invariably follows. But the negative conditions of any phenomenon can be all summed up in one phrase, 'the absence of preventing or counteracting causes', which spares us the wordy labour of listing them separately. The convenience of this form of expression is mainly based on the fact that in most cases cause C_1 's effects in counteracting cause C_2 can with strict scientific exactness be regarded as a mere extension of C_1 's own proper and separate effects. When gravity retards the upward motion of a projectile and deflects

it into a different trajectory, it is producing the very same •kind of effect—and even (as mathematicians know) the same •quantity of effect, as it does when causing an unsupported body to fall to the ground. When an alkaline solution mixed with an acid destroys its sourness, and prevents it from reddening vegetable blues, it's because the specific effect of the alkali is to combine with the acid and form a compound with totally different qualities. Causes of all sorts have this property of preventing the effects of other causes by virtue of the same laws according to which they produce their own.¹ This enables us to do without any mention of negative conditions: we can •establish the general axiom that all causes are liable to be counteracted in their effects by one another, and •limit the notion of cause to

- the sum of the positive conditions, and
- one negative condition, always the same one, namely the absence of counteracting causes;

·and just because the negative condition is always the same it can be silently understood, and in that spirit dropped from the story·.

§4.In most cases of causation a distinction is commonly drawn between •something that acts and •some other thing that is acted upon; between an •agent and a •patient [see Glossary]. Everyone would agree that both of these are •conditions of the phenomenon, but it would be thought absurd to call the latter the 'cause', that label being reserved for the former. But when we look into this we find that this distinction vanishes, or rather turns out to be only verbal. Its

¹ There are a few ·apparent· exceptions; for some properties of objects seem to be purely preventive, e.g. the property of opaque bodies by which they intercept the passage of light. This, as far as we can understand it, seems to be a case of an agency that shows up *only* in defeating the effects of another agency. If we knew what other relations to light, or what peculiarities of structure, opacity depends on, we might find that this is only an apparent exception to the general proposition in the text, not a real one. Either way, it needn't affect the practical application. The formula that includes all the negative conditions of an effect in the single one of the absence of counteracting causes is not violated by such cases as this; though if all counteracting agencies were like this there would be no point in employing the formula.

source is a mere fact about wording: the object that is said to be acted on—and is regarded as the scene in which the effect occurs—is usually included in the phrase by which the *effect* is spoken of, so that if it were also counted as part of the cause there would be the appearance of something's being incongruously said to cause itself. Return to falling bodies, and the question: 'What is the cause that makes a stone fall?' If the answer had been 'the stone itself', that would seem to contradict the meaning of the word 'cause'. So the stone is conceived as the patient, and the earth is represented as the agent or cause. [Mill wrote 'the earth (or, according to the common and most unphilosophical practice, an occult [see Glossary] quality of the earth)'.] But this is a superficial matter: we can conceive the stone as causing its own fall, as long as we word this so as to avoid the mere verbal incongruity. We might say: 'The stone moves toward the earth by the properties of the matter composing it'; and then there's nothing wrong with calling the stone itself the 'agent'. (Wanting to save the established doctrine that matter is inactive, men have usually preferred to say that the cause is not •the stone itself but •its weight or gravitation—an occult quality.)

Those who have defended a radical distinction between agent and patient have generally had this thought:

The 'agent' is what causes some state of, or some change in the state of, another object that is called the 'patient'.

But a little reflection will show that our way of speaking of phenomena as *states* of the various objects that take part in them. . . . is simply a sort of logical fiction, sometimes useful as one among several formulations, but never to be mistaken for an expression of a scientific truth. Even the attributes of an object that might seem with greatest propriety to be called 'states' of the object—its sensible qualities, its colour, hardness, shape, and the like—are really. . . . phenomena of

causation, in which the substance is the agent or producing cause, and the patient is our own organs and those of other sentient beings. What we call 'states of' objects are sequences of events into which the objects enter, usually as antecedents or causes; and things are never more *active* than when they are producing the phenomena in which they are said to be 'acted on'! According to the theory of gravitation, a falling stone is as much an agent as the earth is—the earth attracts the stone but is also attracted by it. When a sensation is produced in our organs, the laws of our organisation—and even the laws of our minds—are as directly operative in determining the effect as are the laws of the external object. We call prussic acid the 'agent' of a person's death, but the whole of his vital and organic properties are as actively instrumental as the poison in the chain of effects that kills him. In the process of education, we may call the teacher the 'agent', and the pupil the material that is acted on; but actually all the facts that pre-existed in the pupil's mind *act* either for or against the teacher's efforts. The agent in vision isn't light alone but light coupled with the active properties of the eye and brain, and with those of the visible object. The 'agent'/'patient' distinction is purely verbal; patients are always agents; in most natural phenomena, indeed, they are agents to such a degree that they react forcibly on the causes that acted on them. . . . All the positive conditions of a phenomenon are alike agents, alike active; and in any account of the cause that professes to be complete, none of them should be excluded except ones that have already been implied in the words used for describing the effect. . . .

§5. I should deal separately with the case of causation where the effect is to invest an object with a certain property—i.e. not to produce a certain phenomenon but to fit something

else for producing it. When sulphur, charcoal, and nitre are put together in a certain way and in certain proportions, the effect is not an explosion but a mixture that is explosive—i.e. that will explode in certain circumstances. The various natural and artificial causes that educate the human body or the human mind have for their principal effect not •to make the body or mind immediately do anything but •to endow it with certain properties—i.e. to ensure that in certain circumstances certain results will take place •in it or •as consequences of it. . . . Painting a wall white doesn't merely produce the sensation of white in those who see the wall; it confers on the wall the permanent property of giving that kind of sensation. In relation to the sensation, the painting of the wall is *a condition of a condition*; it is a condition of the wall's causing that particular fact. The wall may have been painted years ago, but it has acquired a property that has lasted till now, and will last longer; the antecedent condition needed to enable the wall to become in its turn a condition has been fulfilled once for all. In a case like this, where the immediate effect is a property produced in an object, no-one these days thinks that the property is a substantive entity—a special kind of *thing*—'inherent' in the object. What has been produced could be called a state of preparation in an object for producing an effect. . . . In the case of the gunpowder •this state of preparation consists in the particles' coming to be close to one another. In the example of the wall, •it consists in a new spatial closeness of the wall to the paint. In the example of the moulding influences on the human mind, •its involving spatial relations is only conjectural; even if we assume the materialistic hypothesis, there's still a *question* as to whether the increased ease with

which the well-trained brain sums up a column of figures is a result of some permanent new arrangement of some of its material particles. So we must content ourselves with what we know, and include among the effects of causes the *capacities* given to objects of being causes of other effects. A 'capacity' isn't a real thing existing in the object; it's merely a name for our belief that the object *will* act in a certain way *if* certain new circumstances arise. We can give this assurance of future events a fictitious objective existence by calling it 'a state of the object'. But unless the state consists in a spatial arrangement of particles (as with the gunpowder), it expresses no present fact and is merely a contingent future fact re-presented under another name. . . .

§6. I now present a distinction that is of first-rate importance both for clarifying the notion of cause and for blocking a plausible objection that is often brought against the view I have taken of the subject.

When I define the *cause* of x—in the only sense in which this book has any concern with causes—to be 'the antecedent that x invariably **follows**' I do not mean 'the antecedent that x invariably **has followed** in our past experience'. An account of causation in terms of 'has followed' would be open to Reid's very plausible objection that then night must be the cause of day, and day the cause of night, because day and night *have* invariably followed one another from the beginning of the world. It's essential to our use of 'cause' that we should believe not only •that the antecedent always *has* been followed by the consequent, but that as long as the present constitution of things¹ endures, it always *will* be so. And this isn't true of day and night. We believe that night will be followed by day not

¹ By 'the present constitution of things' I mean the ultimate laws of nature (whatever they may be), as distinct from the derivative laws and from the collocations [see Glossary]. The daily revolution of the earth (for example) is not a part of the constitution of things, because it could be terminated or altered by natural causes.

- under all imaginable circumstances, but only
- provided the sun rises above the horizon.

If the sun stopped rising (and for all we know, its doing so may be perfectly compatible with the general laws of matter), night could be eternal. On the other hand, if the sun is above the horizon, its light not extinct, and no opaque body between it and us, we firmly believe that unless there's a change in the properties of matter •this combination of antecedents will always be followed by *day*; that •if this combination were indefinitely prolonged, it would always be day; and that •if the same combination had always existed, it would always have been day, quite independently of night as a previous condition. That's why we don't call night the 'cause' of day, or even a 'condition' of day. The only conditions of day are •the existence of the sun (or some such luminous body), and •there being no light-blocker in a straight line¹ between that body and the part of the earth where we are situated; and the combination of these, without any superfluous details, constitutes the cause. This is what writers mean when they say that the notion of cause involves the idea of *necessity*. If 'necessity' has any agreed meaning it is *unconditionalness*. To say that x is necessary, that x *must* be, is to say that x *will* be, no matter what else happens. The succession of day and night is obviously not 'necessary' in this sense, because it is conditional on the occurrence of other antecedents. If x will be followed by y when and only when z is the case, x isn't the cause of y even if no instance of x has ever occurred without y following it.

...So we can define the *cause* of a phenomenon to be the antecedent (or combination of antecedents) which it invariably *and unconditionally* follows. Or if we adopt

the convenient usage in which 'cause' is confined to the combination of *positive* conditions, then we must replace 'unconditionally' by 'subject to no conditions except negative ones'.

Some may want to object:

'The sequence of night and day is invariable in our experience; we have as much ground in •this case as experience can give in •any case for recognising the two phenomena as cause and effect. To say that more is necessary—to require a belief that the succession is unconditional—is to admit that causation involves an element of belief that isn't derived from experience.

I answer that it is experience itself that teaches us that one uniformity of sequence is conditional and another unconditional. When we judge that the succession of night and day is a derivative sequence, depending on something else, we are going by experience. It's the evidence of experience that convinces us that day could exist without being followed by night, and night without being followed by day. To say as Tulloch does that these beliefs are 'not generated by our mere observation of sequence' is to forget that twice in every 24 hours, when the sky is clear, we have decisive evidence that the cause of day is the sun. We have empirical knowledge of the sun that justifies us on empirical grounds in concluding that if the sun were always above the horizon there would •always• be day even if there had been no night, and that if the sun were always below the horizon there would •always• be night even if there had been no day. [Mill adds a reminder that if x is only a conditionally invariable antecedent of y, then x's status as an invariable antecedent of y is fragile.]

¹ I say 'straight line' for brevity and simplicity. In reality the line in question is not exactly straight: because of refraction, we actually see the sun for a short interval during which the opaque mass of the earth is interposed in a direct line between the sun and our eyes. This provides us with a limited version of the luxury of seeing round a corner.

[Mill now offers a paragraph explaining how a combination of causes (unconditionally invariable antecedents) can generate conditionally invariable relations. He uses this to rebut something said by a contemporary philosopher. And other contemporaries—especially Whewell—are criticised in a further paragraph that is all about terminology.]

§7. Does a cause always relate to its effect as antecedent to consequent? Don't we often speak of two simultaneous facts as cause and effect—fire as the cause of warmth, the sun and moisture as the cause of vegetation, and the like? Since a cause doesn't necessarily go out of existence because its effect has been produced, the two do very generally coexist; and there are appearances, and common expressions, that seem to imply not only that causes *can* but that they *must* be contemporaneous with their effects. The scholastics have had as a dogma 'When the cause ceases, so does the effect' [Mill gives this in Latin], and there was a time, it seems, when it was generally believed that the continuance of an effect requires the continued existence of the cause. Kepler's numerous attempts to explain the motions of the planets on mechanical principles were doomed by his always supposing that the agency that set them in motion must continue to operate in order to keep up the motion it at first produced. Yet there have always been many familiar examples of effects continuing long after their causes had ceased. Sun-stroke gives a person brain-fever; will the fever go off as soon as he is moved into the shade? A sword is run through someone's body; must the sword remain in his body for him to continue to be dead? Once a plough has been made, it remains a ploughshare without any continuance of heating and hammering. . . . On the other hand, the pressure that forces up the mercury in a vacuum-tube must be continued in order to keep it up there. This (it may be replied)

is because another force—the force of gravity—is acting continually, and would bring the mercury down again if it weren't counterbalanced by an equally constant force. Well, then: a tight bandage causes pain, which will sometimes stop as soon as the bandage is removed. The illumination that the sun diffuses over the earth ceases when the sun goes down.

So there's a distinction to be drawn. Sometimes the conditions needed for the first production of a phenomenon are also needed for its continuance; though more often its continuance requires no condition except negative ones. •Most things, once produced, continue as they are until something changes or destroys them; but •some require the permanent presence of the agencies that produced them at first. We could choose to regard these as instantaneous phenomena, needing to be renewed at each instant by the cause that first generated them. The illumination of any given point of space has always been regarded as an instantaneous fact, which perishes and is perpetually renewed as long as the required conditions obtain. This way of talking spares us the necessity of admitting that the continuance of a cause is *ever* required to maintain the effect; because we can say that the cause is required not to •maintain the effect but to •reproduce it or else to •counteract some force tending to destroy it. This may be a convenient terminology, but that's all it is—terminology. The fact remains that in some cases (though only a minority) the continuance of the conditions that produced an effect is necessary for the continuance of the effect.

Is it strictly necessary that a cause should precede—by ever so short an instant—the production of its effect? For my present purposes this doesn't matter. There certainly are cases where the effect follows with no interval that we can detect; and when there is an interval we can't tell how

many intermediate links, imperceptible to us, may fill it up. But even if an effect *can* start simultaneously with its cause, this doesn't affect the view of causation that I am defending. Whether or not the cause and its effect must be successive, •it's the beginning of a phenomenon that implies a cause, and •causation is the law of the succession of phenomena. If these two axioms are granted, we can drop the words 'antecedent' and 'consequent' as applied to cause and effect, though I don't see any need to do so. I have no objection to defining a *cause* as an assemblage of phenomena such that, when it occurs, some other phenomenon invariably starts or has its origin. It doesn't matter whether the effect •coincides with the last of its conditions, or •immediately follows it. It doesn't *precede* it; and when we are wondering which of two coexistent phenomena is the cause and which the effect, we rightly regard the question as answered if we can ascertain which of them came first.

§8. . . . A single phenomenon is often seen to be followed by several different sorts of effects that happen simultaneously. . . .

- The sun produces the motions of the planets, daylight, and heat.
- The earth causes the fall of heavy bodies, and the phenomena of the magnetic needle.
- A crystal of galena [lead sulphide] causes the sensations of hardness, weight, cubic shape, gray colour, and many others between which we can trace no interdependence.

The terminology of 'properties' and 'powers' is specially adapted to this sort of case. When a single phenomenon is followed. . . by effects of radically different kinds, it's usual to say that each different sort of effect is produced by a different 'property' of the cause. Thus we distinguish

the gravitational 'property' of the earth from its magnetic 'property'; the gravitational, light-making and heat-making 'properties' of the sun; a crystal's 'properties' of colour, shape, weight, and hardness. These are mere phrases: they don't explain anything or add anything to our knowledge of the subject; but considered as abstract names denoting the connection between an object and the different effects it produces, they're a powerful instrument of abridgment and of the acceleration of thought that abridgment brings about.

All this leads to a conception that we'll find to be important, namely that of a permanent cause, or original [see Glossary] natural agent. A number of permanent causes have existed for as long as the human race has, and indeed longer—probably vastly longer. The sun, the earth, and the planets are such permanent causes, as are their various constituents—air, water, and other simple or compound substances of which nature is made up. These have existed, and their effects have taken place, from the very beginning of our experience. But we can't explain the origin of the permanent causes themselves. Why *these* particular natural agents and no others existed originally, why they occur in such-and-such proportions, why they are distributed in such-and-such a way throughout space—we don't know the answers to any of this. Furthermore, we can't discover anything regular in the distribution itself. Given how these causes or agents are distributed in •one part of space, there is no uniformity or law that will let us conjecture what the distribution is •elsewhere. So the coexistence of this and that primeval cause is a mere casual coincidence, so far as we are concerned; so we don't classify as a case of causation or a law of nature any regularity (of following or coexisting) between the effects of this one and the effects of that one. We have no basis for expecting such regularities, except where we have direct evidence about how the relevant

natural agents—the things on whose properties the regularities ultimately depend—are distributed in space. These permanent causes aren't always objects; they are sometimes events, i.e. periodical cycles of events (that's the only sense in which events can be 'permanent'). The earth is a permanent cause or primitive natural agent, and so is its rotation. It's a cause that has always produced. . . .the succession of day and night, the ebb and flow of the sea, and many other effects; and because we can't assign a cause for the rotation itself (except by guessing!), it is entitled to be classified as a primeval or original cause. But it's only the *origin* of the rotation that is mysterious to us; once the rotation has begun, its continuance is accounted for by •the first law of motion (that of the permanence of rectilinear motion once it has been started) combined with •the gravitation of the parts of the earth toward one another. [In that paragraph, 'produced the succession of day and night' replaces Mill's 'produced (by the aid of other necessary conditions) the succession of day and night'. Throughout §8, almost every statement about x causing y includes a clause about other necessary conditions; these clauses are omitted here in the interests of brevity.]

All phenomena that begin to exist, i.e. all but the primeval causes, are immediate or remote effects those primitive facts or of some combination of them. Throughout the known universe no thing comes into existence and no event happens that isn't connected by a uniformity or invariable sequence with one or more phenomena that preceded it; so that it

will happen again as often as those phenomena occur again and no counteracting cause coexists. These antecedent phenomena were connected in a similar way with some that preceded them, . . . and so on until we reach our limit, the properties of one or more primeval causes. The whole of the phenomena of nature were therefore the necessary—i.e. the unconditional—consequences of some former collocation [see Glossary] of the permanent causes.

The state of the whole universe at any instant, I believe, is the consequence of its state at the previous instant; so that someone who knew

- all the agents that exist right now,
- their collocations in space, and
- all their properties, i.e. the laws of their agency,

could predict the whole history of the universe from now on, unless some new volition of a power capable of controlling the universe should take over. [How did *volition* get into the picture? Mill is talking about a change in the *basic* causal organisation of the universe, and is assuming that if that were to change it would have to be because God—'a power capable of controlling the universe'—decided to change it.]¹ [Mill adds the somewhat isolated remark that if any one total state of the universe came around a second time, the whole history of the universe would repeat itself for ever, 'like a circulating decimal'. Then he gets back on track:] The whole series of events in the history of the universe, past and future, is intrinsically capable of being constructed *a priori* by anyone who is acquainted with the

¹ In this footnote Mill mentions those who think that human volition constitutes an exception to the determinist thesis that whatever happens is caused to happen. He says that he'll deal with this thoroughly in VI.2, and right now will make just one point:] These metaphysicians base their objection 'to determinism' on the claim that it conflicts with our consciousness. I think they mistake the proposition that consciousness testifies against. If they look into themselves carefully, they'll find that what their consciousness objects to is the thesis that human actions and volitions are *necessary* in the everyday sense of that term. I agree with them about that. But the statement 'A person's actions necessarily follow from his character' really means what is meant in *any* statement about causation, namely that •the person invariably *does* act in conformity to his character, and •that anyone who thoroughly knew his character could predict with certainty how he would act in any supposable case. They probably wouldn't find *this* contrary to their experience or revolting to their feelings. And no-one claims more than this, except for Asiatic fatalists.

original distribution of all natural agents, and with all their properties—i.e. the laws of succession existing between them and their effects. Of course this would require superhuman powers of combination and calculation, even for someone who knew all the original facts.

§9. Because everything that occurs is determined by laws of causation and collocations of the original causes, it follows that coexistences among effects can't be themselves the subject of any similar set of laws distinct from laws of causation. There are uniformities of coexistence and succession among effects; but these must all result from the identity or the coexistence of their causes; if the causes didn't coexist, nor could the effects. And these causes being effects of still earlier causes, and these of others, . . . until we reach the primeval causes, it follows that (apart from effects that can be traced back to a single cause) the coexistences of phenomena can't be universal unless the coexistences of their primeval causes can be reduced to a universal law; but we have seen that they can't. So there are no original and independent—i.e. no *unconditional*—uniformities of coexistence, between effects of different causes; if they coexist, it's only because their causes have happened to coexist. The only independent and unconditional coexistences that are invariable enough to have a claim to be laws are between different and mutually independent effects of a single cause, i.e. between different properties of the same natural agent. This portion of the laws of nature will be treated of in the chapter 22.2 .

§10. Since the first edition of this work, the sciences of physical nature have made a great advance in generalisation, through the doctrine of the conservation or persistence of force. Building and laying out this imposing edifice of theory has for some time been the principal occupation of the most systematic physicists. It consists of two stages: **(1)** one

consisting of ascertained fact, and **(2)** one containing, along with some fact, a large element of hypothesis.

(1) It is proved by numerous facts, some experimental, some informal, that agencies that had been regarded as distinct and independent sources of force—heat, electricity, chemical action, nervous and muscular action, momentum of moving bodies—are interchangeable with one another in definite and fixed quantities. It had long been known that these dissimilar phenomena had the power, under certain conditions, to produce one another; what is new in the theory is a more accurate estimation of what this production consists in. What happens is that phenomena of one kind disappear and are replaced by phenomena another kind, and that there is an equivalence in quantity between the phenomena that have disappeared and the ones that have replaced them; so that if the process is reversed, the exact same quantity that had disappeared will re-appear. For example, the amount of heat that will raise the temperature of a pound of water one degree centigrade will, if used in the expansion of steam, lift a weight of 772 pounds one foot, or a weight of one pound 772 feet; and the same quantity of heat can by certain means be recovered through the expenditure of exactly that amount of mechanical motion.

The establishment of this comprehensive law has led the scientific world to change how it speaks about what are called the 'forces of nature'. Before this correlation between very different phenomena had been discovered, their unalikehood had caused them to be regarded as upshots of such-and-such distinct forces. Now that they are known to be convertible into one another without loss, they are spoken of as all the results of one single force, showing itself in different ways. This force (it is said) can only produce a limited and definite quantity of effect, but it always *does* produce that definite quantity; and produces it (according

to circumstances) in one or another of the forms, or divides it among several, but so as always to make up the same sum; and no one of the manifestations can be produced except by the disappearance of the equivalent quantity of another, which in its turn, in appropriate circumstances, will re-appear undiminished. This mutual interchangeability of the forces of nature according to fixed numerical equivalents is the part of the new doctrine that rests on unchallengeable fact. (The judgments about equivalents are based on a scale of numerical equivalents established by experiment.)

An indefinite and perhaps immense interval of time may elapse between the disappearance of the force in one form and its re-appearance in another. A stone thrown up into the air with a given force and falling back immediately will, by the time it reaches the earth, recover the exact amount of mechanical momentum which was expended in throwing it up (minus a small portion of motion given to the air). But if the stone lands on a high ledge it may not fall back for years, perhaps ages, and until then the force used in raising it is temporarily lost, being represented only by what the language of the new theory calls 'potential energy'. The coal embedded in the earth is considered by the theory as a vast reservoir of force that has remained dormant for many geological periods, and will remain dormant until by being burned it gives out the stored-up force in the form of heat. . . . This means simply that when the coal does at last. . . .generate a quantity of heat (transformable like all other heat into mechanical momentum and the other forms of force), this heat is the re-appearance of a force derived from the sun's rays, expended myriads of ages ago in the growth of the organic substances that were the material of the coal.

(2) The theory of conservation of force has a part that is a combination of fact and hypothesis. Briefly, it is as follows:

The conservation of force is really the conservation of motion. In each interchange between forms of force it is always motion that is transformed into motion.

This requires the assumption of motions that are hypothetical. The supposition is that •there are molecular motions that appear to our senses only as heat, electricity, etc.—oscillations, invisible to us, among the minute particles of bodies; and that •these molecular motions can be changed into molar motions (motions of masses), and vice versa. We do have positive evidence of the existence of molecular motion in these manifestations of force. In chemical reactions, for instance, the particles separate and form new combinations, often with a great visible disturbance of the mass. And with heat the evidence is equally conclusive, since heat expands bodies (i.e. causes their particles to move apart), and enough heat changes them from solid to liquid, or from liquid to gaseous. Again, the mechanical actions that produce heat—friction, and the collision of bodies—must from the nature of the case produce a shock, i.e. an internal motion of particles, which we often find is so violent as to break them apart. Such facts are thought to justify the conclusion that we were wrong when we thought that heat causes the motion of particles, and that really the motion of particles causes heat; the original cause of both being the earlier motion—molar or molecular, collision of bodies or burning of fuel—that formed the heating agency. This conclusion already contains an hypothesis; but at least the supposed cause, the internal motion of molecules, is a *vera causa* [see Glossary]. But in order to reduce the conservation of force to conservation of motion, it was necessary to attribute to motion the heat propagated through apparently empty space from the sun. This required the supposition (already made for the explanation of the laws of light) of a superfine *ether* pervading space; we can't feel it, but it must have the

property that constitutes matter, namely *resistance*, because waves are propagated through it by an impulse from a given point. This ether must be supposed (and the theory of light doesn't require this) to penetrate into the minute crevices in all bodies. The story goes like this:

Vibratory motion in the heated mass of the sun is passed on to the particles of the surrounding ether, and through them to the particles of the same ether in the gaps and crevices of terrestrial bodies; and this is done with enough mechanical force to make the particles of those bodies vibrate strongly enough to make the bodies expand and create the sensation of heat in sentient creatures.

This is all hypothesis, but I'm not expressing doubt as to its legitimacy as hypothesis. It seems to follow from this theory that *force* can and should be defined as *matter in motion*. But this can't be right because, as we have seen, the matter doesn't have to be actually moving. It isn't necessary to suppose that the motion manifested when the coal burns is actually taking place among the molecules of the coal during its time in the earth;¹ certainly not in the stone at rest on the high ledge onto which it has been thrown. The true definition of *force* must be **potentiality of motion**; and what the doctrine. . . amounts to is not that •there is at all times the same quantity of actual motion in the universe, but that •the possibilities of motion are limited to a definite quantity that can't be added to and can't be exhausted; and that all actual motion is a draft on this limited stock [this is a metaphor based on the idea of a bank-draft saying 'Withdraw £n from account number. . .']. It needn't all have existed *ever* as actual motion. There's a vast amount of potential motion in the

universe in the form of gravitation, and it would be a great abuse of hypothesis to suppose that *that* was stored up by the expenditure of an equal amount of actual motion in some former state of the universe! Nor does the motion produced by gravity take place, as far as we know, at the expense of any other motion of any kind.

If we adopt this theory as a scientific truth, thus accepting its change in our conception of the most general physical agencies, does this require •any change in the view I have taken of causation as a law of nature? As far as I can see, •none whatever. The manifestations that the theory regards as modes of motion are just as distinct and separate phenomena when attributed to •a single force as when attributed to •several. Whether the phenomenon is called a transformation of force or the generation of one, it has its own set(s) of antecedents with which it is connected by invariable and unconditional following; and that set, or those sets, of antecedents are its cause. [Mill now embarks on a long discussion of how he should word his theory in the light of the conservation theory. He refers in friendly terms to a detailed discussion of this by Bain, to whom he attributes the conclusion that:]

In the assemblage of conditions that constitutes the cause of a phenomenon we must distinguish two elements—•the presence of a force and •the collocation or position of objects that is required for the force to undergo the particular transmutation that constitutes the phenomenon.

[Mill accepts this, sort of, but argues at great length that this doesn't really require him to alter any of his formulations. He accepts and indeed insists that the cause of any change

¹ I understand that the accredited authorities *do* suppose that molecular motion, equivalent in amount to what will be manifested when the coal burns, is actually taking place during the whole of that long interval, not in the coal but in the oxygen that will then combine with it. You can see how purely hypothetical this supposition is; and I venture to say that it is unnecessarily and extravagantly hypothetical.

must include a change: 'To produce a bonfire there must not only be fuel and air and a spark, which are collocations [see Glossary], but chemical action between the air and the materials, which is a force.' But the insistence on motion, he says, is simply wrong unless we include potential motion; and we must be careful about what we mean by that: 'The force said to be laid up and merely potential is no more a really existing thing than any other properties of objects are really existing things. The phrase "potential force" or "potential motion" is a mere linguistic device that is convenient for describing the phenomena.' He concludes:]

We thus see that no new general conception of causation is introduced by the conservation theory. The indestructibility of force doesn't interfere with the theory of causation any more than the indestructibility of matter does. . . . It only enables us to understand better than before the nature and laws of some of the sequences.

This better understanding, however, lets us join Bain in accepting the expenditure or transfer of energy as one of the tests for distinguishing causation from mere concomitance. If the effect being explained includes matter's beginning to move, then any of the objects present that has lost motion has contributed to the effect; and this is the true meaning of the thesis that the cause is the one of the antecedents that exerts active force.

§11. This is the place to discuss a rather ancient doctrine about causation that has been revived in recent years and now shows more signs of life than any other theory of causation that conflicts with the one I have been defending.

According to the theory in question, the only cause of phenomena is *mind*, or more exactly *will*. The type [see Glossary] of causation, as well as the only source for our idea of it, is our own voluntary agency. The theory's friends say:

Our voluntary agency is our only direct evidence of causation. We know that we can move our bodies. Regarding the phenomena of inanimate nature, all we have other direct knowledge of is what happens before what. But in our voluntary actions we're conscious of •power before we have experience of •results. An act of volition, whether or not followed by an effect, is accompanied by a consciousness of effort, 'of force exerted, of power in action, which is necessarily causal or causative' [quoted from Francis Bowen]. This feeling of energy or force, inherent in an act of will, is knowledge *a priori*; assurance before experience that we have the power to cause effects. So volition is more than an •unconditional antecedent; it is a •cause in a different sense of 'cause' from that in which physical phenomena are said to cause one another. It is an efficient cause [see Glossary].

It's easy to move from this to the doctrine that volition is the *only* efficient cause of all phenomena. 'It is inconceivable that dead force could continue unsupported for a moment beyond its creation. We cannot even *conceive of* change or phenomena without the energy of a mind' [quoted from R. H. Hut-ton]. And another writer of the same school says: 'The word "action" itself has no real significance except when applied to the doings of an intelligent agent. Let anyone conceive, if he can, of any power, energy, or force inherent in a lump of matter' [Bowen again]. Phenomena may seem to be produced by physical causes but they are really produced, say these writers, by the immediate agency of mind. Everything that doesn't proceed from a human (or, I suppose, an animal) will proceed, they say, directly from divine will. The earth is not moved by the combination of a •centripetal and a •projectile force; this is a mere way of speaking that helps to make our conceptions easier. The earth (they say) is moved by the

direct volition of an omnipotent Being, on a path coinciding with the one that we deduce from the hypothesis of •these two forces.

As I have so often said, the general question of the existence of •efficient causes doesn't fall within the limits of my present subject; but a theory that represents •them as capable of being known by humans, and passes off as efficient causes what are only physical or phenomenal causes, belongs to logic as much as to metaphysics, and is a fit subject for discussion here.

As I see it, a volition isn't an efficient cause but simply a physical cause. Our will 'causes' our bodily actions in exactly and only the sense in which cold causes ice, or a spark causes an explosion of gunpowder. The volition, a state of our mind, is the antecedent; the motion of our limbs in conformity to the volition is the consequent. This sequence is not a subject of direct consciousness in the sense intended by the theory. The antecedent and the consequent are indeed subjects of consciousness; but the connection between them is a subject of experience. Our consciousness of the volition doesn't contain in itself any *a priori* knowledge that the muscular motion will follow. If our nerves of motion were paralysed. . . .and had been so all our lives, I don't see the slightest reason to think that we would ever (unless told by other people) have known anything of volition as a physical power, or been conscious of any tendency in feelings

of our mind to produce motions of our body. . . . Would we in that case have had the physical feeling that I suppose these writers mean by 'consciousness of effort'? I don't see why not, because that physical feeling is probably a state of nervous sensation beginning and ending in the brain, without involving the motor apparatus; but we certainly wouldn't have called it anything like 'effort'. . . . If we were conscious of this sensation, we'd have been conscious of it, I think, only as a kind of uneasiness accompanying our feelings of desire.

Hamilton argues well against the theory in question, thus:

'It is refuted by the consideration that between •the overt fact of corporeal movement that we know and •the internal act of mental determination that we also know, there intervenes a series of intermediate agencies of which we know nothing; so we *can't* be conscious of any causal connection between the volition and the movement, as this hypothesis asserts. . . . A paralytic learns after the volition that his limbs don't obey his mind; and it's only after the volition that the healthy man learns that his limbs do obey the mandates of his will.'¹

Those I am arguing against have never produced, and don't claim to produce, any positive evidence that the power of our will to move our bodies would be known to us independently of experience.² What they say about this is that

¹ . . . This acute thinker has a theory of causation that is all his own. As far as I know it has never been analytically examined, but I think it can be refuted as completely any one of the false or insufficient psychological theories that strew the ground under his potent metaphysical scythe. (Since writing that I have examined and controverted his theory in my *Examination of Sir William Hamilton's Philosophy*, ch. 16.)

² Bowen disagrees: 'The result to be accomplished is preconsidered or meditated, and is therefore known *a priori* or before experience.' This merely says that when we will a thing we have an idea of it. . . .but that doesn't imply a prophetic knowledge that it will happen. You may object: 'The first time we exerted our will, when we had no *experience* of any of our powers, we must have known that we had those powers, because we can't will something that we don't believe to be in our power.' But that's a merely verbal impossibility. We can **desire** something that we don't know to be in our power, and when we find by experience that our bodies move according to our desire *then* we can pass into the more complicated mental state that is termed **will**. . . .

the production of physical events by a will seems to carry its own explanation with it, while the action of matter on matter seems to require something else to explain it, and is even 'inconceivable' unless we suppose that some *will* intervenes between the apparent cause and its apparent effect. So they base their argument on an appeal to what they think to be the inherent laws of our conceptive faculty, mistaking for the innate laws of that faculty its acquired habits based on the spontaneous tendencies of its uncultured state. The sequence from the will to move a limb and the actual motion is one of the most direct and instantaneous of all the sequences we observe, and is familiar to every moment's experience from our earliest infancy. It is more familiar to us than any succession of events exterior to our bodies, and especially more so than any other case of the apparent beginning (as distinguished from the mere passing on) of motion. Our mind naturally tends to be constantly trying to help its conception of unfamiliar facts by assimilating them to familiar ones. And so our voluntary acts. . . . in the infancy and early youth of the human race are spontaneously taken as the type [see Glossary] of causation in general, and all phenomena are supposed to be directly produced by the will of some sentient being. I shan't describe this primitive idol-worship in the words of Hume or of any of his followers; rather, I'll take the words of a religious metaphysician, Reid, in order to bring out that all competent thinkers are unanimous on this topic.

·START OF QUOTATION FROM REID·

'When we turn our attention to external objects and begin to exercise our rational faculties about them, we find that there are some motions and changes in them that we have power to produce, and that many must have some other cause. Either the objects must have life and active power, as

we have, or they must be moved or changed by something that has life and active power, as external objects are moved by us.

'Our first thoughts seem to be that the objects in which we perceive such motions have understanding and active power as we have. "Savages", says the Abbé Raynal, "wherever they see motion that they can't account for, postulate a soul." All men can be considered as "savages" in this respect, until they can be taught and can use their faculties better than savages do. . . .

'Raynal's remark is sufficiently confirmed both from fact and from the structure of all languages.

'Primitive nations really do believe that the sun, moon, and stars, the earth, sea, and air, and fountains and lakes have understanding and active power. Savages find it natural to bow down to these things and beg for their favour, as a kind of idolatry.

'All languages carry in their structure the marks of their having been formed at a time when this belief prevailed. The division of verbs and participles into active and passive, which is found in all languages, must have been originally intended to distinguish what is really active from what is merely passive; and, in all languages we find active verbs applied to the sorts of things in which, according to Raynal, savages think there is a soul.

'Thus we say "The sun rises and sets", "The moon changes", "The sea ebbs and flows", "The winds blow". Languages were formed by men who believed these objects to have life and active power in themselves, and so for them it was proper and natural to report such motions and changes with active verbs.

'There's no surer way of tracking what nations believed before they had records than by the structure of their language; despite the changes produced in it by time, a language will

always bear traces of the thoughts of those who invented it. When we find the same beliefs indicated in the structure of *all* languages, those beliefs must have been common to the whole human species when languages were being invented.

'When a few people with superior intellectual abilities find leisure for speculation, they begin to do science [Reid writes: 'to philosophise'], and they soon discover that many of the things they used to regard as thinking and active are really lifeless and passive. This is a very important discovery. It elevates the mind, frees men from many ignorant superstitions, and opens the door to further discoveries of the same kind.

'As science advances, life and activity in natural objects retreats, leaving the objects dead and inactive. We find that rather than •moving voluntarily they •are moved necessarily; rather than •acting they are •acted-upon; and nature appears as one great machine in which one wheel is turned by another, that by a third; and the scientist doesn't know how far back this necessary sequence may reach.' [Reid, *Essays on the Active Powers of Man* IV.3.]

•END OF QUOTATION FROM REID•

So there's a spontaneous tendency of the intellect to explain all cases of causation by assimilating them to the intentional acts of voluntary agents like itself. This is the instinctive philosophy of the human mind in its earliest stage, before it has become familiar with invariable sequences other than those between volitions and voluntary acts. As the notion of fixed laws of succession among *external* phenomena gradually takes hold, the propensity to explain all phenomena in terms of voluntary agency slowly gives way. But the suggestions of daily life continue to be more powerful than those of scientific thought, so the original instinctive philosophy maintains its ground in the mind, underneath the growths obtained by cultivation [see Glossary], and keeps

up a constant resistance to their driving their roots deep into the soil. The theory I'm attacking is fed by that substratum. Its strength lies not in argument but in its link with an obstinate tendency of the infancy of the human mind.

There's plenty of evidence that this tendency isn't the result of an inherent mental law. The history of science right from the beginning shows that mankind haven't been unanimous in thinking either that **(i)** the action of matter on matter *wasn't* conceivable or that **(ii)** the action of mind on matter *was*. To some thinkers, ancient and modern, **(ii)** has seemed much more inconceivable than **(i)**. Sequences that are entirely physical and material, as soon as they became familiar, came to be thought perfectly natural, and were regarded not only as not needing to be explained but as being able to explain other sequences—and even of serving as the ultimate explanation of things in general.

One of the ablest recent supporters of the volitional theory [Hutton] has provided an historically true and philosophically sharp account of the Greek philosophers' failure in physical inquiry—an account in which, it seems to me, he unconsciously depicts his own state of mind:

'Their stumbling-block concerned the nature of the evidence they expected for their conviction. . . They hadn't grasped that they mustn't expect to understand the •processes of external causes but only their •results; so the whole physical philosophy of the Greeks was an attempt to identify mentally the effect with its cause, to probe for a •connection that was not only necessary but natural—meaning that •it would carry within itself some reason why this antecedent should produce this consequent—and they confined themselves to looking such reasons.'

That is, they weren't content merely to know that one phenomenon x was always *followed* by another y; they thought

that science's true aim was to perceive something in x's nature from which they could have known or presumed *previous to trial* that it would be followed by y. . . . To complete his statement of the case, the quoted writer should have added that these early speculators not only •had that aim but •thought they had achieved it. The writer can see plainly that this was an error, because he doesn't believe that any relations between material phenomena can account for their producing one another; but the Greeks' persistence in this error shows that their minds were in a very different state; the assimilation of physical facts to other physical facts gave them the kind of mental satisfaction that we connect with the word 'explanation'. . . . When Thales and Hippo held that moisture was the universal cause and external element of which all other things were merely sensible manifestations; when Anaximenes said the same thing about air, Pythagoras about numbers, and the like; they all thought they had found a real explanation and were content to settle for this as ultimate. The ordinary sequences of the external universe seemed to them. . . .to be inconceivable without the supposition of some universal agency to connect the antecedents with the consequents; but they didn't think that mental volition was the only agency that fulfilled this requirement. Moisture, or air, or numbers, carried to their minds a precisely similar impression of making intelligible what was otherwise inconceivable, and gave the same full satisfaction to the demands of their conceptive faculty.

It wasn't only the Greeks who 'wanted to see some reason why the physical antecedent should produce this particular consequent'. . . . Among modern philosophers, Leibniz laid it down as a self-evident principle that *all* physical causes must contain in their own nature something making it intelligible that they should be able to produce the effects that they do produce. Far from admitting volition as the only kind of

cause that carries internal evidence of its own power, and as the real bond of connection between physical antecedents and their consequents, he demanded some naturally and intrinsically efficient [see Glossary] physical antecedent as the bond of connection between volition itself and its effects. He clearly refused to admit the will of God as a sufficient explanation of anything but miracles; and insisted on finding something that would account better for the phenomena of nature than a mere reference to divine volition.

And the action of mind on matter (which, we're being told, needs no explanation and itself explains all other effects) has seemed to some thinkers to be itself the grand inconceivability. This was the difficulty the Cartesians were trying to solve with the system of 'occasional causes'. They couldn't conceive that thoughts in a mind could produce movements in a body, or that bodily movements could produce thoughts. They couldn't see any necessary connection, any *a priori* relation, between a motion and a thought. And their insistence—greater than any other philosophical school before or since—that their own minds were measure of all things led them to refuse on principle to believe that Nature had done what they couldn't see any reason why she *must* do, so they said it was impossible that a material and a mental fact could be causes one of another. They regarded them as mere 'occasions' on which the real agent, God, thought fit to exert his power as a cause. When a man wills to move his foot, they said, it's not his will that moves it; God moves it on the occasion of the man's will. And when they looked more carefully into the action of matter on matter they found *this* inconceivable too, and therefore (according to their logic) impossible. The *deus ex machina* [see Glossary] was ultimately called in to produce a spark on the occasion of a flint and steel coming together, or to break an egg on the occasion of its falling on the ground.

All this shows that mankind in general is disposed not to be satisfied with knowing that one fact is invariably antecedent and another consequent, but to look for something that may seem to *explain* their being so. But we also see that this demand can be completely satisfied by a purely physical agency, provided it's much more familiar than what it is invoked to explain. To Thales and Anaximenes it seemed inconceivable that the antecedents that we see in nature should *produce* the consequents, but perfectly natural that water or air should produce them. The writers I am opposing in this section declare this to be inconceivable, but they can conceive that mind or volition is an efficient cause; while the Cartesians couldn't conceive even that, but briskly declared that the only conceivable mode of production of any fact whatever is the direct agency of an omnipotent being; all of which is further evidence for something that finds new confirmation in every stage of the history of science—namely that •what people can conceive and what they can't is very much an affair of accident, and depends entirely on their experience and their habits of thought; that •by cultivating the required associations of ideas people can make themselves unable to conceive any given thing, and make themselves able to conceive most things, however inconceivable these may at first appear; and that •the facts in each person's mental history that determine what is or isn't conceivable to him also determine which sequences in nature will appear to him so natural and plausible as to need no other proof of their existence, and to be evident by their own light independently of experience and of explanation.

By what rule can we decide between one theory of this sort and another? The theorists don't direct us to any external evidence; each appeals to his own subjective feelings.

One (X) says:

The succession C, B appears to me more natural, conceivable, and intrinsically credible than the succession A, B; so you are wrong in thinking that B depends on A; I am certain—though I can't give any other evidence of it, that C comes between A and B, and is the real and only cause of B.

Another (Y) answers:

The successions C, B and A, B appear to me equally natural and conceivable, or the latter more so than the former. A is quite capable of producing B without any other intervention.

A third (Z) says:

Like X I can't conceive that A can produce B; but I don't share his view that C produces B, because I find the sequence D, B more natural than C, B, so I prefer my D theory to the C theory.

The only universal law operating here is the one saying that each person's conceptions are governed and limited by his individual experiences and habits of thought! We're justified in saying of all three, what each of them already believes of the other two, namely that they exalt into an original law of the human intellect and of outward nature one particular sequence of phenomena that they find more natural and more conceivable than other sequences, only because it is more familiar. And I apply this judgment to the theory that volition is an efficient cause.

Before leaving this subject I must mention the additional fallacy contained in the inference from this theory that because volition is an efficient cause, therefore it is the only cause and the direct agent in producing even what is apparently produced by something else. Volitions are not known to produce anything directly except activity in the nerves, because the will influences even the muscles

only through the nerves. Suppose we grant that every phenomenon has an efficient [see Glossary] cause and not merely a phenomenal cause, and that volition in the case of the special phenomena that are known to be produced by it is that efficient cause; are we therefore to say (as these writers do) that because we know of no other efficient cause, and oughtn't to assume one without evidence, *there is no other*, and volition is the direct cause of all phenomena? [Mill scornfully dismisses this. If our volition is an efficient cause, it's the only one we *can* be conscious of because 'it is the only one that exists within ourselves'; and it's absurd to infer that volition is the only efficient cause in the universe. Mill likens this to the inference that because we know for sure that there is life on this planet, we can infer that there is life on every heavenly body. He concludes:] I ascribe to certain other creatures a life like my own, because they show the same sort of signs of it as I do. . . . Earth, fire, mountains, trees, are remarkable agencies, but their phenomena don't conform to the same laws as my actions do, so I don't attribute animal life to them. But the supporters of the volition theory ask us to infer that volition causes everything simply because it causes one particular thing; although that one thing. . . . is utterly special, its laws being enormously unlike those of any other phenomenon, organic or inorganic.

·SUPPLEMENTARY NOTE TO CHAPTER 5·

[In this densely learned four-page note, Mill responds to critics who accuse him of misrepresenting the views of Thales and Anaximenes, and of Descartes and Leibniz. (He points out that he didn't mention Descartes, only the Cartesians.) Mill's response opens with all cannons firing: 'A greater quantity of historical error has seldom been comprised in a single sentence.' Regarding the ancient philosophers he adduces more evidence, and also shows that his critics'

rival views about Thales and Anaximenes are based on misreadings of ancient texts, and ignorance of what Aristotle and others thought about who had had what theory.

[Mill side-tracks at some length into Aristotle's views about causation in the natural world, mainly so as to highlight two aspects of them. **(a)** Aristotle held that *chance* is an efficient cause (though not of everything). We now know that this was an error, Mill writes, but it wasn't a disreputable one:] Chance had as good a claim to real existence as many other of the mind's abstract creations; it had been given a name, and why should it not be a reality? **(b)** The parts of nature that Aristotle regards as representing evidence of *design* are the uniformities—the phenomena that conform to laws. The common interpretation of nature—we could call it the instinctive, religious interpretation of nature—is the reverse of this. The events in which men spontaneously see the hand of a supernatural being are the ones that they *can't* bring under physical laws. Events that they can clearly connect with physical causes, and especially ones they can predict, seem to them not to bear so obviously the mark of a divine will (though they may think that God is responsible for those too). . . . Some eminent writers on natural theology [see Glossary] . . . think that although design is present everywhere, the irresistible evidence of it is to be found not in the laws of nature but in the collocations [see Glossary], i.e. in the part of nature that shows no signs of any law. A few properties of dead matter might, they think, conceivably account for the regular and invariable succession of effects and causes; but they see proof of a divine providence in the way the different kinds of matter have been so placed as to promote beneficent ends. [Very roughly: It might be possible to explain how your body works without bringing in God, but we have to appeal to God to explain how there comes to be such a material configuration as your body.]

[Mill shows that he was certainly right in what he wrote about the views of Leibniz. We needn't spend time on this, except to note the tone of anger:] The critics say that what Leibniz found to be inconceivable was not •that mind acts on matter but •how it does so. This is an abuse of the privilege of writing confidently about authors without reading them! If my critics knew *anything* about Leibniz they would know that for him 'the inconceivability of *how*' and 'the impossibility of *the thing*' were equivalent expressions. . . .

[Regarding the Cartesians, Mill focuses on Malebranche, the best known Cartesian and the chief expositor (though not the inventor) of 'the system of occasional causes', and easily shows that he (Mill) was right about his views. He concludes: 'If Malebranche hadn't believed in an omnipotent Being, he would have held all action of mind on body to be a demonstrated impossibility.'

[There's a further half-page of tussle with the critics, but we can safely by-pass this efficient operation of garbage-removal.]

Chapter 6. The composition of causes

§1. To complete the general notion of causation. . . ., one distinction still remains to be pointed out. It is so radical and so important that it requires a chapter to itself.

We're now familiar with the case in which several agents or causes jointly produce an effect. It is indeed the usual case: very few effects are produced by just one agent. Suppose that two agents, operating jointly, are followed (under certain collateral conditions) by a given effect. If either of them had operated alone (under those same conditions), some effect would probably have followed, an effect •different from the joint effect of the two and •more or less dissimilar to it. When we know what would be the effect of each cause acting alone, we can often arrive deductively—i.e. *a priori*—at a correct prediction of what will arise from their joint agency. We can do this just so long as

the law expressing the effect of each cause acting by itself also correctly expresses that cause's part of the effect that follows from the two together.

That's how things stand in the important class of phenomena

commonly called 'mechanical', namely the phenomena of the communication of •motion from one body to another (or of •pressure, which is tendency to motion). In this class of cases it never happens that one cause defeats or frustrates another; both have their full effect. If a body is propelled by one force tending to drive it to the north and by another to the east, it is caused to move in a given time exactly as far in both directions as the two forces would separately have carried it. It ends up precisely where it would have arrived if it had been acted on first by one of the two forces and then by the other. In dynamics this law of nature is called the principle of the composition of forces; and in imitation of that well-chosen label I shall give the name 'composition of causes' to the principle that is exemplified whenever the joint effect of several causes is identical with the sum of their separate effects.

This principle doesn't prevail in all parts of nature. The chemical combination of two substances produces, as is well known, a third substance with properties different from

those of either of the two substances separately. . . . No trace of the properties of hydrogen or of oxygen is observable in their compound, water. The taste of lead acetate isn't the sum of the tastes of its component elements, acetic acid and lead or its oxide; nor is the colour of blue vitriol a mixture of the colours of sulphuric acid and of copper. This explains why mechanics is a deductive or demonstrative science, and chemistry is not. In mechanics we can compute the effects of combinations of causes from the laws that we know to govern those causes when acting separately, because they conform to the same laws when in combination that they conform to when acting separately. . . . Not so in the phenomena that are the special subject of the science of chemistry. There most of the uniformities the causes conform to when separate cease altogether when they are conjoined, and we can't (at least in the present state of our knowledge) foresee what result will follow from any new combination until we have tried the specific experiment.

If this is true of chemical combinations, it's even more true of the far more complex combinations of elements that constitute organised bodies—combinations out of which arise the extraordinary new uniformities called the 'laws of life'. All the parts of organised bodies are similar to the parts of inorganic things, and have themselves existed in an inorganic state; but the phenomena of life that result from the juxtaposition of those parts in a certain manner are utterly unlike all the effects that would be produced by the action of the component substances acting as mere physical agents. No imaginable knowledge of a living body's ingredients, however wide-ranging and complete, could enable us to predict the events of the living body itself from our knowledge of the separate actions of its elements. The

tongue, for instance, is composed of gelatine, fibrin, and other products of the chemistry of digestion; but from no knowledge of the properties of those substances could we ever predict that the tongue could taste, unless gelatine or fibrine could themselves taste, for no elementary fact can be in the conclusion that wasn't in the premises.

Thus the combined action of several causes can belong to either of two types, from which arise two ways in which laws of nature can conflict or interfere with one another. Take a case where at a given point of time and space there are two or more causes which, if they acted separately, would produce effects contrary to—or at least conflicting with—each other, one of them tending to undo some or all of what the other tends to do. Examples:

- The expansive force of the gases generated by the ignition of gunpowder tends to launch a bullet toward the sky, while its gravity tends to make it fall to the ground.
- A stream running into a reservoir at one end tends to raise its level higher and higher, while a drain at the other end tends to empty it.

In cases such as these, although the two causes exactly annul one another the laws of both are still fulfilled; the effect is the same as if the drain had been open for half an hour first, and the stream had flowed in for half an hour afterward.¹ Each agent produces the same amount of effect as if it had acted separately, though the contrary effect occurring at the same time obliterated it as fast as it was produced. Here, then, are two causes producing by their joint operations an effect that •at first seems quite unlike the effects they produce separately but •on examination proves to be really the sum of those separate effects. . . .

¹ Strictly speaking, in the second case the draining would be a little slower because there would be less pressure to create it, but that doesn't affect the truth of what I'm saying, because that would involve a change in the conditions under which the drain was acting.

So there's one kind of mutual interference of laws of nature in which, even when the joint causes annihilate each other's effects, each exerts its full efficacy according to its own law as a separate agent. In the other kind of case, the agencies that are brought together cease entirely, and a totally different set of phenomena arise—e.g. when two liquids are mixed in certain proportions they instantly become not a larger amount of liquid but a solid mass.

§2. This difference between. . . •laws that work together without alteration and •laws which, when called on to work together, cease and give place to others, is one of the fundamental distinctions in nature. The former case (the composition of causes) is the usual one; the other is always special and exceptional. There are no objects that don't obey the principle of the composition of causes with regard to *some* of their effects. For instance, a body retains its weight in all the combinations in which it is placed. The weight of a chemical compound, or of an organised body, is equal to the sum of the weights of the elements composing it. The weight will be lessened if the body is moved further from the centre of attraction, but it will be the same lessening for the compound as for the elements. The component parts of a plant or animal don't lose their mechanical and chemical properties as separate agents when they are spatially inter-related in the special way such that they, as an aggregate whole, acquire physiological or vital properties in addition. Those bodies still obey mechanical and chemical laws, because the operation of those laws isn't counteracted by the new laws that govern the bodies as organised beings. To put that in another way: when two or more causes jointly operate in a way that calls into action new laws with no resemblance to any we can find in the separate operation of the causes, the new laws, while superseding one portion of the previous laws,

may coexist with another portion, and may even compound the effect of those previous laws with their own.

Also, laws that were themselves generated in the second way may generate others in the first. The laws of chemistry and physiology (for example) owe their existence to a breach of the principle of composition of causes, but these *heteropathic* laws, as we might call them, are capable of composition with one another. The causes which by one combination had their laws altered may carry their new laws with them *unaltered* into further combinations. So we needn't despair of eventually raising chemistry and physiology to the condition of deductive sciences; for though it's impossible to deduce all chemical and physiological truths from the laws or properties of simple substances or elementary agents, perhaps they are deducible from laws that come into play when these elementary agents are brought together into some moderate number of not very complex combinations. The **laws of life** will never be deducible from the mere laws of the ingredients, but the prodigiously complex **facts of life** may all be deducible from comparatively simple laws of life—which do indeed depend on combinations, but comparatively simple ones. These laws of life may in more complex circumstances be strictly compounded with one another *and* with the physical and chemical laws of the ingredients. We already know enough about the vital phenomena to know of countless cases where they enter into the composition of causes; and the more precisely we study these phenomena the more reason we seem to get for believing that the laws that operate in the simpler combinations of circumstances do in fact continue to be observed in more complex ones. This will be found equally true in the phenomena of mind; and even in social and political phenomena, which are results of the laws of mind. It's with chemical phenomena that the least progress has been made, so far, in bringing the special

laws under general ones from which they can be deduced; but even in chemistry there are many circumstances to encourage the hope that such general laws will eventually be discovered. There's no chance that the different actions of a chemical compound will ever be found to be the sums of the actions of its separate elements; but between the properties of the compound and those of its elements there may be some constant relation that would enable us to foresee the sort of compound that will result from a new combination before we have actually tried it, and to judge what sort of elements some new substance is compounded of before we have analysed it. (The relation would of course have to be discovered by a sufficient induction.) The law of definite proportions, first presented in its full generality by [John] Dalton in his atomic theory, is a complete solution of this problem so far as quantity is concerned; and for quality we already have some partial generalisations suggesting that we may eventually get further. We can know in advance some properties of the kind of compounds that result from combining, in each of the small number of possible proportions, any acid with any base. We also have the curious law discovered by Berthollet: two soluble salts mutually decompose one another whenever the new combinations that result produce a compound that is less soluble than either of the original two. . . . Thus it appears that even heteropathic laws—laws of combined agency that aren't derived by simple addition from the laws of the separate agencies—are in some cases derived according to some fixed principle from the separate laws. So there may be laws governing the generation of laws from others that are unlike them; and in chemistry these undiscovered laws of the dependence of a compound's properties on the properties of its elements may, together with the laws of the elements themselves, provide the premises by which chemistry is perhaps destined

eventually to be made deductive.

So it seems that the composition of causes occurs in every class of phenomena; that as a general rule causes in combination produce exactly the same effects as when acting singly; but that this rule, though general, isn't universal because in some instances, at some particular points in the transition from separate to united action, the laws change and an entirely new set of effects occur in place of (or in addition to) the effects arising from the separate agency of those same causes; and that the laws of these new effects are in their turn capable of composition. . . .

§3. That *effects are proportional to their causes* is laid down by some writers as an axiom in the theory of causation. It has been worked hard in reasonings about the laws of nature, though it is burdened with many difficulties and apparent exceptions which much ingenuity has been expended in showing not to be real ones. What truth there is in this 'axiom' is just a special case of the composition of causes—the case where compounded causes are homogeneous, so that their joint effect might be expected to be the sum of their separate effects. The 'axiom' is illustrated by this:

A force equal to 100lb will raise a certain body a certain distance along an inclined plane; a force equal to 200lb will raise two such bodies the exact same distances; so the effect is proportional to the cause.

But the 200lb force contains two forces each equal to 100lb—forces each of which would raise one of the bodies if it were employed separately. So the fact used to illustrate the 'axiom' results from the composition of causes; it's a mere instance of the general fact that mechanical forces are subject to the law of composition. And it's the same in every other conceivable case. The doctrine of the proportionality of effects to their causes obviously can't apply to cases where

adding something to the cause alters the *kind* of effect. . . . Suppose that the application of a certain quantity of heat to a body merely •increases its size, that a double quantity •melts it, and a triple quantity •decomposes it: because these three effects are heterogeneous, there can't be *any* ratio between them, let alone one that matches the ratio among the quantities of heat applied. Thus the 'axiom' of the proportionality of effects to their causes fails at the precise point where the principle of the composition of causes fails. . . .

This is the end of my general remarks on causation, which I thought were needed as an introduction to the theory of the inductive process—a process that is essentially an inquiry into cases of causation. All the uniformities in the

•succession of phenomena, and most of the uniformities in their •coexistence, are either laws of causation or consequences of such laws. If we could determine what causes are correctly assigned to what effects, and what effects to what causes, that would virtually amount to knowing the whole course of nature. All the uniformities that are mere results of causation might then be explained, and every individual fact or event might be predicted, provided we knew the relevant facts about the circumstances that preceded it.

To ascertain, therefore, what laws of causation there are in nature—to determine the effect of every cause, and the causes of all effects—is the main business of induction. And the chief object of inductive logic is to point out how this is done.

Chapter 7. Observation and experiment

§1. One upshot of what I have been saying is that the process of ascertaining •what consequents are invariably connected with what antecedents, i.e. •what phenomena are related to each other as causes and effects, is a sort of process of analysis. We can take it as certain that every fact that begins to exist has a cause, and that this cause is some fact (or facts) that immediately preceded it. The totality of present facts is the infallible result of the totality of past facts, and more immediately of all the facts that existed a moment ago. So we have here a great sequence that we know to be uniform: if the whole moment-ago state of the entire universe could occur again, it would again be followed by the present state. How, then, are we to resolve [see Glossary] this complex uniformity into the simpler uniformities that compose it, and assign to each part of the vast antecedent

the part of the consequent that comes from it?

This operation, which I have called 'analytical' because it's the resolution of a complex whole into its component elements, is more than a merely mental analysis. We shan't get what we want merely by thinking about the phenomena, dividing them by the intellect alone. But such a mental partition is an indispensable first step. At first glance the order of nature looks at every moment like a chaos followed by another chaos! We must decompose each chaos into single facts. We must learn to see in the chaotic antecedent a multitude of distinct antecedents, in the chaotic consequent a multitude of distinct consequents. But this won't tell us which of the antecedents produces each consequent. To determine *that* we must try to separate the facts from one another, not only in our minds but in nature. The mental

analysis, however, must be done first. And we all know that intellects differ immensely in how they do this. It is of the essence of the act of *observing*; because the observer doesn't merely see the thing before his eyes but sees what parts it is composed of. The ability to do this well is rare:

- one person, from inattention or attending only in the wrong place, overlooks half of what he sees;
- another sets down much more than he sees, mixing it up with what he imagines or infers;
- a third takes note of the *kind* of all the circumstances, but because he's inexpert in estimating their *degree* he leaves the quantity of each vague and uncertain;
- a fourth sees the whole, but makes such an awkward division of it into parts—throwing into one mass things that should be separated, and separating others that would be better considered as one—that the result is no better, perhaps worse, than if he hadn't attempted any analysis.

We might discuss what qualities of mind and kinds of mental culture equip someone to be a good observer; but that belongs not logic but to the theory of *education* in the broadest sense of that term. There's no art of observing if 'art' is being used properly. There may be rules for observing. But these—like rules for *inventing*, are really instructions for how to put one's own mind into the state in which it will be most fitted to observe, or most likely to invent. So they are essentially rules of self-education, which is different from logic. They don't teach how to *do* the thing but how to *make ourselves capable of doing* it. They're an art of strengthening the limbs, not of using them.

How wide and how detailed does the observation have to be? How far down do we have to go in the mental analysis? That depends on the purpose in view. To ascertain the state of the whole universe at any moment is impossible, and

would also be useless. When making chemical experiments we don't think it necessary to note the position of the planets, because experience has shown. . . .that in such cases *that* detail isn't relevant to the result. Thus, at times when men believed in the occult influences of the heavenly bodies it might have been unphilosophical [here = 'unscientific'] to fail to check on the precise condition of those bodies at the moment of the experiment. As for the degree of minuteness of the mental subdivision, if we had to break down what we observe into •its very simplest elements. . . ., it would be hard to say where we would find •them; we can hardly ever affirm that our divisions of any kind have reached the ultimate unit. But fortunately this doesn't matter either. The point of the mental separation is to suggest the required physical separation, as something to be done by us or sought for in nature; and we needn't go beyond the point at which we can see what observations or experiments we require. What *does* matter is this: at whatever point our mental decomposition of facts has stopped, we should be ready and able to carry it further if there's a need for that, not allowing the freedom of our discriminating faculty [= 'our ability to make distinctions'] to be imprisoned by the straps and bindings of ordinary classification. That's what happened with all early speculative inquirers, the Greeks included. It seldom occurred to them that something called by •one abstract name might actually be •several phenomena, or that the facts of the universe might be decomposable into elements other than the ones already recognised in ordinary language.

§2. Suppose that we have ascertained the different antecedents and consequents and have discriminated them from one another as far as the case requires, we now face the question: Which is connected with which? There are

always many antecedents and many consequents. If the antecedents couldn't be separated from one another except in thought, or if the consequents were never found apart, it would be impossible for us to distinguish the real laws empirically, or to assign to any cause its effect, or to any effect its cause. To do that we have to encounter some of the antecedents apart from the rest, and observe what follows from them; or some of the consequents, and observe what they are preceded by. In short, we must follow the Baconian rule of **varying the circumstances**. This is indeed only the first rule of physical inquiry, and not the sole rule, as some have thought; but it is the foundation of all the rest.

If we want to vary the circumstances, we can rely on •observation or •experiment; we can either •find in nature an instance suited to our purposes, or •make one by an artificial arrangement of circumstances. The value of the instance depends on what it is in itself, not on how it is obtained; its role in induction depends on the same principles in each case, just as the uses of money are the same whether it is inherited or earned. So there's no difference in kind, no real logical distinction, between the two processes of investigation. But there are practical differences that it's important not to overlook.

§3. The most obvious difference is that experiment is an immense extension of observation. As well as enabling us to produce many more variations in the circumstances than nature spontaneously offers, it also (in thousands of cases) enables us to produce just exactly the sort of variation we need for discovering the law of the phenomenon we are studying. Nature is seldom so friendly as to *give* us that, because it's not constructed on a plan of helping us to study it!

For example, in order to ascertain what principle [see Glossary] in the atmosphere enables it to sustain life, we need a living animal to be immersed in each component element of the atmosphere separately. But **nature** doesn't supply either oxygen or nitrogen in a separate state. We are indebted to **artificial** experiments for our knowledge that it's oxygen, not nitrogen, that supports respiration; and for our knowledge of the very existence of those two ingredients.

Everyone realises that experimentation has the advantage over simple observation that it enables us to •obtain ever so many combinations of circumstances that aren't found in nature, and so •add to nature's experiments a multitude of experiments of our own. But many people *don't* realise that there's another superiority. . . .of artificially obtained instances over spontaneous ones—of our own experiments over even the same experiments when made by nature—which is at least as important.

When we produce a phenomenon artificially, we can (as it were) take it home with us, and observe it in circumstances [see Glossary] that we know about in detail. When we want to know what the effects are of the cause A, if we can produce A by means at our disposal we can generally determine at our own discretion. . . .the whole of the circumstances that are present along with it; and because this lets us know exactly the simultaneous state of everything else that could interact with A, we have only to observe what alteration is made in that state by the presence of A.

For example the electric machine lets us produce in thoroughly known circumstances the phenomena that nature displays on a grander scale in the form of lightning and thunder. Think about it: How much could mankind have learned about the effects and laws of electric agency from the mere observation of thunder-storms? And compare that with what they have learned and may expect to learn from

electrical and galvanic experiments! What makes this example especially striking is the fact—as we now have reason to believe—that electric action is of all natural phenomena (except heat) the most pervading and universal. This might lead you to think that electricity has the least need to be artificially produced in order to be studied; but the fact is the reverse of that—without the electric machine, the Leyden jar, and the voltaic battery we would probably never have suspected the existence of electricity as one of the great agents in nature; the few electric phenomena we would have known of would have gone on being regarded either as supernatural or as a sort of anomaly, an eccentricity in the order of the universe.

When we have insulated the phenomenon we're investigating by placing it among known circumstances, we can vary the circumstances in any way we like, choosing the variations that we think have the best chance of bringing the laws of the phenomenon into a clear light. By introducing one well-defined circumstance after another into the experiment, we discover how the phenomenon behaves in an indefinite variety of possible circumstances. Thus, when chemists have obtained some newly-discovered substance in a pure state. . . . they •introduce various other substances one by one, to discover whether it will combine with them or decompose them, and with what result; and also •apply heat, or electricity, or pressure, to discover what will happen to the substance in each of these circumstances.

But if we can't produce the phenomenon, and have to look for occurrences of it in nature, the task before us is very different.

Rather than *choosing* what the concomitant [see Glossary] circumstances shall be, we now have to *discover* what they are; and it's next to impossible to do this with any precision and completeness except in the simplest and most accessible

cases. Here's a phenomenon that we have no means of making artificially—a *human mind*. **Nature** produces many; but because we can't produce them by **art** we can see a human mind developing or acting on other things only when it is surrounded and obscured by an indefinite multitude of undiscoverable circumstances, making the use of ordinary experimental methods almost delusive. To get a sense of the scope of this difficulty, consider the fact that whenever nature produces a human mind she produces a body closely connected with it; i.e. a vast complex of physical facts, with no two of these complexes being exactly alike (probably), and most of them being radically out of the reach of our means of exploration (except for the mere structure, which we can examine in a coarse way after it has ceased to act [i.e. in an autopsy]). And if instead of a human mind we try to investigate a human society or a state, we encounter all the same difficulties—the same only worse.

We are now within sight of a conclusion that later chapters will (I think) make shingly evident: in the sciences dealing with phenomena in which •artificial experiments are impossible (such as astronomy), or in which •they have a very limited range (as in psychology, social science, and even physiology), induction from direct experience is practised at a disadvantage that usually amounts to impossibility. If those sciences are to learn anything worth learning, therefore, their methods must be largely and perhaps principally deductive. This is already known to be the case with astronomy; that it's not generally recognised as true of the others is probably one reason why they aren't in a more advanced state.

§4. Although pure ·hands-off· observation is at a great disadvantage compared with artificial experimentation in one branch of the direct exploration of phenomena, there's another branch where the advantage is all on the other side.

Because inductive inquiry aims to learn what causes are connected with what effects, we can begin this search at either end of the road: we can inquire either into •the effects of a given cause or into •the causes of a given effect. The fact that light blackens silver chloride could have been discovered either

- (a) by experiments on light, trying what effect it would have on various substances, or
- (b) by observing that portions of the chloride had repeatedly become black, and investigating the circumstances.

The effect of the poison curare could have become known either

- (a) by administering it to animals, or
- (b) by examining how it came about that the wounds the Indians of Guiana inflict with their arrows are always fatal.

A quick look at those examples, with no need for theoretical discussion, shows that artificial experimentation is possible only with the (a) procedures. We can take a cause and *try* what it will produce; we can't take an effect and *try* what it will be produced by. We can only watch till we see it produced, or are enabled to produce it by accident.

This wouldn't matter much if it was always up to us to choose which end to start from. But we seldom have any option. We can only travel from the known to the unknown, so we have to start at whichever end we know most about. If the agent is more familiar to us than its effects, we watch for or contrive instances of the agent in whatever varieties of circumstances we can manage, and observe the result. If the conditions on which a phenomenon depends are obscure but the phenomenon itself is familiar, we must start our inquiry from the effect. If we're struck with the fact that silver chloride has been blackened, and have no idea of the cause,

all we can do is to compare instances where the blackening has happened to occur, until through that comparison we discover that in all those instances the substances had been exposed to light. If we knew nothing of the Indian arrows but their fatal effect, accident alone could turn our attention to experiments on the poison; in the regular course of investigation we could only investigate or try to observe what had been done to the arrows in particular instances.

Whenever we have no leads on the cause and therefore have to start from the effect and apply the 'varying the circumstances' rule to the consequents, not the antecedents, we're deprived of the resource of artificial experimentation. But this is a matter of looking for or waiting for cases of the consequent in varying circumstances; we can't *produce* them because the only way to produce an effect is through its cause, and we don't *know* the cause. . . . If nature happens to present us with instances sufficiently varied in their circumstances, and if we can discover something that is always found—either immediately before the effect or some distance further back—when the effect is found and never found when it isn't, we can discover by mere observation and without experiment a real uniformity in nature.

But although this is certainly the most favourable case for •sciences of pure observation, as contrasted with •sciences in which artificial experiments are possible, there's really no case that more strikingly illustrates the inherent *imperfection* of direct induction when not based on experiment. Suppose that by comparing cases of the effect *y* we find an antecedent *x* that appears to be invariably connected with it; we haven't proved *x* to be the cause of *y* until we have reversed the process and used *x* to produce *y*. If we can produce *x* artificially, and if when we do so *y* follows, the induction is complete: we know that antecedent *x* is the

cause of that consequent y .¹ But we got there by adding the evidence of experiment to that of simple observation. [Mill then goes through it all again, with different words but the same content. He sums up:] In short, observation without experiment (and with no aid from deduction) can discover sequences and coexistences but can't prove causation.

[Mill cites zoology as a science in which an enormous amount is known about what follows what and what coexists with what, and yet:] on this vast subject. . . .we have made most scanty progress in discovering any laws of causation. In most of the cases of coexistence of animal phenomena

we don't know for sure which is the cause and which the effect (or whether they aren't related as cause and effect but rather are two effects of causes yet to be discovered, complex results of laws hitherto unknown.

Some of what I have said really belongs later, but I thought that a few general remarks on •how sciences of mere observation differ from sciences of experimentation, and on •the extreme disadvantage that inductive inquiry labours under in the former, would be the best preparation for discussing the methods of direct induction. . . ., a discussion to which I now proceed.

Chapter 8. The four methods of experimental inquiry

§1. The aim is to single out from among the circumstances that precede or follow a phenomenon the ones that it is really connected with by an invariable law. Two ways of doing this are simpler and more obvious than any other others. In the **Method of Agreement** we compare different instances in which the phenomenon occurs. In the **Method of Difference** we compare instances in which it occurs with instances in other respects similar in which it doesn't.

In illustrating these methods. . . ., I'll attend to their use both in •inquiring into the cause of a given effect and •inquiring into the effects or properties of a given cause. . . . I'll denote antecedents by upper-case letters and the corresponding consequents by •italicised• lower-case.

Let A be an agent or cause, and suppose we are trying to ascertain what its effects are. If we can find or produce A in such varieties of circumstances that the different cases have

no circumstance in common except A , then any effect that we find to be produced in *all* our trials is shown to be the effect of A . Suppose, for example, that A is tried along with B and C , and that the effect is $a b c$; and suppose that A is next tried with D and E , but without B and C , and that the effect is $a d e$. Then we may reason thus: b and c are not effects of A , for they weren't produced by it in the second experiment; nor are d and e , for they weren't produced in the first. Whatever is really the effect of A must have been produced in both instances, and the only circumstance that's true of is a

For example, let the A be the contact of an alkaline substance and an oil. We try this combination in several varieties of circumstances that resemble each other only in that they all produce a greasy and soap-like substance; so we conclude that the combination of an oil and an alkali causes the production of a soap. That is how we use the

¹ Unless y was generated not by the x but by the means used to produce the x . But these means are •under our power, so there's some probability that they are also sufficiently •within our knowledge to enable us to judge whether that could be the case.

Method of Agreement to inquire into the effect of a given cause.

In a similar way we can inquire into the cause of a given effect. Let *a* be the effect. Here. . . . we have only the resource of observation without experiment; we can't take a phenomenon of which we don't know the origin and try to find how it is produced by producing it! But if we can observe *a* in two different combinations, *a b c* and *a d e*; and if we know or can discover that the antecedent circumstances in these cases respectively were A B C and A D E, we can conclude by a reasoning similar to that in the 'soap' example that A is the antecedent connected with the consequent *a* by a law of causation. B and C can't be causes of *a* because on its second occurrence they weren't present; nor can D and E, because they weren't present on its first occurrence. A is the only one of the five circumstances that was found among the antecedents of *a* in both instances.

For example, suppose the effect whose cause we want to discover is crystallisation. We compare cases where bodies are known to acquire crystalline structure but have nothing else in common. We find them to have one—and as far as we can see *only* one—antecedent in common, namely the deposition of a solid matter from a liquid state. . . . So we conclude that the solidification of a substance from a liquid state is an invariable antecedent of its crystallisation.

In this example we can go further and say that this is not only •the invariable antecedent of crystallisation but •the cause of it; or at least the immediately preceding event that completes the cause. That's because after detecting the antecedent A we can •produce it artificially, and by finding that *a* follows it •verify the result of our induction. [Mill cites two examples, discoveries about how to produce quartz and how to produce marble. He comments that these are] two admirable examples of the light that can be thrown

upon the most secret processes of Nature by well-contrived interrogation of her.

But if we can't artificially produce A, the conclusion that it's the cause of *a* remains very doubtful. Even if it's an invariable antecedent of *a*, preceding it as day precedes night, it may not be *unconditionally* so.

This uncertainty arises from our inability to be sure that A is the only immediate antecedent common to both the instances. If we could be certain of having ascertained *all* the invariable antecedents, we might be sure that the unconditional invariable antecedent—i.e. the cause—must be among them. Unfortunately it's hardly ever possible to ascertain all the antecedents unless the phenomenon is one we can produce artificially. Even then, the difficulty is merely lightened, not removed; men knew how to raise water in pumps long before they learned what was really the operating circumstance in pumping, namely the pressure of the atmosphere on the open surface of the water. It's much easier to analyse completely •a set of arrangements made by ourselves than •the whole complex mass of agencies that nature happens to be exerting at the moment when a given phenomenon is produced. We may overlook some of the relevant circumstances in an experiment with an electrical machine; but at worst we'll be better acquainted with them than with the circumstances of a thunder-storm.

The way of discovering and proving laws of nature that I have just presented is based on the following axiom:

Whatever circumstances can be absent when the phenomenon is present is not causally connected with it. With such casual circumstances set aside, if only one remains then it is the cause we are searching for; if more than one remains, they either are the cause or contain it among them; and the same thing holds *mutatis mutandis* [see Glossary] for the effect.

As this method proceeds by comparing different instances to ascertain what they agree in, I call it the Method of Agreement, and we can adopt as its regulating principle the following:

FIRST CANON.

If two or more instances of the phenomenon under investigation have only one circumstance in common, that one circumstance is the cause (or effect) of the given phenomenon.

I'll return to the Method of Agreement very soon, but first I proceed to a still more powerful instrument in the investigation of nature, the Method of Difference.

§2. The Method of Agreement required instances that agreed in the given circumstance but differed in every other; the present method requires two instances that resemble one another in every other respect but differ in the presence or absence of the phenomenon we wish to study and, presumably in the presence or absence of the cause of that phenomenon. [That addition is needed to save Mill's account from incoherence. Other instances of the same trouble will be left untreated.] If we're trying to discover the effects of an agent A, we must procure A in some set of known circumstances A B C, note the effects of that, and compare them with the effect of the remaining circumstances B C without A. If the effect of A B C is $a b c$, and the effect of B C is $b c$, it is evident that the effect of A is a . And if we begin at the other end, wanting to investigate the cause of an effect a , we must select an instance $a b c$ in which the effect occurs and the antecedents were A B C, and then look for another instance in which $b c$ occur without a . If in that instance the antecedents are B C, we know that the cause of a must be A—either alone or in conjunction with some other circumstances present.

It's scarcely necessary to give examples of a logical process that gives us almost all the inductive conclusions we draw in daily life. When a man is shot through the heart, the Method of Difference shows us that it was the gunshot that killed him: he was in the fullness of life immediately before, all circumstances being the same as after except the wound.

The axioms implied in this method are evidently the following. An antecedent that can't be excluded without preventing the phenomenon is the cause of that phenomenon or a condition of it; a consequent that can be excluded with no other difference in the antecedents than the absence of a particular one x is the effect of x . Instead of comparing different instances of a phenomenon to see how they agree, this method compares an instance of its occurrence with an instance of its non-occurrence to see how they differ. The regulating principle of the Method of Difference may be expressed thus:

SECOND CANON.

If an instance where the phenomenon y under investigation occurs and an instance where it doesn't occur have every circumstance in common except for one x that occurs only in the former, x is the effect or the cause or an indispensable part of the cause of y .

§3. The two methods I have presented are alike in many ways but also unlike in many way. Both are methods of *elimination*. This term (borrowed from the mathematical theory of equations. . . .) is well suited to express the operation that has been understood since the time of Bacon to be the foundation of experimental inquiry—namely the successive exclusion of the various circumstances that are found to accompany a phenomenon in a given instance, in order to

ascertain which of them can be absent consistently with the existence of the phenomenon. The Method of Agreement is based on the thesis that whatever can be eliminated is not connected with the phenomenon by any law. The Method of Difference has for its foundation that whatever *can't* be eliminated is connected with the phenomenon by a law.

Of these two, the Method of Difference is more particularly a method of artificial experiment; while the Method of Agreement is more especially what we use when experimentation is impossible. A few reflections will prove this, and point out the reason of it.

It is inherent in the unique character of the Method of Difference that the nature of the combinations it requires is much more strictly defined than in the Method of Agreement. The two instances that are to be compared must be exactly similar in all circumstances except the one we're trying to investigate; they must inter-relate as $A B C$ relates to $B C$ (·if we're investigating the effects of A ·) or as $a b c$ relates to $b c$ (·if we're investigating the cause of a ·). This similarity of circumstances needn't be total—it needn't extend to circumstances that we already know to be irrelevant to the result. With most phenomena we learn at once, from the commonest experience, that most of the coexistent phenomena in the universe can be either present or absent without affecting the given phenomenon. . . . Still, even limiting the identity that's required between the two instances $A B C$ and $B C$ to circumstances that aren't already known to be irrelevant, nature very seldom offers two instances that we can be sure are related in that way—i.e. that the *only* difference between them (apart from ones that we know are irrelevant) is the presence of A in one of them and not the other. Nature's spontaneous operations are generally so complicated and so obscure—being out of our reach because they are too vast or too tiny—that we're ignorant of a great part of the facts

that really take place, and even the ones we aren't ignorant of are so numerous and thus so seldom exactly alike in any two cases that a spontaneous experiment [= 'a hands-off observation'] of the kind required by the Method of Difference is usually not to be found. On the other hand, when we obtain a phenomenon by an artificial experiment, a pair of instances such as the method requires is obtained almost as a matter of course, provided the process doesn't last a long time. A certain state of surrounding circumstances existed before we started the experiment; this is $B C$. We then introduce A —e.g. by merely bringing an object from another part of the room—before there has been time for any change in the other elements. Comte was right: it's the very nature of an experiment to introduce into the pre-existing state of circumstances a perfectly definite change. We *choose* a previous state of things that we are well acquainted with, so that it's not likely to change without our noticing; and into this we introduce, as rapidly as possible, the phenomenon x that we want to study; so that in general we're entitled to be quite sure that the state we have produced differs from the pre-existing state only in the presence or the absence of x . If a bird is taken from a cage and instantly plunged into carbonic acid gas, the experimenter can be fully assured (after one or two repetitions) that no circumstance that could cause suffocation had intruded except the change from •immersion in the atmosphere to •immersion in carbonic acid gas. . . . It thus appears that in the study of the various kinds of phenomena that *we* can modify or control, we can in general satisfy the requirements of the Method of Difference; but that those requirements are seldom fulfilled by the spontaneous operations of nature.

With the Method of Agreement the situation is reversed. We don't here require instances of such a special and determinate kind. For the purposes of this method, *any* instances

in which nature presents us with a phenomenon can be examined, and if all such instances agree in *anything*, that's already a useful conclusion. It's true that we usually can't be sure that the one point of agreement is the only one; but this ignorance does not invalidate the conclusion, as it would with the Method of Difference. . . . We have ascertained one invariable antecedent or consequent, however many other invariable antecedents or consequents may still remain unascertained. If A B C and A D E and A F G are all equally followed by *a* then *a* is an invariable consequent of A. If *a b c* and *a d e* and *a f g* all have A among their antecedents, then A is connected as an antecedent with *a*. But to determine whether this invariable antecedent is a cause, or this invariable consequent an effect, we must also be able to produce one of them by means of the other; or at least to obtain an instance in which the effect *a* has come into existence with no change in the circumstances except the addition of A. (That is our only way of being sure that we have *produced* something.) And this, if we can do it, is an application of the Method of Difference, not of the Method of Agreement.

So it seems that the only way direct experience can give us certain results about causes is through the Method of Difference. The Method of Agreement leads only to uniformities which either aren't laws of causation or whose status as causal must for the present remain undecided. (Some writers call these 'laws of phenomena', but that's a bad usage because laws of causation are also laws of phenomena.)

The Method of Agreement is to be used mainly •as a means of suggesting applications of the Method of Difference (as in the last example, where the comparison of A B C and A D E and A F G suggested that A was the antecedent on which to try by experiment whether it could produce *a*); or •as a second-best in cases where the Method of Difference is

impracticable—e.g. because we can't artificially produce the phenomena. So the Method of Agreement—though applicable in theory to either case—is more emphatically the method of investigation in cases where artificial experimentation is impossible, because in them it's usually our only resource of a directly inductive kind, whereas with phenomena that we can produce at will the Method of Difference is generally more effective because it can ascertain •causes as well as •mere laws.

§4. But in many cases our power of producing the phenomenon is complete and yet the Method of Difference can't be used at all, or only with a previous use of the Method of Agreement. This occurs when our only way of producing the phenomenon involves a combination of antecedents that we can't separate from each other and exhibit apart. Suppose, for instance, that we want to investigate the cause of the double refraction of light. We can produce this phenomenon at will, using any one of the many substances that we know to refract light in that special manner—Iceland spar, for example—but we can't use the Method of Difference because we can't find another substance precisely resembling Iceland spar except in some one property. The only way to push this inquiry is the one provided by the Method of Agreement. And that's what *was* used: the physicists compared all the known substances that doubly refract light, and found that they have in common *being crystalline substances*; from which they reasonably inferred. . . . that either •crystalline structure or •the cause of that structure is one of the conditions of double refraction.

[This paragraph will have a good many small omissions not indicated by. . . ellipses. The reasons are purely aesthetic; you can trust the paragraph's content.] Suppose that by using the Method of Agreement we have discovered that there's a connection

between A and *a*. To convert this evidence of *connection* into proof of *causation* by the direct Method of Difference we would need to do things like:

having tested A B C and found that it leads to *a*, we then test B C and observe whether that leads to *a* also.

Now, we often can't do this (see the Iceland spar example), but sometimes we can find out what *would* be the upshot if we could test B C, and that's as good as conducting the test. Here's how we do that:

Having tested a variety of cases where *a* occurred, and found that they all contain A, we now observe a variety of instances where *a* doesn't occur, and find that none of them contains A.

This establishes by the Method of Agreement the same connection between the absence of A and the absence of *a*, which was previously established between their presence. Just as our first work showed that whenever A is present *a* is present, so now we can conclude whenever A is absent *a* is also absent, which means that we have the positive and negative instances that the Method of Difference requires.

This method—call it the 'Indirect Method of Difference' or the 'Joint Method of Agreement and Difference'—consists in two uses of the Method of Agreement, each independent of the other and corroborating it. But it isn't equivalent to a proof by the direct Method of Difference. The Method of Difference requires us to be quite sure •that the instances leading to *a* have nothing in common except A, or •that the instances that don't lead to *a* have nothing in common but the absence of A. This is never possible; and if it were, we wouldn't need the joint method, because either of the two sets of instances separately would prove causation. This indirect method, therefore, can only be regarded as an extension and improvement of the Method of Agreement, but not as

having any part in the more powerful nature of the Method of Difference. Its canon is this:

THIRD CANON.

If two or more instances in which the phenomenon occurs have only one circumstance x in common, while two or more instances in which it doesn't occur have nothing in common except the absence of x, then x is the effect, or the cause or an indispensable part of the cause, of the phenomenon.

[Mill says that the Joint Method of Agreement and Difference has another advantage over 'the common Method of Agreement', but that he needs to postpone discussing this until later, and will] at once proceed to a statement of the other two methods, which will complete the list of the means we have for exploring the laws of nature by specific observation and experience.

§5. The first of these has been well named 'the Method of Residues'. Its principle is very simple. Remove from any given phenomenon all the parts of it that can by virtue of preceding inductions be assigned to known causes, and what's left will be the effect of antecedents which had been overlooked or whose effect was still an unknown quantity.

Suppose again that we have the antecedents A B C followed by the consequents *a b c*, and that by previous inductions (based, let's say, on the Method of Difference) we have discovered the causes of some of these effects or the effects of some of these causes; specifically we have learned that the effect of A is *a*, and that the effect of B is *b*. Subtracting the sum of these effects from the total phenomenon, there remains *c*, and we don't need any new experiments to know that *c* is the effect of C. This Method of Residues is in fact a special adaptation of the Method of Difference. If the instance A B C and *a b c* could have been compared with

a single instance $A B$ and $a b$, we would have proved C to be the cause of c by using the Method of Difference in the ordinary way. In the present case, though, instead of a single instance $A B$ we have had to study separately the causes A and B , and to infer from the effects they produce separately what effect they must produce in the case $A B C$, where they act together. Thus, of the two instances that the Method of Difference requires—one positive, the other negative—the negative one (in which the given phenomenon is absent) is not the direct result of observation and experiment, but has been arrived at by deduction. As one of the forms of the Method of Difference, the Method of Residues shares in its rigorous certainty, provided the previous inductions—the ones that gave the effects of A and B —were obtained by the same infallible method, and provided we're certain that C is the only antecedent that the residual phenomenon c can be connected with, i.e. the only agent whose effect we hadn't already calculated and subtracted. But we can never be quite certain of this, so the evidence derived from the Method of Residues is not complete unless we can obtain C artificially and test it separately, or unless its agency, when once suggested, can be explained and proved deductively from known laws.

Even with these reservations, the Method of Residues is one of our most important instruments of discovery. Of all the methods of investigating laws of nature, this is the most fertile in unexpected results, often informing us of sequences in which neither the cause nor the effect was conspicuous enough to attract the attention of observers. The agent C may be an obscure circumstance, not likely to have been perceived unless sought for, nor likely to have been sought for until attention had been awakened by the insufficiency of the obvious causes to account for the whole of the effect. And c may be so disguised by its intermixture with a and b that

it would scarcely have presented itself spontaneously as a subject of separate study. I'll soon present some remarkable examples of these uses of the Method of Residues. Its canon is as follows:

FOURTH CANON.

Subtract from any phenomenon the part of it that is known by previous inductions to be the effect of certain antecedents, and the remainder of the phenomenon is the effect of the remaining antecedents.

§6. There remains a class of laws that can't be discovered by any of the three methods I have tried to describe, namely the laws of •permanent causes—i.e. •indestructible natural agents—that we can't exclude or isolate, can't hinder from being present or arrange to have present alone. You might think that we can't possibly separate the effects of these agents from the effects of the other agents that they have to coexist with; but in fact for most of the permanent causes no such difficulty arises: although we can't eliminate them as •coexisting facts, we can eliminate them as •influencing agents by simply conducting our experiment in a place outside the reach of their influence. The swing of a pendulum, for example, is disturbed by a nearby mountain; we move the pendulum far enough away from the mountain, and the disturbance ceases. From these data [see Glossary] we can use the Method of Difference to calculate the amount of effect due to the mountain; and beyond a certain distance everything goes on precisely as it would do if the mountain exercised no influence whatever, and we reasonably enough conclude that it doesn't.

But the picture changes when we can't get ourselves ·or our experimental apparatus· out of reach of the influence of a permanent cause. The pendulum can be moved away

from the influence of •the mountain, but it can't be removed from the influence of •the earth; we can't move the earth and the pendulum away from one another, to discover whether it would continue to swing if the earth's action on it were withdrawn. Then what is our evidence that the pendulum's swing is caused by the earth's influence? It can't be anything supported by the Method of Difference, for one of the two instances is lacking—namely the negative instance where the earth's influence isn't a factor. Nor by the Method of Agreement: when any pendulum swings the earth is always present, but so is the sun! Obviously to establish even such a simple fact of causation as this we needed some method other than those I have so far presented.

For another example, consider *heat*. Independently of any theory about the real nature of heat we can be sure of this much: •we can't deprive any body of the whole of its heat, and •no-one ever perceived heat that wasn't being given off by a body. So we can't separate body and heat, and therefore can't vary the circumstances in the way the foregoing three methods require—we can't ascertain by those methods what portion of the phenomena exhibited by any body is due to the heat contained in it. If we could observe a body with its heat, and the same body entirely divested of heat, the Method of Difference would show the effect of the heat, apart from the effect of the body. If we could observe heat under circumstances agreeing only in heat, and therefore not involving the presence of a body, we could use the Method of Agreement to discover the effects of heat by comparing •an instance of heat with a body and •an instance of heat without a body; or we could use the Method of Difference to discover what effect was due to the body, the remainder due to heat being given by the Method of Residues. But we can't do any of these things, so none of the three methods can help us to solve this problem. . . .

. . . .But there is still something we can try. Even when we can't exclude an antecedent altogether, we may be able to produce—or nature may produce for us—some modification in it, by which I mean a change in it not amounting to its total removal. If some modification in the antecedent *A* is always followed by a change in the consequent *a*, the other consequents *b* and *c* remaining the same; or *vice versa* if every change in *a* is found to have been preceded by some modification in *A*, none being observable in any of the other antecedents, we can safely conclude that *a* is at least in part causally connected somehow with *A*. We can't expel heat altogether from any body, but we can modify its amount, increasing or diminishing it; and in doing this we can find by the various methods of experiment or observation that I have discussed that such increase or diminution of heat is followed by expansion or contraction of the body. This brings us to the conclusion that we couldn't have achieved in any other way, that one effect of heat is to make bodies bigger, i.e. to increase the distances between their particles.

A change in a thing that doesn't amount to its total removal—i.e. a change that leaves it still the same thing—must be a change either in •its quantity or in •some of its variable relations to other things; and the main one of these is position in space. We have seen an example depending on quantity; now for one involving spatial position. Question: what influence does the moon exert on the surface of the earth? We can't try an experiment in the absence of the moon. But when we find that all the variations in the moon's position are followed by corresponding variations in the time and place of high tide, the place always being either the part of the earth nearest to the moon or the part furthest from it, this gives us ample evidence that the moon is at least partially the cause that determines the tides. . . .

Similar evidence shows that the swinging of a pendulum is caused by the earth. The swings take place between equidistant points on opposite sides of a line that •is perpendicular to the earth, and therefore •varies with every variation in the earth's position. . . . This method tells us that all terrestrial bodies tend toward the earth, and not towards some unknown fixed point lying in the same direction. In every 24 hours of the earth's rotation, the line drawn from the body at right angles to the earth coincides successively with all the radii of a circle, and in the course of six months the place of that circle changes by nearly 200,000,000 million miles; yet in all these changes of the earth's position the line in which bodies tend to fall—the line down the centre of the pendulum's swing—continues to be directed toward it. This proves that terrestrial gravity is directed towards the earth and not, as some people used to think, towards a fixed point in space.

The method by which these results were obtained may be termed the Method of Concomitant [see Glossary] Variations; it is regulated by the following canon:

FIFTH CANON.

If any phenomenon x varies in some specific way whenever another phenomenon y varies in some specific way, x is either a cause or an effect of y, or is causally connected with it in some other manner.

I add that last clause because when two phenomena match each other in their variations it doesn't follow one is cause and the other effect. If they were two effects of a common cause, they would exhibit concomitant variation; and this method alone can't tell us whether they're related in that way rather than as cause and effect. The only way to answer the question would be—yet again!—by trying to ascertain whether we can produce the one set of variations

by means of the other. In the case of heat, for example, by increasing the temperature of a body we make it bigger, but by making it bigger (e.g. by using an air-pump to decrease the air-pressure on it) we don't increase its temperature; on the contrary, in most cases we diminish it. So heat is not an effect of increase in size but a cause of it. If we can't ourselves produce the variations, we must try—though we'll usually fail—to find them produced by nature in some case in which the pre-existing circumstances are perfectly known to us. . . .

You might think that the Method of Concomitant Variations assumes a new axiom, i.e. a new law of causation in general, namely: *Every modification of the cause is followed by a change in the effect.* And it does usually happen that when a phenomenon A causes a phenomenon a, any variation in A's quantity or relational properties is uniformly followed by a variation in the quantity or relational properties of a. . . . The sun causes a certain tendency to motion in the earth; here we have cause and effect; but that tendency is toward the sun, and therefore varies in •direction as the sun varies in its position relative to the earth; and the tendency also varies in •intensity in a certain numerical correspondence to the sun's distance from the earth—i.e. according to another relation of the sun. So there's not only an invariable connection between the sun and the earth's gravitation, but two of the sun's relational properties—its position relative to the earth and its distance from the earth—are invariably connected as antecedents with the quantity and direction of the earth's gravitation. The cause of the earth's gravitating at all is simply *the sun*; but the cause of its gravitating with a given intensity in a given direction is the existence of the sun at a given distance and in a given direction from the earth. A modified cause is really a different cause, so it's not surprising that it produces a different effect.

But the Method of Concomitant Variations doesn't require as an axiom that

(a) If x is the cause of y , any modification of x is followed by a modification of y .

All it needs is the converse proposition:

(b) If every modification of x is followed by a modification of y , x is the cause of y (or is connected with the cause of y).

It's obvious that (b) is true, because if x has no influence over y then modifications of x can't influence y either. If the stars have no power over the fortunes of mankind, then the conjunctions or oppositions of stars can have no such power.

The most striking uses of the Method of Concomitant Variations occur in cases where the Method of Difference, strictly so-called, is impossible; but its use isn't confined to those cases. It is often useful as a follow-up to the Method of Difference, to give additional precision to a solution that the latter method has found. When we know through the Method of Difference that x produces y , the Method of Concomitant Variations can be usefully called in to determine what law governs the match between x 's quantity and relational properties and y 's.

§7. This method is most widely used in cases where the ·relevant· variations of the cause are variations of *quantity*. It's pretty safe to say that quantitative variations in the cause will be attended by *quantitative* variations in the effect; because the proposition that •more of the cause is followed by •more of the effect follows from the principle of the Composition of Causes, which we saw on page 182 to be the •general rule of causation, whereas counterexamples to it—cases where causes change their properties on being combined—are •special and exceptional. Suppose that when A changes in quantity, a also changes in quantity, and that

we can trace the •numerical relation between parts of the two sets of changes—the parts, that is, that aren't too big or too small for us to observe them. Then with certain precautions we can safely conclude that the same •numerical relation will hold outside those limits. If we find that when A is double, a is double, when A is treble or quadruple, a is treble or quadruple, we can conclude

- (i) that if A were a half or a third, a would be a half or a third, and
- (ii) that if A were annihilated, a would be annihilated; and thus
- (iii) that a is wholly the effect of A or wholly the effect of A 's cause.

And we could infer (iii) for any other numerical relation according to which (ii) A and a would vanish simultaneously—e.g. if a were proportional to A^2 . If on the other hand a is not wholly the effect of A , but still varies when A varies, it is probably a mathematical function not of A alone but of A and something else. For example, its changes may be what you would get if some part of it remained constant or varied on some other principle, while the remainder varied in some numerical relation to the variations of A . In that case, as A diminishes, a will be seen to approach not •zero but •some other limit; and when the series of variations indicates what that limit is, the limit will exactly measure how much of a is the effect of some other and independent cause, and the remainder will be the effect of A (or of the cause of A). That is stated for cases where the limit is constant; if it is variable, replace 'indicates what that limit is' by 'indicates what the law of its variation is'.

But these conclusions mustn't be drawn without certain precautions. In the first place, they can't be drawn at all unless we're acquainted not only with •the variations but with •the absolute quantities both of A and a . If we don't

know the total quantities, we can't determine the numerical relation according to which they vary. So it's an error to conclude (as some have concluded) that because increase of heat expands bodies, i.e. increases the distance between their particles, therefore •that distance is wholly the effect of heat, and •if we could entirely deprive the body of its heat the particles would be in complete contact. This is a mere guess, and wildly risky one rather than a legitimate induction. Because we don't know how much heat there is in any body, or what the real distance is between any two of its particles, we can't judge whether the contraction of the distance follows the diminution of the quantity of heat according to a numerical relation such that the two quantities would reach zero simultaneously.

Now consider a case where the absolute quantities *are* known, namely the case addressed in the first law of motion, which says that *all bodies in motion continue to move in a straight line with uniform velocity until acted upon by some new force*. This is in open opposition to first appearances; all moving terrestrial objects slow down and eventually stop; and the ancients—going by *inductio per enumerationem simplicem* [see Glossary]—imagined that to be the law. But every moving body encounters various obstacles—friction, the resistance of the atmosphere, etc.—which we know by daily experience to be causes that can destroy motion. It was suggested that the lessening of motion might come wholly from these causes. How was this inquired into? With the obstacles entirely removed, the Method of Difference could have come into play. But they couldn't be removed, only lessened, so the case had to be handled by the Method of Concomitant Variations. This was used, and it was found that every lessening of the obstacles lessened the slowing of the motion; and this being a case (unlike the case of heat) where the total quantities of both the antecedent and of

the consequent were known, it was possible to get a fairly accurate estimate of the amount of •the slowing and the amount of •the relevant resistances, and to judge how near each was to zero; and it turned out that the effect dwindled as rapidly as the cause did, so that at each step the two were equally near to annihilation. The swinging of a weight suspended from a fixed point and moved a little out of the perpendicular ordinarily lasts for only a few minutes, but Borda got it to continue for more than thirty hours by going as far as possible towards reducing the friction at the point of suspension and making the body move in a vacuum. That left no reason to hesitate to conclude that the whole of the slowing of motion was due to the influence of the obstacles. With the slowing removed from the total phenomenon, the remainder was a uniform velocity, and the result was the proposition known as the first law of motion.

The inference that *the law of variation that the quantities conform to within our limits of observation will hold beyond those limits* is open to another kind of uncertainty. Actually there are two of them, one being obvious: we don't know what happens in the range outside the limits of our observation, and it *might* be that something comes into play there that spoils our conclusion. This kind of uncertainty comes into virtually all our predictions of effects; it's not specially relevant to the Method of Concomitant Variations. I want to talk about an uncertainty that is characteristic of that method; especially in the cases where our *observable* range is very small compared with the *possible* variations in the quantities of the phenomena. If you know anything of mathematics you know that very different laws of variation can produce numerical results that differ only slightly from one another; and in many cases it's only when the absolute amounts of variation are considerable that we can see the difference between the results given by two rival laws. The upshot is

that when the variations in the quantity of the antecedents that we can observe are small in comparison with the total quantities, there's a great danger of our picking the wrong numerical law, and being led to miscalculate the variations that would occur beyond our limits. That miscalculation would invalidate any conclusion about the dependence of the effect on the cause. There are plenty of examples of such mistakes. Herschel writes: 'The formulae that have been empirically deduced for the elasticity of steam (till very recently), and those for the resistance of fluids and other similar subjects', when relied on beyond the limits of the observations from which they were deduced, 'have almost invariably failed to support the theoretical structures based on them'.

Even when we have this uncertainty, the Method of Concomitant Variation can prove that there is some connection between A and *a*, and . . . can legitimately satisfy us that the

relation we have observed (within our limits) to exist between the variations of A and *a* will hold true in all cases that fall between those same limits. . . .

The four methods that I have tried to describe are the only possible modes of experimental inquiry—of direct induction *a posteriori* as distinguished from deduction. At any rate, I don't know of any others and can't imagine any others. And the Method of Residues (I remind you) isn't independent of deduction; but I include it among methods of direct observation and experiment because as well as deduction it also requires specific experience.

. . . .In chapter 10 I'll come to certain circumstances that make the use of these methods much more complicated and difficult than I have so far indicated. Before coming to that, though, I shall illustrate the use of the methods by suitable examples drawn from actual physical investigations.

Chapter 9. Examples of the four methods

§1. First example: I'll start with an interesting bit of theory by one of the most eminent theoretical chemists, Baron Liebig. The objective is to discover the immediate cause of the death produced by metallic poisons.

Arsenious acid and the salts of lead, bismuth, copper, and mercury, if introduced into the animal organism in anything but the smallest doses, destroy life. These facts have long been known, as separate and unconnected truths that are as ungeneral as generalisations can be. It was left to Liebig, by an apt employment of the Methods of Agreement and Difference, to connect these truths with one another by a higher induction, revealing the property that •is common to

all these harmful substances and •is the operative cause of their fatal effect.

(a) When solutions of these substances are placed in close enough contact with many animal products—albumen, milk, muscular fibre, and animal membranes—the acid or salt leaves the water it was dissolved in and enters into combination with the animal substance; and this substance, after being thus acted upon, is found to have lost its tendency to putrefy [see Glossary].

(b) Observation also shows, in cases where death has been produced by these poisons, that the parts of the body that the poisonous substances have been brought into contact

with don't afterwards putrefy.

(c) And, finally, when the poison has been supplied in too small a quantity to destroy life. . . ., certain superficial portions of the tissues are destroyed and afterwards thrown off by the process of recovery in the healthy parts.

These three sets of instances can be handled according to the Method of Agreement. In all of them the metallic compounds are brought into contact with the substances that compose the human or animal body; and the instances seem to have nothing else in common. The remaining antecedents are as different—even opposite—as they could possibly be made; for in some the animal substances exposed to the action of the poisons are in a state of life, in others only in a state of organisation, in others not even in that. And the result in *all* the cases is the conversion of the animal substance (by combination with the poison) into a chemical compound that is held together by force so powerful that it resists the subsequent action of the ordinary causes of decomposition. Now, organic life (the necessary condition of sensitive life) consists in a continual state of decomposition and recomposition of the different organs and tissues, so anything that prevents this decomposition destroys life. Thus the immediate cause of the death produced by poisons of this kind is ascertained, as far as the Method of Agreement can ascertain it.

Let us now use the Method of Difference to test our conclusion. This will involve a comparison. On one hand we have:

cases where the antecedent is the presence of substances that combine with the tissues to form a compound that can't putrefy (and therefore can't support life), the consequent being death of the whole organism or of some part of it.

We are to compare these with

cases as much like the former ones as possible except that they don't have the death of anything as their effect.

Many insoluble basic salts of arsenious acid are known not to be poisonous. The substance called 'alkargen', discovered by Bunsen, which contains a great amount of arsenic and is very like the organic arsenious compounds found in the body, hasn't the slightest injurious action upon the organism. Now when these substances are brought into contact with the tissues in any way, they don't combine with them, and don't stop their progress towards decomposition. What these instances seem to show is that when the effect is absent that's because of the absence of the antecedent that we already had good reason to consider as the immediate cause.

But the rigorous conditions of the Method of Difference aren't yet satisfied; for we can't be sure that these unpoisonous bodies differ from the poisonous substances *only* in not combining with animal tissues to form a compound that resists decomposition. To make the method strictly applicable, we need an instance not of a different substance but of one of the very same substances, in circumstances that prevent it from combining with the tissues to form the sort of compound in question; and *then*, if death doesn't follow, our case is made out. Instances of this kind are provided by the antidotes to these poisons. For example, if hydrated peroxide of iron is administered along with poisonous arsenious acid, the destructive agency of the latter is instantly checked. Now, this peroxide is known to combine with the acid to form a compound that is insoluble, and so can't act at all on animal tissues. Thus, sugar is a well-known antidote to poisoning by salts of copper; and sugar turns those salts into something that doesn't combine with animal matter. The disease called 'painter's colic', so common in factories making white lead, is unknown where

the workmen regularly take (as a preservative) a solution of sugar made acid by sulphuric acid. Now, diluted sulphuric acid has the property of •decomposing all compounds of lead with organic matter or •preventing them from being formed.

[Mill then describes a set of facts about 'soluble salts of silver' which, when applied externally, have about the same effect as arsenious acid, but aren't poisonous when ingested. The explanation is that the animal stomach contains common salt and muriatic acid, which turn the soluble salts into something virtually insoluble and therefore unable to combine with the tissues to fatal effect,]

Those instances have shown us a very conclusive induction that illustrates the two simplest of our four methods; though it doesn't rise to the maximum of certainty that a *perfect* example of the Method of Difference can provide. Remember that the positive instance and the negative one strictly ought to differ only in the presence or absence of one single circumstance. And in the foregoing argument they differ in the presence or absence not of a single •circumstance but of a single •substance; every substance has countless properties; so there's no knowing how many real differences are involved in what is apparently only one difference. It is conceivable that the antidote. . . counteracts •the poison through some property other than that of forming an insoluble compound with •it; and if that were so the theory would collapse so far as it rests on that instance. This source of uncertainty is a serious hindrance to all extensive generalisations in chemistry; but in our present case it is reduced to almost the lowest possible degree when we find that *many* substances can act as antidotes to metallic poisons, and that all these share the property of forming insoluble compounds with the poisons and can't be ascertained to share any other

property whatsoever. So we have in favour of the theory all the evidence that can be obtained by the Joint Method of Agreement and Difference [see page 196]; and though the evidence it produces can't amount to that of the Method of Difference properly so-called, it can approach indefinitely near to that.

§2. Second example: The aim is to discover the law governing 'induced electricity'—i.e. to learn under what conditions a body that is positively or negatively electrified gives rise to the opposite electric state in some other body adjacent to it.¹

The most familiar kind of example of the phenomenon to be investigated is the following. Around the prime conductors of an electrical machine the nearby atmosphere or any conducting surface suspended in it is found to be in the electric condition opposite to that of the prime conductor itself: near and around the positive prime conductor there's negative electricity, and near and around the negative prime conductor there's positive electricity. When a pith ball (or a human hand) is brought near to one of the conductors, it becomes electrified with the opposite electricity to it—either •receiving a share from the already electrified atmosphere by conduction, or •acted on by the direct inductive influence of the conductor itself—and then it is attracted by the conductor to which it is opposite or by any other oppositely charged body. Now, we have no evidence that a charged conductor can be suddenly discharged except by the approach of a body with the opposite charge. In the case of the electric machine, therefore, it appears that the accumulation of electricity in an insulated conductor is always accompanied by the excitement of the opposite electricity in •the surrounding atmosphere and in things in •it. It does not seem possible,

¹ For this bit of theorising as for many of my other scientific illustrations I am indebted to Bain, whose treatise on Logic is full of apt illustrations of all the inductive methods.

in this case, to produce one electricity by itself. [That last sentence is verbatim from Mill.]

Let us now examine all the other instances we can get that resemble this one in the given consequent, namely *the occurrence of an opposite electricity in the neighbourhood of an electrified body*. One remarkable instance is the Leyden jar; another is the magnet, in which it is impossible to produce one kind of electricity by itself, i.e. to charge one pole without charging another pole with the opposite electricity at the same time. (That holds both for natural magnets and for electromagnets. In counting magnets as relevant to my topic, I am relying on Faraday's splendid experiments decisively showing that magnetism and electricity are basically the same thing.) We can't have a magnet with one pole; if we break a natural lodestone into a thousand pieces, each piece will have its two oppositely electrified poles complete within itself. In the voltaic circuit, again, we can't have one current without its opposite. In the ordinary electric machine, the glass cylinder or plate acquires an electrical charge opposite to that of the rubber.

From all these instances, treated by the Method of Agreement, a general law appears to result. The instances cover all the known ways in which a body can get an electric charge; and in all of them there is found, as a concomitant or consequent, the excitement of the opposite electric charge in some other body or bodies. It seems to follow that the two facts are invariably connected, and that a necessary condition of a body's acquiring an electric charge is the simultaneous excitement of the opposite charge in some neighbouring body.

As the two opposite charges can only be produced together, so they can only cease together. This can be shown by an application of the Method of Difference to the Leyden jar. In the Leyden jar electricity can be accumulated and retained

in considerable quantity, by the device of having two conducting surfaces of equal extent, and parallel to each other through the whole of that extent, with a non-conducting substance such as glass between them. When one side of the jar is charged positively, the other is charged negatively (which is why I cited the Leyden jar as an instance in our use of the Method of Agreement). Now, it's impossible to discharge one of the coatings unless the other is discharged at the same time. A conductor held to the positive side can't convey away any electricity unless an equal quantity is allowed to pass from the negative side; if one coating is perfectly insulated, the charge is safe. . . .

The law that this strongly indicates can be corroborated by the Method of Concomitant Variations. The Leyden jar can receive a much higher charge than can ordinarily be given to the conductor of an electrical machine. Now, in the Leyden jar the metallic surface that receives •the induced electricity is a conductor exactly like that which receives •the primary charge, and is therefore as capable of receiving and retaining one charge as the opposite surface is of receiving and retaining the other; but in the machine the neighbouring body that is to get the opposite charge is the surrounding atmosphere or a nearby object; and as these can usually hold only a much smaller charge than the conductor itself, their limited power imposes a corresponding limit to the conductor's capacity for being charged. As the neighbouring body's ability to support the opposition increases, a higher charge becomes possible; and that appears to explain the great superiority of the Leyden jar.

One of Faraday's experiments provides a further and most decisive confirmation by the Method of Difference. [Mill's account of the experiment and the conclusion drawn from it is hard to follow. It speaks of 'two opposite •electric• currents. . . .both accommodated in one wire', and it's hard

to see what Mill has in mind. We can slide past this example without harm to our grasp of the rest of what he has to say.]

§3. Our **third example** will be extracted from Herschel's *Discourse on the Study of Natural Philosophy*, a work full of well-selected examples of inductive processes from almost every branch of physical science. . . . The present example is described by Herschel as 'one of the most beautiful specimens' that can be cited 'of inductive experimental inquiry lying within a moderate compass'—namely, the theory of *dew* that is now accepted by all scientific authorities.

[Mill devotes four pages to this, much of it in direct quotations from Herschel. We can afford to excuse ourselves from going through all the details. Mill shows that the series of tests and experiments make clear use of three of his methods (the exception being the Method of Residues). At a certain point he arrives at this:]

It thus appears that the various instances in which much dew is deposited agree in this (and as far as we can see *only* this): they either •radiate heat rapidly or •conduct it slowly; and those two qualities have nothing in common except that by virtue of either of them the body tends to lose heat from the surface faster than it can be restored from within. And the instances where little or no dew is formed have nothing in common (as far as we can see) except *not* having this same property. So we seem to have detected the characteristic difference between the substances on which dew is produced and those on which it isn't produced. We have done this by using the Joint Method of Agreement and Difference; and the data were prepared for that by the Methods of Agreement and of Concomitant Variations. . . .

Can we be quite sure that the substances on which dew is produced differ from those on which it isn't in *nothing* but the property—I'll call it R—of losing heat from the surface

faster than the loss can be repaired from within? No, but this matters less than you might think. Suppose there is an undiscovered property Q that is present in all the substances that contract dew and absent from those that don't, Q must be present in all the substances that have R and in none of the substances that don't. That much match between two properties creates a strong presumption that they have the same cause and therefore will invariably go together. And if that is right, then the property R—being a better radiator than conductor—if it isn't itself the cause almost certainly always *accompanies* the cause, and for purposes of prediction we can safely treating it as if it really were such.

At an earlier stage of the inquiry we found that whenever dew is formed the surface on which it forms is colder than the surrounding air. Was this coldness the cause of dew or an effect of it? We can now answer this. We have found that when dew forms, the substance on which it forms is one which, by its own properties or laws, would if exposed in the night become colder than the surrounding air. •The coldness is accounted for independently of the dew, while it is proved that •there is a connection between the two; so it must be the case that •the dew depends on the coldness, i.e. that the coldness is the cause of the dew.

This law of causation, already so amply established, can be further corroborated in no less than three ways. **(i)** First, by the Deductive Method. I won't be ready to deal with that until chapter 11, but I'll say enough here to firm up the results concerning dew. It is known by direct experiment that only a limited quantity of water can remain suspended as vapour at each degree of temperature, and that this maximum goes down as the temperature falls. From this it follows deductively that if the air already has much vapour as it can contain at its existing temperature, any lowering of that temperature will cause a portion of the vapour to be

condensed. And we also know deductively, from the laws of heat, that the air's contact with a body colder than itself must lower the temperature of the layer of air immediately against its surface, and will therefore cause it to part with some of its water. And this, by the ordinary laws of gravitation or cohesion—deduction again!—will attach itself to the surface of the body, constituting dew. This deductive proof has the advantage of proving causation as well as coexistence; and it has the further advantage of explaining the exceptions, the cases where the body is colder than the air but no dew is deposited—by showing that this must be the case when the air has too little vapour to give any of it up. That's why in a very dry summer there are no dews, and in a very dry winter no hoar-frost. This is a condition of the production of dew that wasn't detected by the other methods; it might have remained still undetected if we hadn't set out to *deduce* the effect from the known properties of the agents known to be present.

(ii) The second corroboration is by direct experiment according to the canon of the Method of Difference. By cooling the surface of a body we can find the temperature at which dew begins to be deposited. Here again the causation is directly proved. We can accomplish this only on a small scale, but we have ample reason to conclude that the same operation, if conducted in •nature's great laboratory, would equally produce the effect.

(iii) Even on •that great scale we can verify the result. This

is one of the rare [see page 194] cases where nature works the experiment for us in the same way that we ourselves perform it, introducing into the previous state of things a single perfectly definite new circumstance, and producing the effect so rapidly that there's no time for any other material [see Glossary] change in the pre-existing circumstances. Herschel writes:

'It is observed that dew is never copiously deposited in situations much screened from the open sky, and not at all in a cloudy night; but if the clouds withdraw even for a few minutes and leave a clear opening, dew starts to appear almost at once, and goes on increasing. . . Dew formed in clear intervals often evaporates when the sky becomes thickly overcast.'

So we have complete proof that the presence or absence of an uninterrupted communication with the sky causes the deposition or non-deposition of dew. Now, a clear sky is merely the absence of clouds, and we know that clouds. . . tend to raise or keep up the surface temperature of a nearby object by radiating heat to it; so we see at once that the disappearance of clouds will cause the surface to cool. Thus, in this case nature produces a change in the antecedent by definite and known means, and the consequent follows accordingly—a natural experiment that satisfies the requirements of the Method of Difference!¹

The accumulated proof that has been found for the theory of dew is a striking example of the fullness of assurance that

¹ This example may seem to count against my claim that the Method of Difference doesn't apply well to cases of pure observation •as distinct from controlled experiments•; but really it doesn't. Nature seems to have imitated man's type of experiment, but has succeeded only in copying man's most imperfect experiments—namely, those in which he succeeds in producing the phenomenon only by using complex •means that he can't perfectly analyse and therefore can't tell what parts of the effects may be due not to the supposed cause but to some unknown agency of the •means by which that cause was produced. In the natural experiment in question here, the •means was the clearing off a canopy of clouds; and we don't know enough about this process. . . to be certain *a priori* that it couldn't operate upon the deposition of dew independently of any effect on the temperature of the earth's surface. Thus, even in a case as favorable as this to Nature's experimental talents, her experiment is of little value except in corroboration of a conclusion already reached through other means.

the inductive evidence of laws of causation can achieve in cases where the invariable sequence is far from obvious at first glance.

§4. Fourth example: The admirable physiological investigations of Brown-Séguard provide brilliant examples of the use of the inductive methods in a class of inquiries in which—for reasons I'll give soon—direct induction is done under special difficulties and disadvantages. I select his theorising. . . . about the relations between •muscular irritability [see Glossary], •*rigor mortis* [see Glossary], and •putrefaction.

The law that Brown-Séguard's investigation tends to establish, is this:

The greater the degree of muscular irritability at the time of death, the later the *rigor mortis* sets in, and the longer it lasts, and also the later putrefaction appears, and the more slowly it progresses.'

At first glance you'd think that this must be work for the Method of Concomitant Variations, but that is wrong—it's an illusion arising from the fact that the conclusion to be tested is itself a fact about concomitant variations. For the establishment of *that* fact any of the ·four· Methods may be put to work, and it will turn out that the fourth Method—the Method of Concomitant variations—has a real but subordinate place in this investigation.

The items of evidence by which Brown-Séguard establishes the law can be enumerated as follows:

Firstly: (a) •Paralysed muscles have greater irritability than healthy muscles. And (b) paralysed muscles are later in entering *rigor mortis* than healthy muscles, the *rigor* lasts longer, and putrefaction sets in later and proceeds more slowly.

Brown-Séguard proved both these propositions by experiment. He established (a) in various ways, but most

decisively by comparing the duration of irritability in a paralysed muscle and in the corresponding healthy muscle on the opposite side when they are both submitted to the same stimulus. He often found that the paralysed muscle remained irritable up to four times as long as the healthy one. This is induction by the Method of Difference. Because the two limbs were those of the same animal, they were presumed not to differ in any circumstance relevant to the case except the paralysis, so that the presence and absence of paralysis was the source of the difference in the muscular irritability. The assumption that there was only one relevant difference between the legs wasn't safe in any *one* pair of experiments, because the two legs of any given animal might happen to differ in other relevant respects; but if. . . .the experiment was repeated often enough with different animals to exclude the supposition that any abnormal circumstance could be present in them all, the conditions of the Method of Difference were well enough satisfied.

Brown-Séguard also proved the proposition (b) concerning *rigor mortis* and putrefaction. Having. . . .cut some nerves so as to produce paralysis in one hind leg of an animal but not the other, he found that muscular irritability lasted much longer in the paralysed limb, *rigor* set in later and ended later, and putrefaction began later and progressed more slowly than on the healthy side. This is a routine use of the Method of Difference, requiring no comment. An important corroboration was obtained by the same method. When the animal was killed not •soon after the nerves were cut but •a month later, the effect was reversed; *rigor* set in sooner and lasted a shorter time in the paralysed limb than in the healthy one. What had happened was this: During the month before death the paralysed muscles were of course •resting, and thereby •losing much of their irritability and eventually becoming *less* irritable than the muscles on the

healthy side. This gives the

A B C — *a b c* and

B C — *b c*

of the Method of Difference. When one antecedent (increased irritability) was changed and the other circumstances kept the same, the consequent didn't follow; and when a new antecedent was provided, contrary to the first, it was followed by a contrary consequent. This has the special advantage of proving that the delay and slowing of *rigor mortis* don't depend directly on the paralysis, because that was the same in both cases, but on one effect of the paralysis, namely the increased irritability—they stopped when it stopped, and were reversed when it was reversed.

Secondly: Lowering the temperature of muscles before death increases their irritability, and also delays *rigor mortis* and putrefaction.

It was Brown-Séguard himself who made these truths known, through experiments that conform to the Method of Difference. There's nothing in the nature of the process that requires comment.

Thirdly: When muscular exercise is continued to exhaustion, that lessens the muscular irritability. This is a well-known truth that depends on the most general laws of muscular action and is proved by constantly repeated experiments using the Method of Difference. Now, observation has shown that if cattle are driven too hard and then killed before they recover from their fatigue, their bodies become rigid and putrefy in a surprisingly short time. The same thing has been observed in animals hunted to death, cocks killed during or shortly after a fight, and soldiers slain in battle. The only thing involving the muscles that all these have in common is their having just been subjected to exhausting exercise. Under the canon of the Method of Agreement, therefore, we can infer that there is a connection between

the two facts. We have seen that the Method of Agreement can't prove causation; but we already know that what we're dealing with here is causation. It's certain that the body's state after death must somehow depend on its state at the time of death; so we are justified in concluding that the single circumstance shared by all the instances is the part of the antecedent that causes that particular consequent.

Fourthly: In proportion as the nutrition of muscles is in a good state, their irritability is high; this is supported also by laws of physiology based on many familiar applications of the Method of Difference. Now, when someone (or some animal) dies from accident or violence, with his muscles in a good state of nutrition, •the muscular irritability continues long after death, •*rigor* sets in late, and •it continues for a long time without putrefaction. On the other hand, in cases of disease where nutrition has been diminished for a long time before death, all these effects are reversed. This satisfies the conditions of the Joint Method of Agreement and Difference. These cases of delayed and long continued *rigor* agree only in being preceded by a high state of nutrition of the muscles; the cases of rapid and brief *rigor* agree only in being preceded by a low state of muscular nutrition; so a connection is inductively proved between •the degree of the nutrition and •the slowness and prolongation of the *rigor*.

Fifthly: Convulsions lessen the muscular irritability, like exhausting exercise but even more. When death follows violent and prolonged convulsions—as in tetanus, hydrophobia, some cases of cholera, and certain poisons—*rigor* sets in very rapidly and after a very little time gives place to putrefaction. This involves the Method of Agreement in the same way as 'Thirdly. . . '.

Sixthly: The last series of instances that I'll present is more complex and requires a more finely detailed analysis.

It has long been observed that in some cases of death by lightning *rigor mortis* either doesn't occur at all or doesn't last long enough to be noticed, and that in these cases putrefaction is very rapid; whereas in other cases the usual *rigor mortis* appears. There must be some difference in the cause to account for this difference in the effect. [Mill reports the experimental work by Brown-Séquard that located the line between the two kinds of death by lightning (which he brought within experimental reach by substituting artificial galvanic shocks for natural lightning), namely: When and only when the eclectic shock produced muscular convulsions throughout the body, the irritability of the muscles went down, and the duration of the *rigor* went down with it. We can safely spare ourselves the details, and rejoin Mill when he quotes Brown-Séquard's summing up of his findings from all the work described in this section:]

'When the degree of muscular irritability at the time of death is considerable, either because of

- a good state of nutrition, as in persons who die in full health from an accidental cause, or
- rest, as in cases of paralysis, or
- the influence of cold,

rigor mortis sets in late and lasts long, and putrefaction appears late and progresses slowly; but when the degree of muscular irritability at the time of death is slight, either because of

- a bad state of nutrition, or
- exhaustion from overexertion, or
- convulsions caused by disease or poison,

rigor mortis sets in and ceases soon, and putrefaction appears and progresses quickly.'

These facts completely satisfy the conditions of the Joint Method of Agreement and Difference. Early and brief *rigor* takes place in cases that agree only in having a low state

of muscular irritability. *Rigor* begins late and lasts long in cases that agree only in the opposite circumstance of high and unusually prolonged muscular irritability. It follows that there's a causal connection between the degree of muscular irritability after death and the tardiness and length of the *rigor mortis*.

This investigation shines a strong light on the value and efficacy of the Joint Method. We have seen that the defect of that Method—as of the Method of Agreement—is that it can't prove causation. But in the present case (as in one of the steps in the argument leading up to it) causation is already proved; because there could never be any doubt that the *rigor* and the ensuing putrefaction are caused by death; the empirical basis for this is too familiar to need analysis, and falls under the heading of the Method of Difference. So we know beyond doubt that the aggregate antecedent, the death, is the actual cause of the whole sequence of consequences; and we can get more fine-grained results—'The death's being of *this* kind is the cause of such-and-such a feature of the upshot'—when variations in the manner of death can be shown to match corresponding variations in the effect we are investigating. . . .

§5. Some more examples: The examples I have presented offer such a clear conception of the use and practical management of three of the four methods of experimental inquiry that there's no need to give further examples of them. There remains the Method of Residues, which hasn't yet made an appearance in this chapter. I shall quote from Herschel some examples of that method, with the remarks by which they are introduced.

'It is by this process that science in its present advanced state is chiefly promoted. Most natural phenomena are very complicated; and when the effects of all known causes

are estimated exactly and set aside, the residual facts are constantly appearing in the form of entirely new phenomena that lead to the most important conclusions.

'For example, the return of the comet predicted by Professor Encke a great many times in succession, and the general good agreement of its calculated place with its observed place during any one of its periods of visibility, would lead us to say that its gravitation toward the sun and planets is the sole and sufficient cause of all the facts about its orbital motion; but when the effect of this cause is strictly calculated and subtracted from the observed motion, there remains a residual phenomenon that would never have been known to exist if this method weren't used. This residue is a small diminution of the comet's periodic time that can't be accounted for by gravity, and whose cause is therefore to be inquired into. Such a diminution would be caused by the resistance of a medium disseminated through the celestial regions; and as there are other good reasons for believing this to be a *vera causa* [see Glossary] it has therefore been ascribed to such a resistance.' [The idea is: resistant medium → slower movement → less propulsive force relative to centripetal force → greater tendency towards the sun → shorter journey → shorter time.]

[Herschel's next example is actually not a use of the Method of Residues, Mill says. There are several more, but we can settle for one more, introduced again by Herschel:]

'Unexpected and striking confirmations of inductive laws frequently occur in the form of residual phenomena, during investigations that are nothing like the ones that led to the inductions themselves. An elegant example is the unexpected confirmation of the law of *the development of heat in elastic fluids by compression*, which is provided by the phenomena of sound. The inquiry into the cause of sound had led to conclusions about its mode of propagation, from which its velocity in the air could be precisely calculated. The

calculations were performed, and the results were near enough to right to show the general correctness of the theory about the cause and the mode of propagation; but this theory couldn't be shown to account for all the sound's velocity. There was still a residual velocity to be accounted for, and for a long time this remained a puzzle. Eventually Laplace had the nice idea that it might come from the heat developed by the condensation that necessarily takes place at every vibration by which sound is conveyed. This matter was subjected to exact calculation, and the immediate result was the complete explanation of the residual phenomenon, and a striking confirmation of the general law of the development of heat by compression, under circumstances beyond artificial imitation.'

§6. Whewell has expressed an unfavourable opinion of the utility of the Four Methods, as well as of the aptness of the examples by which I have tried to illustrate them. He writes:

'The obvious thing to say about these methods is that they take for granted the very thing that it's hardest to discover, the reduction of the phenomena to formulae such as are here presented to us. When we have any set of complex facts offered to us. . . ., and we want to discover the law of nature that governs them—or, if you want to put it this way, the feature in which all the cases agree—where are we to look for our A, B, C, and *a, b, c*? Nature doesn't present the cases to us in this form; and how are we to reduce them to this form? You say when we find the combination of A B C with *a b c* and A B D with *a b d*, then we may draw our inference. Granted; but when and where are we to find such combinations? Even now that the discoveries are made, who will point out to us what are the A, B, C, and *a, b, c*, elements of

the cases that have just been enumerated? [He has cited ones from astronomy, mechanics, optics, and chemistry.] Who will tell us which of the methods of inquiry those historically real and successful inquiries exemplify? Who will carry these formulae through the history of the sciences, as they have really grown up, and show us that these four methods have been operative in their formation; or that any light is thrown upon the steps of their progress by reference to these formulae?

He adds that in this work of mine the methods haven't been applied 'to a large body of conspicuous and undoubted examples of discovery, extending along the whole history of science'; which ought to have been done if the methods were to be shown to have the advantage. . . .of being those 'by which all great discoveries in science have really been made'.

These objections against the Canons of Induction are strikingly like the 18th century objections, by men as able as Whewell, against the acknowledged Canon of Ratiocination. Those who protested against the Aristotelian logic said of the syllogism what Whewell says of the inductive methods, namely that it 'takes for granted the very thing that is most difficult to discover, the reduction of the argument to formulae such as are here presented to us'. The great difficulty, they said, is to obtain your syllogism, not to judge its correctness when obtained. On the matter of fact, they and Whewell are right. The greatest difficulty in both cases is •obtaining the evidence and then •reducing it to the form that tests its conclusiveness. But if we try to reduce it without knowing what it's to be reduced to we're not likely to make much progress. It's harder to solve a geometrical problem than to judge whether a proposed solution is correct; but if people couldn't judge the solution when it was found, they would have little chance of finding it. And it can't be maintained that to judge an induction once it has been found

is perfectly easy, a thing for which aids and instruments are superfluous; for erroneous inductions, •false inferences from experience, are quite as common as—and on some subjects much commoner than—•true ones. The business of inductive logic is to provide rules and models (such as the syllogism and its rules are for ratiocination) such that inductive arguments are conclusive if, and only if, they conform to them. That's what the four methods claim to be, and what I believe they are considered to be by all experimental philosophers, who had practised all of them long before anyone tried to reduce the practice to theory.

The assailants of the syllogism also anticipated Whewell in the other branch of his argument. They said that no discoveries were ever made by syllogism; and Whewell seems to say, that none were ever made by the four methods of induction. To the former objectors Whately gave a good answer, namely that if their argument was any good it was good against the reasoning process altogether; for whatever can't be reduced to syllogism isn't reasoning. And Whewell's argument, if good at all, is good against all inferences from experience. In saying that no discoveries were ever made by the four methods, he affirms that none were ever made by observation and experiment; for assuredly if any were, it was by processes reducible to one or other of those methods.

This difference between us accounts for his dissatisfaction with my examples, which I didn't select with a view to showing that observation and experiment are ways of acquiring knowledge. In choosing them I was thinking only of •illustration, and of •making methods easier to grasp by examples. If I had wanted to justify the processes themselves as means of investigation, I wouldn't have needed to look far off or use recondite or complicated instances. As a specimen of a truth ascertained by the Method of Agreement, I could have chosen the proposition 'Dogs bark'. This dog and that

dog and the other dog answer to A B C and A D E and A F G. The circumstance of being a dog answers to A. Barking answers to *a*. As a truth made known by the Method of Difference, 'Fire burns' might have sufficed. Before I touch the fire I am not burned; this is B C; I touch it, and am burned; this is A B C and *a* B C.

Whewell doesn't regard such familiar experimental processes as inductions; but they are perfectly homogeneous with the ones on which, even on his own showing, the pyramid of science is based. He tries to escape from this conclusion by arbitrarily restricting the range of examples that can serve as instances of induction: they must not be

- things that are still matters of discussion,
- drawn from mental and social subjects, or
- drawn from ordinary observation and practical life.

They must all concern generalisations by which scientific thinkers have ascended to great and comprehensive laws of natural phenomena. Now it is seldom possible in these complicated inquiries to go much beyond the first steps without making use of deduction and the temporary aid of hypotheses—this being something that Whewell and I have maintained against the purely empirical school—so that such cases wouldn't serve well as illustrations of the principles of mere observation and experiment. Whewell is misled by their absence into representing the experimental methods as serving no purpose in scientific investigation, forgetting that if those methods hadn't supplied the first generalisations there would have been no materials for his own conception of induction to work on.

But it's easy to answer his challenge to say which of the four methods are involved in certain important scientific developments. The planetary paths, as far as they are a case of induction at all [see page 147], involves the Method of Agreement. The law of 'falling bodies', namely that they

cover distances proportional to the squares of the times, was historically a deduction from the first law of motion; but the experiments that verified it and could have led to its discovery involved the Method of Agreement; and the apparent variation from the true law caused by air-resistance was cleared up by experiments *in vacuo*, involving the Method of Difference. . . . The movements of comets were determined by highly complex processes of thought in which deduction was predominant, but the Methods of Agreement and of Concomitant Variations had a large part in establishing the empirical laws. Every case. . . is a well-marked example of the Method of Difference. To anyone acquainted with the subjects—to Whewell himself—there wouldn't be the slightest difficulty in setting out 'the A B C and *a b c* elements' of these cases.

If discoveries are ever made by observation and experiment without deduction, the four methods are methods of •discovery; but even if they weren't, they would still be the sole methods of •proof; and they could serve as proofs even of the results of deduction. The great generalisations that begin as hypotheses must end by being proved, and in due course I'll show that they are in fact proved by the four methods. Now logic is principally concerned with proof as such. This approach has no chance of finding favour with Whewell, because his system has the special feature that it doesn't recognise any need for proof in cases of induction. If an hypothesis is carefully collated with facts, and nothing inconsistent with it turns up—i.e. if experience doesn't disprove it—Whewell is content, at least until we find a simpler hypothesis that is equally consistent with experience. If this is induction, doubtless there is no need for the four methods. But to suppose that it is induction seems to me a radical misunderstanding of the nature of the evidence for physical truths.

There's a real practical *need* for a test for induction, like the syllogistic test of ratiocination. Inferences that defy the most elementary notions of inductive logic are confidently presented by persons eminent in physical science, as soon as they are off the factual ground that they know. . . . As for educated persons in general, I doubt that they are better judges of a good or a bad induction than they were before Bacon wrote. The improvement in the •results of thinking has seldom extended to the •processes; and if it *has* reached any process it has been that of investigation only and not that of proof. No doubt a knowledge of many laws of nature has been arrived at by forming hypotheses and finding that the facts corresponded to them; and many errors have been cured by coming to know facts that were

inconsistent with them, but not by discovering that the mode of thought that led to the errors was itself faulty and could have been known to be faulty independently of the facts that disproved the specific conclusion. The upshot is that while mankind's thoughts on many subjects have worked out well in practice, the thinking power remains as weak as ever. In all subjects where the facts that would check the result are not accessible—e.g. in what relates to the invisible world, and even. . . .to the visible world of the planetary regions—men with the greatest scientific acquirements argue as pitifully as the merest ignoramus. They have made many sound inductions, but they haven't learned from them—and Whewell thinks there is no need for them to learn—the principles of inductive *evidence*.

Chapter 10. Plurality of causes, and the intermixture of effects

§1. In my account of the four methods of observation and experiment by which we contrive to sort out among a mass of coexistent phenomena the particular effect of a given cause, or the particular cause of a given effect, I have had to suppose for simplicity's sake that this analytical operation doesn't run into difficulties other than the ones that are essentially inherent in its nature. So I have represented every effect as connected exclusively with a single cause, and as incapable of being confusingly mixed in with any other coexistent effect. I have regarded *a b c d e*, the aggregate of the phenomena existing at any moment, as consisting of dissimilar facts—*a* and *b* and *c* and *d* and *e*—for each of which we need to look for just one cause; the difficulty being only that of singling out this one cause from the multitude of antecedent circumstances A, B, C, D, and E. The cause may indeed not

be simple; it may consist of an assemblage of conditions; but I have supposed that there's only one possible assemblage of conditions from which the given effect could result.

If that were right, it would be comparatively easy to investigate the laws of nature. But the supposition is false in both its parts. **(i)** It's not true that the same phenomenon is always produced by the same cause; the effect *a* may sometimes arise from A, sometimes from B. **(ii)** And the effects of different causes are often not dissimilar but homogeneous, and not demarcated by any assignable boundaries; A and B, instead of producing *a* and *b*, may produce different parts of an effect *a*. Investigation of the laws of phenomena is made much harder and darker by the need to take account of these two circumstances: intermixture of effects, and plurality of causes. I'll take the latter first, because it is the simpler of

the two. I'll start on the intermixture of effects in section 4.

Here's the situation that we face. It's not true that one effect must be connected with only one cause or assemblage of conditions; i.e. that each phenomenon can be produced in only one way. There are often several independent ways in which the same phenomenon could have originated. . . . Many causes can produce mechanical motion; many causes can produce some kinds of sensation; many causes can produce death. It can happen that a given effect was produced by a certain cause but could perfectly well have been produced without it.

§2. One of the principal consequences of this fact of plurality of causes is to bring uncertainty into the Method of Agreement. I illustrated that method by supposing two instances:

- A B C followed by $a b c$, and
- A D E followed by $a d e$.

To avoid a difficulty that isn't relevant to my present theme, let us suppose that we know for sure that the two cases have no antecedent in common except A. Then it might seem that we have a basis for concluding that A is an invariable antecedent of a , and even that it is its unconditional invariable antecedent, i.e. its cause. But the moment we admit the possibility of a plurality of causes, that conclusion fails. Why? Because it tacitly assumes that a must have been produced in both instances by the same cause. If there could have been two causes, they might have been (for example) C and E; with C causing a in the former of the instances and E in the other, and A having no influence in either case.

Suppose we investigate the circumstances of the upbringing and history of two great artists (or it could be two great philosophers, two extremely selfish men, or two extremely generous men) and find that their antecedents agree only in one circumstance x ; would it follow that x was what caused

each to be a great artist (or a great philosopher or . . .)? Not at all! The causes that can produce any type of character are very numerous; and the two persons could have been just as alike in character without there being *any* resemblance between their previous histories.

This is a **characteristic imperfection** of the Method of Agreement, from which the Method of Difference is free. For if we have two instances A B C and B C, of which B C gives $b c$, and the addition of A converts it into $a b c$, it's certain that at least in this instance A was either the cause of a or an indispensable portion of its cause, even if in other instances a is produced by entirely different causes. Plurality of causes, therefore, doesn't make the Method of Difference less reliable, and doesn't even require a greater number of observations or experiments; two instances, one positive and the other negative, are still enough for a complete and rigorous induction. Not so with the Method of Agreement. The conclusions that *it* yields when the number of instances is small are of no real value unless they function as *suggestions* that may lead either to experiments bringing them to the test of the Method of Difference or to reasonings that can explain and verify them deductively.

When the instances are indefinitely multiplied and varied and *still* suggest the same result, then (and only then) we have an independently valuable result. If the only instances of production of a are A B C and A D E, though these instances have nothing in common except A, the effect a may have been produced in the two cases by different causes so that there's at most only a slight probability in favour of A; there *may be* causation but it's almost equally probable that there was only a coincidence. But the oftener we repeat the observation, varying the circumstances, the more we advance toward a solution of this doubt. For if we try A F G, A H K, etc., all unlike except in containing A, and if we find

the effect *a* appears in all these cases, we must suppose one of two things: •that *a* is caused by A, or •that it has as many different causes as there are instances. With each addition, therefore, to the number of instances, the presumption is strengthened in favour of A. The inquirer will take any chance he gets to exclude A from one of these combinations—let's say from A H K—and by trying H K separately bring the Method of Difference to the aid of the Method of Agreement. Only the Method of Difference can show us that A is the cause of *a*; but the Method of Agreement, provided the instances are numerous and sufficiently various, can put it beyond any reasonable doubt that A is either the cause of *a* or an effect of the cause of *a*.

How many varied instances with only A in common does it take to •rule out the supposition of a plurality of causes and •make it virtually certain that *a* is connected with A? We mustn't dodge this question, but the consideration of it belongs to the theory of probability, which I'll come to in chapter 17. Still, we can see right away •that the conclusion does amount to a practical certainty after a sufficient number of instances, and thus •that the method isn't radically discredited by the **characteristic imperfection**. There are two upshots to these considerations, **(1)** We see a new source of inferiority in the Method of Agreement, and new reasons for never resting content with results obtained by it without trying to confirm them either by the Method of Difference or by connecting them deductively with some law already ascertained by that superior method. **(2)** We learn the true theory of the value of mere number of instances in inductive inquiry. The plurality of causes is the only reason why mere number is of any importance. Unscientific inquirers tend to rely too much on number, without analysing the instances—without looking into their nature closely enough to see what circumstances are or aren't eliminated by means

of them. Most people hold their conclusions with a degree of assurance proportioned to the mere mass of the experience they appear to rest on, overlooking the fact that by adding instances to instances, differing from one another only in details already recognised as immaterial, nothing whatever is added to the force of the conclusion. A single instance eliminating some antecedent that existed in all the other cases is of more value than the greatest multitude of instances that are reckoned by their number alone. We do of course have to assure ourselves, by repetition of the observation or experiment, that we haven't committed any error concerning the individual facts observed; and until we are sure about this our primary need is not to vary the circumstances but to repeat the same experiment or observation, very carefully, without any change. But once we have this assurance, the multiplication of instances that don't exclude any more circumstances is entirely useless, provided there have been already enough to exclude the supposition of plurality of causes.

This is important: . . . the Joint Method of Agreement and Difference is not affected by the characteristic imperfection of the Method of Agreement. In the joint method it is supposed not only that •the instances in which *a* is •an effect• agree only in containing A, but also that •the instances in which *a* is not •an effect• agree only in not containing A. If that's how things stand, A must be not only •the cause of *a* but •the only possible cause; for if there were another—say, B—then in the instances in which *a* is not •an effect• B must have been absent as well as A, and it wouldn't be true that these instances agree only in not containing A. This is an immense advantage of the •joint method over the •simple Method of Agreement. It may seem, indeed, that the advantage belongs to the negative part of the joint method rather than to the method as a whole. The Method of Agreement, when

applied to negative instances (i.e. ones where a phenomenon does not take place), is certainly free from the characteristic imperfection which affects it in the affirmative case. So you might think that the negative premise could be worked as a simple case of the Method of Agreement, with no need for an affirmative premise to go with it. But though this is true in principle, it's usually impossible to work the Method of Agreement by negative instances without positive ones, because it's so much harder to exhaust the field of negation than the field of affirmation. For example: if we are inquiring into what makes bodies transparent, what are our chances of success if we try to discover what the various substances that *aren't* transparent have in common. We are more likely to succeed in seizing some point of resemblance among the comparatively few and definite kinds of things that *are* transparent; and when we've done this our natural next task is to look into whether the absence of this one circumstance isn't precisely the respect in which all opaque substances will be found to be alike.

So the Joint Method of Agreement and Difference. . . is, after the Direct Method of Difference, the most powerful of the instruments of inductive investigation that I haven't yet discussed; and in the sciences that depend on pure observation with little or no aid from experiment, this method—so well illustrated by the theorising about the cause of dew—is the primary resource, so far as direct appeals to experience are concerned.

§3. Up to here I have treated plurality of causes only as a *possible* supposition that makes our inductions uncertain until we have eliminated it; and have considered how we can eliminate it in cases where there isn't in fact a plurality of causes. But we must also consider it as something that actually occurs in nature, and find ways for our methods of

induction to be able to identify the cases where it does occur. We don't need any special method for doing this. When an effect really could be produced by either of two (or more) causes, the process for detecting them is exactly the same as the process for discovering single causes. They may (first) be discovered as separate sequences, by separate sets of instances—i.e. of observations or experiments—showing that the causes of heat include

- the sun,.
- friction,
- percussion,
- electricity,
- chemical action,

with each of these being shown by its own special set of instances. Or (secondly) the plurality may come to light when we are collating a number of instances in an attempt to find something that they all have in common. A *failing* attempt: we can't find anything that is common to all instances of heat; we find that no one antecedent is present in all the instances, no one of them indispensable to the effect. But when we look harder we find that although no *one* is always present, *one or other of several* always is. If on further analysis we can detect any common element in these, we may be able to ascend from them to some one cause that is the really operative circumstance in them all. Thus it is now thought that a single *ultimate* source is at work in the production of heat by friction, percussion, chemical action, etc. But if (as continually happens) we can't take this further step, the different antecedents must be noted provisionally as distinct causes each of which is sufficient, unaided, to produce the effect.

I now move from the plurality of causes to the still more special and more complex case of •the intermixture of effects and •the interference of causes with one another. This is the

principal source of complication and difficulty in the study of nature; and we'll soon see that the four inductive methods that I have presented—the only *possible* methods of directly inductive investigation by observation and experiment—are for the most part quite unable to cope with it. Our only means for unravelling the complexities proceeding from the intermixture of effects and the interaction amongst causes is *deduction*; and the four methods can't do much more than supply premises for our deductions and check their conclusions.

§4. A concurrence of two or more causes, not separately producing each its own effect but interfering with or altering one another's effects, happens in two ways. **(1)** In one, exemplified by the joint operation of different forces in mechanics, the separate effects of all the causes continue to be produced, but are compounded with one another and make one total. **(2)** In the other, already mentioned on page 184 and illustrated by the case of chemical action, the separate effects cease entirely and are succeeded by phenomena that are altogether different and governed by different laws.

Of these **(1)** is by far the more frequent, and also the more likely to elude the grasp of our experimental methods. The exceptional case **(2)** is basically open to being handled by them. When

- the laws of the original agents cease to be applicable because a new phenomenon appears that doesn't offer a hand-hold for those laws, e.g. when
- two gaseous substances, hydrogen and oxygen, are brought together and throw off their special properties and produce water,

in such cases the new fact can be subjected to experimental inquiry, like any other phenomenon; and the elements that

are said to compose it can be regarded as the mere agents of its production—the conditions on which it depends, the facts that make up its cause.

The *effects* of the new phenomenon—e.g. the *properties* of water—are as easily found by experiment as the effects of any other cause. But to discover the *cause* of it, i.e. the particular conjunction of agents from which it results, is often difficult enough. **(a)** The origin and actual production of the phenomenon are usually out of reach of our observation. If we couldn't have learned the composition of water until we *found* instances where it was actually produced from oxygen and hydrogen, we'd have been forced to *wait* until someone had the random idea of passing an electric spark through a mixture of the two gases, or inserting a lighted taper into it, merely to see what would happen. **(b)** Many substances can be analysed but can't be recomposed by any known artificial means. **(c)** Even if we could have learned by the Method of Agreement that oxygen and hydrogen are both present when water is produced, no experiments with oxygen and hydrogen separately—no knowledge of their separate laws—could have enabled us to infer deductively that they would produce water. For that we need a specific experiment on the two combined.

Given these difficulties, you might expect that our knowledge of the causes of this class of effects comes either from accident or from the gradual progress of experimentation on the different combinations that the producing agents are capable of. But we can often do better than that, because effects of this kind have the special feature that under certain combinations of circumstances they reproduce their causes. Water results from putting hydrogen and oxygen really closely and intimately together, and correspondingly hydrogen and oxygen result from placing water in certain situations. When that happens the new laws—i.e. the laws of water—abruptly

cease and the two gases re-appear separately with their own properties. What is called 'chemical analysis' is the process of searching for the causes of a phenomenon among its effects, or rather among the effects of subjecting it to some other causes.

Lavoisier, by heating mercury to a high temperature in a closed vessel containing air, found that the mercury increased in weight and became 'red precipitate', while the air when examined after the experiment turned out to have •lost weight and •lost its ability to support life or combustion. When red precipitate was exposed to a still greater heat it became mercury again and gave off a gas which *did* support life and flame. . . .

Where two phenomena between whose laws or properties no connection can be traced are thus cause-effect and effect-cause, each capable in its turn of being produced from the other, and each when it produces the other ceases itself to exist (as water is produced from oxygen and hydrogen, and oxygen and hydrogen are reproduced from water); this causation of the two phenomena by one another, each being generated by the other's destruction, is strictly *transformation*. The idea of chemical composition is an idea of transformation, but of a transformation that is *incomplete* because we consider the oxygen and hydrogen to be present in the water as oxygen and hydrogen, discoverable in it if our senses were keen enough. That's a mere a supposition, based solely on the fact that the weight of the water is the sum of the separate weights of the two ingredients. This fact about the weights is an exception to the entire disappearance in the compound of the laws of the separate ingredients. . . .

In these cases, where the heteropathic effect (as I called it on page 184) is merely a transformation of its cause—i.e. where an effect and its cause are also a cause and its effect, and are mutually convertible into each other—the problem

of finding the cause resolves itself into the far easier one of finding an effect, which is the kind of inquiry that *can* be performed by direct experiment. But in some cases of heteropathic effects this can't be done. Consider for instance the heteropathic laws of mind; the part of the phenomena of our mental nature that are analogous to •chemical rather than to •dynamical phenomena; as when a complex passion is formed by the coalition of several elementary impulses, or a complex emotion by several simple pleasures or pains of which it is the result without being the aggregate or in any way homogeneous with them. in these cases the product is generated by its various factors, but the factors can't be reproduced from the product. . . . We can't ascertain what simple feelings any of our complex states of mind are generated from, in the way we ascertain the ingredients of a chemical compound by making it generate them. So our only way to discover these laws is the slow process of studying the simple feelings themselves, and learning by experimenting on the various combinations they're capable of, what they can generate by their interactions.

§5. One might have thought that the other and apparently simpler sort of the mutual interference of causes, where each cause continues to obey the laws that it conformed to in its separate state, would have presented **fewer** difficulties to the inductive inquirer than does the one I have been discussing. In fact, however, it presents—so far as direct induction without help from deduction is concerned—ininitely **greater** difficulties. When a concurrence of causes gives rise to a new effect that has no relation to the separate effects of those causes, the resulting phenomenon stands out undisguised. . . ., presenting no obstacle to our recognising its presence or absence among any number of surrounding phenomena. So it can easily be brought under the canons

of induction if instances of the right kind can be obtained. And the non-occurrence of such instances, or the lack of means to produce them artificially, is the real and only difficulty in such investigations—not a *logical* difficulty but in some way a *physical* one. What I (in chapter 5) called the 'composition of causes' is not like that. There, the effects of the separate causes don't terminate and give place to others, thereby ceasing to form any part of the phenomenon to be investigated; on the contrary, they hold their place but are intermingled with, and disguised by, the homogeneous and closely allied effects of other causes. They are no longer

a, b, c, d, e

existing side by side and continuing to be separately discernible; rather, they are

$+a, -a, \frac{1}{2}b, -b, 2b$, etc..

some of which cancel one another, while many others don't appear separately but merge in one sum. Between •their over-all result and •the causes by which it was produced there's often an insurmountable difficulty in seeing any fixed relation whatever.

We have seen that according to the general idea of the composition of causes:

Two or more laws interfere with one another, apparently frustrating or modifying one another's operation, yet in reality all are fulfilled, the collective effect being the exact sum of the effects of the causes taken separately.

As a familiar example, think about a body in equilibrium by two equal and opposite forces. One on its own would carry the body in an hour one mile westward, the other on its own would carry it in an hour one mile eastward; and the result of the equilibrium is just the same as if the body had first been carried westward by one force and then back eastward by the other, being finally left where it was at first.

Every causal law L_1 is liable to be counteracted—seemingly frustrated—by coming into conflict with another law L_2 (or more than one) the separate result of which is opposite to L_1 's, more or less inconsistent with it. The result of that is that many instances in which L_1 really is entirely fulfilled don't at first sight seem to involve its operation at all. An example of that is the west-east one that I just offered: a 'force' in mechanics means precisely a 'cause of motion', but it can happen that the sum of the effects of two causes of motion is rest = motionlessness. Another example: a body subjected to two forces in different directions moves in the diagonal; and it seems a paradox to say that motion in the diagonal is the sum of two motions in two other lines. [The 'diagonal' referred to here is, as Mill explained on page 161, the diagonal of a parallelogram whose sides represent the direction and strength of those two forces.] Motion, however, is merely change of place, and at every instant the body is in the exact place it would have been in if the forces had acted during alternate instants instead of acting in the same instant (except that we must of course allow the forces double the time if they're to do successively what they in fact do simultaneously). So it's clear that each force has had during each instant its own full effect, and that the *modifying* influence that C_2 is said to exercise with respect to C_1 can be seen as exerted not over the action of C_1 itself but over the effect after C_1 has done its work. For all purposes of predicting, calculating, or explaining their joint result, causes that compound their effects can be treated as if •they produced their own separate effects simultaneously, and •all these effects coexisted visibly.

Because the laws of causes are just as completely fulfilled •when the causes are 'counteracted' by opposing causes as they are •when they are left to their own undisturbed action, we must take care not to express the laws in terms that would make the assertion of their being fulfilled in those

cases a contradiction. For example, if we said that there's a law of nature according to which

any body to which a force is applied moves in the direction of the force with a velocity directly proportional to the force and inversely proportional to its own mass,

when in fact

some bodies to which a force is applied don't move at all, and the ones that do move (at least in the region of our earth) are from the very first held back by the action of gravity and other resisting forces, and eventually stopped altogether,

it's clear that the general proposition—the supposed 'law'—though true under a certain hypothesis doesn't express the facts as they actually occur. To get the expression of the law to fit the real phenomena we must say not that the object *moves* in the direction and with the velocity specified but that it *tends* to move in that way. We could guard our expression in a different way by saying that the body moves in that manner *unless prevented*, or *except to the extent that it is prevented*, by some counteracting cause. But that's less good, because the body doesn't merely

move in that manner unless counteracted;

it also

tends to move in that manner even when it is counteracted.

It still exerts, in the original direction, the same energy of movement as if its first impulse had been undisturbed, and produces by that energy an exactly equivalent quantity of effect. This is true even when the force leaves the body as it found it, in a state of absolute rest. Suppose we are trying to raise a body weighing three tons with a force equal to one ton; if while we're doing this wind or water or any other agent supplies an additional force of just over two tons, the

body will be raised—proving that the force we applied exerted its full effect by neutralizing an equivalent part of the total weight. And if, while we're exerting this force of one ton on the object in a direction contrary to that of gravity, it is put onto a scale and weighed, it will be found to have lost a ton of its weight—i.e. to press downward with a force equal to only the difference between the two forces.

These facts are correctly indicated by the term 'tendency'. Because laws of causation can be counteracted, they should all be stated in terms of tendencies only, not actual results. . . .

The habit of neglecting this needed element in the precise expression of the laws of nature has given rise to the popular [see Glossary] prejudice that all general truths have exceptions; and this has brought much unmerited distrust to the conclusions of science when they have been submitted to the judgment of minds insufficiently disciplined and cultivated [see Glossary]. The rough generalisations suggested by common observation usually do have exceptions; but principles of science—i.e. laws of causation—don't. Let me quote from an earlier work of mine [from here to the end of this section]: 'What is thought to be an exception to a principle is always some *other* principle cutting into the former, some other force that impinges against the first force and pushes it off-course. We do *not* have this:

a law and an exception to it, the law acting in 99 cases and the exception in one.

What we do have is this:

Two laws, each possibly acting in the whole 100 cases, bringing about a common effect by their joint operation.

If in a single case the less conspicuous force—called the "disturbing" force—prevails sufficiently over the other force to create what is commonly called an "exception", the same

disturbing force probably acts as a *modifying* cause in many other cases that no-one will call exceptions.

'Thus if it were said to be a law of nature that **all heavy bodies fall** to the ground, it would probably be said that the resistance of the atmosphere that prevents a balloon from falling makes the balloon an "exception" to that supposed law of nature. But the real law is that **all heavy bodies tend to fall**; and there are no exceptions to this, not even the sun and moon, because (as every astronomer knows) they tend toward the earth with a force exactly equal to that with which the earth tends toward them. . . .' [From Mill's *Essays on Some Unsettled Questions of Political Economy*, Essay 5.]

§6. We now have to face the question: *How* are we to study these complex effects made up of the effects of many causes? *What* enable us to trace each effect back to the concurrence of causes in which it originated, and learn what the circumstances are in which it may be expected to recur?

The conditions of a phenomenon that arises from a composition of causes can be investigated either •deductively or •experimentally.

It's obvious that the deductive mode of investigation is appropriate to this kind of case. The law governing an effect of this sort, x, is an upshot of the separate laws governing the causes that jointly produced x, so it is in itself capable of being deduced from these laws. This is called the *a priori* method. The *a posteriori* method claims to proceed according to the canons of experimental inquiry. Considering the whole assemblage of causes that jointly produced x as one single cause, it tries to ascertain the cause in the ordinary manner, by a comparison of instances. There are two varieties of this second method. If it merely assembles and compares instances of the effect, it's a method of pure observation. If it operates on the causes and tries different combinations

of them in hopes of eventually hitting the precise combination that will produce the given total effect, it is a method of experiment.

So we have three methods: deductive, observational, and experimental. In order get clearer about the nature of each, and determine which of them deserves preference, I shall 'clothe them in circumstances' (Lord Eldon's phrase). I'll select for this purpose a case that hasn't yet provided a brilliant example of the success of any of the three methods, but does illustrate the difficulties inherent in them. Let's suppose we are inquiring into. . .the conditions of recovery from a given disease x; and for a start let our question be limited to: 'Is mercury a remedy for x?'

•The deductive method would set out from known properties of mercury and known laws of the human body, and by reasoning from these would try to discover whether mercury will act on an x-afflicted body in a manner that would tend to restore health. •The experimental method would simply administer mercury in as many cases as possible, noting the age, sex, temperament, and other details of bodily constitution, the particular form or variety of x, the particular stage of its progress etc., noting in which of these cases it led to a salutary effect, and what circumstances it was combined with on those occasions. •The method of simple observation would compare instances of recovery, to find whether they agreed in having been preceded by the administration of mercury; or would compare instances of recovery with instances of failure, to find cases that were like the others except that in them mercury had been administered, or. . .that it hadn't.

§7. No-one has ever seriously contended that the last of these three methods of investigation could work with composite causes. No useful conclusions on a subject of such intricacy were ever obtained in that way. The most one could

get would be •a vague general impression for or against the efficacy of mercury, and •that would be practically useless unless it were confirmed by one of the other two methods. The results that this method tries to obtain would be of the utmost possible value *if* they could be obtained. If in an examination of a great number of instances all recoveries were cases in which mercury had been administered, we could generalise with confidence from this experience, and would have obtained a conclusion of real value. But in a case of this sort we have no chance of getting a basis for such a generalisation. Why not? because of what on page 215 I called the 'characteristic imperfection' of the Method of Agreement, namely the plurality of causes. Even if mercury does tend to cure the disease, so many other natural and artificial causes also tend to cure it that there are sure to be abundant instances of recovery in which mercury has not been administered. . . .

When an effect results from the union of many causes, no one of them can have a large role in determining •whether the effect follows or, if it does, •what it is like in detail. Recovery from a disease is an event that always comes from many influences acting together. Mercury may be one such influence; but there are bound to be cases where it is administered but the patient doesn't recover because other needed influences aren't at work, or where mercury isn't given but the patient recovers because the other favourable influences are powerful enough to do this in its absence. . . . About the best that this method could do for us is to show—by multiplied and accurate hospital records and the like—that

there are rather more recoveries and rather fewer failures when mercury is administered than when it isn't. But that result would have little value as a guide to practice, and virtually none as a contribution to the theory of the subject.¹

§8. Having recognised the inapplicability of the method of simple observation to ascertain the conditions of effects that have many concurring causes, let us now ask whether any greater benefit can be expected from the other branch of the *a posteriori* method—the one that directly *tries* different combinations of causes. . . .and takes note of their effects; e.g. trying the effect of mercury in as many different circumstances as possible. This method differs from the previous one in turning our attention directly to the causes or agents, instead of turning it to the effect, recovery from the disease. As a general rule the effects of causes are easier to study than the causes of effects, so it's natural to think that this method has a better chance of succeeding than the previous one.

The method now under consideration is called the Empirical Method; and to estimate it fairly we must take it to be *completely* empirical, without any input from any deductive operation. We might do this:

Try experiments with mercury on a healthy person in order to ascertain the general laws of its action upon the human body, and then reason from these laws to determine how it will act on persons affected with a particular disease,

and this might be a really effectual method; but it is deduction. The experimental method doesn't derive the law

¹ Bain rightly says that though the Methods of Agreement and Difference are not applicable to these cases, the Method of Concomitant Variations is of some use with them: 'If a cause happens to vary alone, the effect will also vary alone: a cause and effect may be thus singled out under the greatest complications. . . .' [Mill says that this is correct in theory, but:] when there are many influencing causes, no one of them greatly predominating over the rest, and especially when some of them are continually changing, it is scarcely ever possible to trace a relation between •the variations of the effect and •those of any one cause in a way that would enable us to assign to that cause its real share in the production of the effect.

of a complex case from the simpler laws that jointly produce it, but experiments directly on the complex case. We must set aside entirely all knowledge of the simpler tendencies of mercury in detail. Our experimentation must try to get a direct answer to the specific question: 'Does mercury tend to cure the particular disease or doesn't it?'

Let us see how far the rules of experimentation that have to be followed in other cases can be followed in the 'multiple-cause' case. [Mill's handling of this question is long, detailed and demanding, and leads him to conclude that the rules *can't* be obeyed in this 'case'. At every turn we encounter possibilities of error that we can't exclude because of the complexities of the multiple-cause situation. This sums it up:]

Anything like a scientific use of the method of experiment in these complicated cases is out of the question. We can generally, even in the most favourable cases, only discover by a series of trials that a certain cause is *very often* followed by a certain effect. Anything like a scientific use of the method of experiment is therefore out of the question in these complicated cases. Even in the most favorable cases we can generally only discover by a series of trials that a certain cause is *very often* followed by a certain effect. . . .

If so little can be done by the experimental method to understand multiple-cause situations in **medical science**, still less is this method applicable to a class of phenomena even more complicated than those of physiology—the phenomena of **politics and history**. In that region plurality of causes exists in almost boundless excess, and most effects are inextricably interwoven with one another. To make things still worse, most inquiries in political science relate to the production of very large-scale effects such as •the public wealth, •public security, •public morality and the like; and these items are open to being affected—directly or indirectly,

helped or hindered—by nearly every fact that exists or event that occurs in human society. The vulgar notion that the safe methods on political subjects are those of Baconian induction—that the true guide is not general reasoning but specific experience—will some day be referred to as one of the clearest signs of a low state of theoretical thinking in any age in which it is accepted! Nothing can be more ludicrous than the parodies of experimental reasoning that we encounter not only in popular discussion but also in solemn treatises about the affairs of nations:

- 'How can an institution be bad, when the country has prospered under it?'
- 'How can such-and-such a cause have contributed to the prosperity of one country, when another has prospered without it?'

Anyone who argues like this, not intending to deceive, should be sent back to learn the elements of some one of the easier physical sciences! Such reasoners ignore the fact of plurality of causes in the very case that provides the most obvious example of it. [Mill adds that there's little 'reason for regret' in our inability to perform experiments in this area; because even if we could perform them, we would be comprehensively defeated by the scope and complexity of the material.] The nearest approach to an *experiment*—in the philosophical [here = 'scientific'] sense of the term—in politics is the introduction of a new operative element into national affairs by some particular identifiable measure of government, such as the enactment or repeal of a particular law. But where there are so many influences at work, it takes time for the influence of a new cause on national phenomena to become apparent; and as the causes operating in such an extensive sphere are not only •infinitely numerous but •in a state of perpetual alteration, it is always certain that before the effect of the new cause becomes conspicuous enough to be a subject of

induction, many of the other influencing circumstances will have changed, wrecking the experiment.¹

Thus, two of the three possible methods for the study of phenomena resulting from a combination of many causes are, from the very nature of the case, inefficient and illusory.

That leaves us with the third—the method that considers the causes separately and infers the effect from the balance of the different tendencies that produce it; i.e. the deductive or *a priori* method. A detailed consideration of this intellectual process requires a chapter to itself.

Chapter 11. The deductive method

§1. Given that the direct methods of observation and experiment can't help us to grasp the conditions and laws of recurrence of the more complex phenomena, our main source of knowledge of those phenomena has to be the Deductive Method. It consists of three operations; •direct induction, •ratiocination, and •verification.

First operation, Induction: I call the first step in the process an 'inductive' operation, because there must be a direct induction as the basis of the whole. The role of the induction may often be played by •a prior deduction, but the premises of •this must have been derived from induction.

The problem of the Deductive Method is to find the law of an •effect from the laws of the different tendencies of which •it is the joint result. What is needed first, then, is to know the laws of those •separate• tendencies, i.e. the law of each of the concurrent causes; and this requires, for each cause

separately, a previous process of observation or experiment or else a previous deduction whose ultimate premises come from observation or experiment. If we're investigating social or historical phenomena, the premises of the Deductive Method must be the laws of the causes that determine such phenomena; and those causes are •human actions together with •the general external circumstances by which mankind are influenced. . . . So the Deductive Method as applied to social phenomena must start by investigating. . . .the laws of human action, and the properties of external things that determine the actions of human beings in society. Some of these general truths will be obtained by observation and experiment, others by deduction (e.g. deducing the more complex laws of human action from the simpler ones); but the simple or elementary laws must have been obtained by a directly inductive process.

¹ Though Bain generally agrees with the views expressed in this chapter, he seems to estimate more highly than I do the scope for specific experimental evidence in politics. He is right when he says that there are some cases 'when an agent suddenly introduced is almost instantaneously followed by some other changes, as when the announcement of a diplomatic rupture between two nations is followed the same day by a derangement of the money-market'. But this 'experiment' would be quite inconclusive merely as an experiment. It can only serve, as any experiment can, to verify the conclusion of a deduction. Unless we already knew from our knowledge of business men's motives that the prospect of war *tends* to derange the money-market, we would never have been able to prove a connection between those two facts. What if we ascertained through historical study that one followed the other in a great number of instances? Anyone who has carefully examined any of the attempts—they're continually being made—to prove economic doctrines by such a recital of instances knows very well how futile they are. It turns out that the circumstances of the cases have hardly ever been fully stated, and that the records have omitted as many or even more instances that would have tended to an opposite conclusion.

Learning what the separate causes are that must be studied in this way is sometimes hard, sometimes easy. In the social-phenomena case it is easy. There could never have been any doubt that social phenomena depend on *the acts and mental impressions of human beings*, however little may have been known about what laws govern those impressions and actions, or what social consequences their laws naturally lead to. ·Another easy case:· After physical science had achieved a certain development, there was no real doubt about where to look for the laws on which *the phenomena of life* depend: they had to be the mechanical and chemical laws of the solid and fluid substances composing the organised body and the medium in which it lives, together with the special life-laws of the various tissues constituting the organic structure. ·A hard case:· With celestial phenomena it was much less obvious in what direction the causes were to be looked for (although the relevant causal structures were far simpler than those of society and of life). What happened was this: scientists combined the laws of certain causes and discovered that those laws •explained all the facts that experience had proved concerning the heavenly motions, and •led to predictions that always turned out to be true. It was only then that mankind knew that those *were* the causes. But whether we can put the question before we can answer it (society, life) or can't state it until we have become able to answer it (celestial phenomena), either way it must be answered, and the laws of the different causes ascertained, before we can deduce from them the conditions of the effect.

The mode of ascertaining those laws is the fourfold method of experimental inquiry that I already discussed—no other way is even possible. All I need add are a few remarks on the application of that method to cases of the composition of causes.

Obviously, we can't expect to find the law of a tendency

through an induction from cases where the tendency is counteracted. The laws of motion couldn't have been brought to light from the observation of bodies kept motionless by the equilibrium of opposing forces. Even when the tendency is not •counteracted but merely •modified by having its effects combined with the effects of some other tendency, we are still not well placed to extract from the tangle the law of the tendency itself. It would hardly have been possible to discover the law that *every body in motion tends to continue moving in a straight line* by an induction from instances in which the motion is deflected into a curve by some external force. In such cases the Method of Concomitant Variations can give some help; but still the principles of a judicious experimentation prescribe that the law of each of the tendencies should be studied, if possible, in cases where that tendency operates •alone or •in combination only with agencies whose effects can be calculated and allowed for.

Accordingly, in cases where the causes can't be separated and observed apart it's *very* hard to lay down with due certainty the inductive foundation needed to support the deductive method. (It's bad luck that such cases are numerous and important.) This difficulty is especially conspicuous in the case of *physiological phenomena*, because it's seldom possible to separate the different agencies that collectively compose an organised body, without destroying the very phenomena we are trying to investigate:

—following life, in creatures we dissect,
We lose it, in the moment we detect.

(Alexander Pope)

For this reason I'm inclined to think that •physiology (greatly and rapidly progressive though it now is) is troubled by greater natural difficulties, and is probably capable of less ultimate perfection, than even •the social sciences. We can

make a better job of studying the laws and operations of one human mind apart from other minds than we can of studying the laws of one organ or tissue of the human body apart from the other organs or tissues.

Pathological facts—i.e. in common language, *diseases*—in their different forms and degrees provide physiologists with the most valuable equivalent to experimentation strictly so-called, because they often show us a definite disturbance in some one organ or organic function, the remaining organs and functions being unaffected, at least for a while. It's true that. . . there can't be a prolonged disturbance in any one function without eventually involving many of the others; and as soon as this happens the experiment loses most of its scientific value. Everything depends on observing the early stages of the disturbance, which unfortunately are bound to be the least conspicuous. But if organs and functions that aren't disturbed at first become affected in a fixed order of succession, that throws *some* light on one organ's influence over another; and we occasionally get a series of effects that we can with some confidence attribute to the original local disturbance; but to get this benefit we have to know that the original disturbance *was* local. If instead it was 'constitutional' (as they say)—i.e. if we don't know where in the animal's system it started, or exactly what it consisted of—we can't determine which of the various upsets was cause and which was effect—which of them were produced by one another, and which by the direct (perhaps delayed) action of the original cause.

. . . We can also produce pathological facts artificially. We can try *experiments*, even in the popular sense of the term, by subjecting the living being to some external agent such as the mercury of my former example or cutting a nerve so as to ascertain the functions of different parts of the nervous system. This experimentation isn't intended

to obtain a direct solution of any practical question, but rather to discover general laws from which the conditions of any particular effect can be obtained later by deduction; so it's best to select cases whose circumstances can be best ascertained; and those are usually not ones in which there's any practical object in view. The experiments are best tried not in a state of disease, which is essentially a changeable state, but in the condition of health, which is comparatively stable. In sickness, unusual agencies are at work and we can't predict their results; in health the usual course of the physiological phenomena would remain undisturbed if it weren't for the disturbing cause that we introduce.

Such are our inductive resources for ascertaining the laws of the causes considered separately when we confront them only in complexes and don't have any way of separating them out and then investigating them separately. (Actually, the Method of Concomitant Variations sometimes comes to our aid; but it is as burdened as the more elementary methods are by the special difficulties of the subject.

These resources are so glaringly inadequate that the backward state of the science of physiology is no surprise. Indeed our knowledge of causes in that science is so imperfect that it's no surprise we can't explain, and couldn't without specific experience have predicted, many of the facts that we know about from ordinary observation. Fortunately, we're much better informed than this concerning the empirical laws of the phenomena, i.e. the uniformities that we can't yet decide how to classify—whether as cases of causation or merely results of it. Not only do we have this:

The order in which the facts of organisation and life successively manifest themselves, from the first germ of existence to death, has been found to be uniform and very accurately ascertainable

but also

By a great application of the Method of Concomitant Variations to the entire facts of comparative anatomy and physiology, the characteristic organic structure corresponding to each class of functions has been determined with considerable precision.

We are quite ignorant as to whether these organic conditions are *all* the conditions—and in many cases we don't even know whether they are conditions at all rather than mere collateral effects of some common cause. And we're not likely ever to know this unless we can construct an organised body and try whether it would live.

Those are the obstacles we encounter in cases of this complex-cause sort when we try to take the first step, the inductive step, in applying the Deductive Method to complex phenomena. But fortunately things are usually not as bad as that. In general, the laws of the causes on which the effect depends can be obtained by an induction from comparatively simple instances. . . . By 'simple instances' I mean of course ones in which the action of each cause is not much intermixed or interfered with other causes whose laws we don't know. The use of the Deductive Method to ascertain the laws of a complex effect has sometimes had brilliant results, but only when the induction supplying the premises for the Deductive Method has rested on simple instances of that kind.

§2. Second operation, Ratiocination: When the laws of the causes have been ascertained, and the first stage of our great logical operation satisfactorily completed, the second part follows: determining from the laws of the separate causes what the effect will be of any given combination of those causes. This is a process of *ratiocination*, and it often involves processes of *calculation* in the narrow sense in which it = *numerical calculation*. When our knowledge of

the causes is so perfect as to extend to the exact numerical laws that they conform to in producing their effects, the ratiocination may include among its premises the theorems of the science of number—and I'm speaking of the whole immense extent of that science. We often need the most advanced truths of mathematics to be able to compute an effect of which we already know the numerical law; and even with the help of those advanced truths we can't get very far. Here's a simple problem:

Given the locations and masses of three bodies that are gravitating toward one another, with a force directly proportional to their mass and inversely proportional to the square of the distance, how do we calculate what their locations will be after n seconds?

All the resources of the calculus haven't yet been able to provide a general solution—only approximations. [And in 2012 there is *still* no complete general solution to the 'three-body problem'.] A slightly more complex case, though still one of the simplest that arise in practice, is that of plotting the motion of a projectile. Even if we know and have numerical values for all the causes that affect the velocity and range of a cannonball—the force of the gunpowder, the angle of elevation, the density of the air, the strength and direction of the wind—it's an extremely difficult mathematical problem to put these together so as to calculate their combined effect.

Besides the theorems of number, those of geometry also come in as premises, when we are trying to solve problems in mechanics, optics, acoustics, or astronomy—where the effects take place in space and involve motion and extension. But when the complication increases, and the effects depend on so many and such shifting causes that there's no place for fixed numbers or for straight lines and regular curves, the laws of number and extension are applicable, if at all, only on a large scale where precision of details becomes

unimportant. I'm thinking here of physiology, and even more of mental and social phenomena. Although mathematical laws play a conspicuous part in the most striking examples of the investigation of nature by the Deductive Method—e.g. in the Newtonian theory of the celestial motions—they are by no means an indispensable part of every such process. All that is essential is reasoning from a general law to a particular case—i.e. determining by means of the particular circumstances of that case what result is required in that instance to fulfill the law. Thus in the Torricellian experiment [in effect, the discovery of the barometer], if the fact that air has weight had been previously known it would have been easy, without any numerical data, to deduce from the general law of equilibrium that the mercury would stand in the tube at a height such that the column of mercury would exactly balance a column of the atmosphere of equal diameter. . . .

By such ratiocinations from the separate laws of the causes we can to some extent answer either of the following questions:

- Given a certain combination of causes, what effect will follow?
- What combination of causes, if it existed, would produce a given effect?

In the one case, we determine the effect to be expected in any complex circumstances of which the different elements are known; in the other case we learn according to what law—under what antecedent conditions—a given complex effect will occur.

§3. Third operation, Verification: You may want to say:

Those arguments that you used to dismiss as illusory the methods of direct observation and experiment when applied to the laws of complex phenomena—don't they apply with equal force against the Method of

Deduction? When in every single instance a multitude (often an unknown multitude) of agencies are clashing and combining, how can we be sure that in our *a priori* computation we have taken them all into account? Aren't there certain to be many that we don't know anything about? Aren't we likely to have overlooked some that we *do* know of? And even if we did take account of them all, that would be useless unless we knew the precise numerical law of each, which we usually don't. And if we did, we would need to make a calculation which, in any but very simple cases, surpasses the utmost power of mathematical science with all its most modern improvements.

These objections have real weight, and would be unanswerable if it weren't for the fact that when we are using the Deductive Method there's a *test* that enables us to judge whether we have committed any of those errors. The application of this test constitutes Verification, the third essential component part of the Deductive Method, without which all the results the method can give amount to little more than conjectures. We aren't entitled to rely on the general conclusions arrived at by deduction unless careful comparison shows us that they fit the results of direct observation wherever that can be had. If we have relevant experience and it confirms them, we may safely trust to them in other cases of which we don't *yet* have specific experience. But if our deductions have led to the conclusion that from a particular combination of causes C a given effect E would result, then in all known cases where C is not followed by E we must be able to show (or at least to make a probable surmise about) what *blocked* E; and if we can't do that the theory is imperfect and not yet to be relied upon. And the verification isn't complete unless some of the cases where the theory is confirmed by the observed result are at least as

complex as any where its application could be called for.

If direct observation and the assembling of instances have provided us with any relevant empirical laws (whether true in all observed cases, or only true for the most part), the best verification the theory could have would be its leading deductively to those empirical laws, so that the complete or incomplete uniformities that were observed among the phenomena were *accounted for* by the laws of the causes. . . . It was very reasonably thought to be an essential requirement of any true theory of the causes of the celestial motions that it should lead by deduction to Kepler's laws—which the Newtonian theory did.

Something else that is important for the verification of theories obtained by deduction. . . . is that the phenomena should be described in the most comprehensive and accurate manner possible; . . . as when the series of the observed places of a planet was first expressed by a circle, then by a system of epicycles, and subsequently by an ellipse.

Complex instances would have been no use for the discovery of the simple laws into which we ultimately analyse their phenomena, but when they have served to verify the analysis they become additional evidence for the laws themselves. Although we couldn't have discovered the law from complex cases, still when the law—discovered in some other way—is found to be in accordance with the result of a complex case, that case becomes a new experiment on the law, and helps to confirm something that it didn't help to discover. . . . This was strikingly conspicuous in the example [page 211] in which the difference between the observed and the calculated velocity of sound was found to result from the heat developed by the condensation that happens in each sound-vibration. This was a trial, in new circumstances, of the law of the development of heat by compression, and it added materially to the proof of the universality of that law. Accordingly, any

law of nature is regarded as having become more certain by being found to explain some complex case that hadn't previously been thought of in connection with it; and this indeed is a consideration that scientific inquirers customarily value too much rather than too little.

To the Deductive Method—with its three constituent parts, Induction, Ratiocination, and Verification—the human mind is indebted for its most conspicuous triumphs in the investigation of nature. To it we owe all the theories by which vast and complicated phenomena are embraced under a few simple laws that could never have been detected by the direct study of those great phenomena. To get a sense of what the method has done for us, consider the case of planetary motions. Of the greater instances of the composition of causes this is one of the simplest, because (except in a few not very important cases) each heavenly body *x* can be considered (without material inaccuracy) to be never attracted by more than two bodies at once, •the sun and •one other planet or satellite. So *x*'s motions depend on only four different agents:

- the sun,
- the other planet or satellite,
- the reaction of *x* itself, and
- the force generated by *x*'s own motion and acting in the direction of the tangent.

This is surely a much smaller number than any of the other great phenomena of nature is determined or modified by. Yet how could we ever have discovered the combination of forces on which the motions of the earth and planets depend by merely comparing the orbits or velocities of different planets, or the different velocities or positions of the same planet ·at different times·? Despite the usual regularity of those motions, and although the periodical recurrence of exactly the same effect shows that all the combinations of

causes that occur at all recur periodically, we wouldn't have known what the causes were if the existence of precisely similar agencies on our own earth hadn't brought the causes themselves within the reach of experimentation under simple circumstances. I'll have occasion in chapter 14 to analyse

this great example of the method of deduction, so I shan't spend time on it here. My next topic is a secondary application of the deductive method, the result of which is not to •prove laws of phenomena but to •explain them.

Chapter 12. Explaining laws of nature

§1. When we use the deductive operation to derive the law of an effect from the laws of the causes that jointly give rise to it, we may be engaged in either of two things: •discovering the law or •explaining a law already discovered. The word 'explanation' occurs so continually in philosophy, and has such an important place in it, that a little time spent in fixing its meaning will be well spent.

An individual fact is said to be *explained* when someone points out its cause, i.e. states the law or laws of causation of which its production is an instance. A fire is explained when it is proved to have arisen from a spark falling onto a heap of dry leaves. Similarly a law or uniformity in nature L is said to be explained when someone points out another law or laws •of which L is a special case and •from which it could be deduced.

§2. There are three sets of circumstances in which a law of causation can be explained from other laws—or, as it also is often expressed, *resolved into* other laws.

The **first** is a case that I have already fully considered: a mixture of laws producing a joint effect equal to the sum of the effects of the causes taken separately. The law of the complex effect is explained by being resolved into the

separate laws of the causes that contribute to it. For example, •the law of the motion of a planet is resolved into •the law of the acquired force that tends to produce a uniform motion in the tangent and •the law of the centripetal force that tends to produce an accelerating motion toward the sun, the real motion being a compound of those two.

In this resolution of the law L of a complex effect, the laws L is compounded of aren't the only elements. It is resolved into the laws of the separate causes *and the fact of their coexistence*. This is as essential as the other ingredients, whether we are discovering L or only explaining it. To deduce the laws of planetary motions, we have to know not only

- the law of a rectilinear force and
- the law of gravitative force, but also
- the fact that both these forces exist in the celestial regions, and even their relative amount.

The complex laws of causation are thus resolved into two distinct *kinds* of elements: •simpler laws of causation and •collocations, the latter consisting in the existence of certain agents or powers in certain places at certain times. Later on I'll need to return to this distinction and discuss it at some length—enough to remove any need to go on about it here. So: the first kind of explanation of laws of causation occurs when the law of an effect is resolved into the various

tendencies of which it is the result, together with the laws of those tendencies. . . .

§3. A **second** kind of explanation of laws occurs when what seemed to be an immediate cause-effect pair turns out to have an intermediate link, a fact caused by the antecedent and in its turn causing the consequent. A seemed to be the ·immediate· cause of C, but we later found that A was the ·immediate· cause only of B and a remote cause of C, and that B was the ·immediate· cause of C. We knew that touching an outward object caused a sensation. We discovered later that •after we have touched the object and •before we experience the sensation •some change occurs in a kind of thread called a 'nerve' that extends from our external organs to the brain. Thus, touching the object is only the remote cause of our sensation—i.e. not *the cause* properly speaking, but *the cause of the cause*—and the real cause of the sensation is the change in the nerve. Future experience may not only •increase our knowledge of the nature of this change, but also •interpolate another link. It may be (for example) that between the contact of the object with our external organs and the change in the nerve there is some electric phenomenon, or some phenomenon unlike anything we now know. No such intermediate link has been discovered up to now, so the touch of the object must be regarded *provisionally* as the immediate cause of the event in the nerve. Thus the sequence

contact with an object → sensation of touch

is discovered not to be an ultimate law; it is 'resolved' (as they say) into two other laws:

contact with an object → event in the nerve, and
event in the nerve → sensation of touch.

Another example: the more powerful acids corrode or blacken organic compounds; this is causation, but ·only· remote

causation; and it's said to be explained when it is shown that there's an intermediate link, namely an event in which chemical elements in the organic structure separate from the rest and combine with the acid. The acid causes this separation of the elements, and the separation of the elements causes the disorganisation and often the charring of the structure. . . .

§4. This is important: when a sequence of phenomena is thus resolved into other laws, they're always laws more general than itself. The law that A is followed by C is less general than either of the laws connecting B with C and A with B. Some very simple points will show that this is so.

All laws of causation are liable to be counteracted or frustrated by the non-fulfillment of some negative condition; so B's tendency to produce C may be defeated. Now the law that A produces B is equally fulfilled whether or not B is followed by C; but the law that •A produces C by means of B is of course fulfilled only when B really is followed by C; so it is less general than the law that •A produces B. It is also less general than the law that B produces C. [Mill's defence of this is the same, *mutatis mutandis*, as the one he has just given. And then he has a paragraph applying all this to the touch-nerve-sensation case, including this:] The law that an event in a nerve produces sensation is more general than the law that contact with an object produces sensation, because the sensation equally follows the change in the nerve when it is produced not by contact with an object but by some other cause. . . .

The laws of •more immediate sequence that the •law of a remote sequence is resolved into are not only more general than that law is but also more to be relied on. . . . The tendency of A to produce C can be defeated by whatever can defeat either the tendency of A to produce B or the tendency

of B to produce C; so it is twice as liable to failure as either of those more elementary tendencies; and the generalisation that A is always followed by C is twice as likely to be found erroneous. . . .

The resolution of one generalisation into two others not only •shows that there are possible failures of the former from which its two elements are exempt, but also •shows where these are to be looked for. As soon as we know that B intervenes between A and C we also know that if there are cases where $A \rightarrow C$ doesn't hold they are most likely to be found by studying the causes and the effects of B.

So we see that in the second of the three ways in which a law can be resolved into other laws, the explaining laws are more general (i.e. cover more cases) and less likely to collide with subsequent experience than is the law they explain. They are

- more nearly unconditional,
- defeated by fewer contingencies, and
- a nearer approach to universal truths of nature.

And all this is still more obviously true of first of the three modes of resolution. When the law of an effect of combined forces is resolved into the separate laws of the causes, the law of the effect must be less general than the law of any of the causes because it only holds when they are combined; whereas the law of any one of the causes holds good both in that combination and out of it. . . .

Here's another strong reason why the law of a complex effect must be less general than the laws of the causes that collaborate to produce it. If two complexes involve the same causes acting according to the same laws, they can still differ in the proportions in which the causes are combined; and that can lead to their having effects that differ not merely in •quantity but in •kind. The combination of •a centripetal force with •a projectile force, in the proportions they have

in all the planets and satellites of our solar system, gives rise to an elliptical motion; but if the ratio between the two forces were slightly different the motion they produced would be in a circle or parabola or hyperbola. . . . The law of each of the concurrent causes remains the same, however their collocations may vary; but the law of their joint effect varies with every difference in the collocations. . . .

§5. There is also a **third** mode in which laws are resolved into one another; and in this it's self-evident that they are resolved into laws more general than themselves. This third mode is the subsuming (as they say) of one law under another, i.e. the gathering up of several laws into one more general law that includes them all. The most splendid example of this occurred when •terrestrial gravity and •the central force of the solar system were brought together under •the general law of gravitation. It had already been proved that the earth and the other planets tend toward the sun; and it had been known from the earliest times that terrestrial bodies tend toward the earth. These were similar phenomena; for them both to be subsumed under one law all that was needed was to prove that as well as being alike in •quality they conform to the same rules as to •quantity. This was first shown to be true of the moon: it resembled terrestrial objects in tending to a centre and indeed in tending toward the earth. After it had been discovered that the moon's tendency toward the earth varied inversely with the square of the distance between them, it was directly calculated that if the moon were as near to the earth as terrestrial objects are, and if the acquired force in the direction of the tangent were suspended, the moon would fall toward the earth through exactly as many feet per second as those objects do by virtue of their weight. So the inference was irresistible •that the moon also tends toward the earth by virtue of its weight, and

•that these two phenomena. . . are cases of one and the same law of causation. But the tendency of the moon to the earth, and the tendency of the earth and planets to the sun, were already known to be cases of the same law of causation; and thus the law of all these tendencies and the law of terrestrial gravity were recognised as identical, and were subsumed under one general law, that of gravitation.

In a similar manner the laws of magnetic phenomena have more recently been subsumed under known laws of electricity. That's how the most general laws of nature are usually arrived at—we climb up to them by successive steps. Here is why. To arrive by correct induction at laws holding under such an immense variety of circumstances, laws so general as to be independent of any changes of space or time that we can see, requires many sets of experiments or observations, conducted at different times by different people. One set of observations teaches us that the law holds good under conditions C_1 , another that it holds good under conditions C_2 , and by combining these we find that it holds good under much more general conditions or even holds universally. The general law is literally the *sum* of all the partial ones: it recognises the same sequence in different sets of instances, and can in fact be regarded as merely one step in the process of elimination. The tendency of bodies toward one another that we now call 'gravity' was at first observed only on the earth's surface, where it shows up only as a tendency of all bodies toward the earth; it might have been regarded as a special property of the earth itself, because one of the circumstances—namely **proximity to the earth**—hadn't been eliminated. To eliminate this required a fresh set of instances in other parts of the universe; we couldn't create these ourselves, and though nature had created them for us, we weren't well-placed to observe them. The making of these observations came within the province of a different set of

scientists from those who studied terrestrial phenomena; and it was a matter of great interest back at a time when the idea of explaining •celestial facts by •terrestrial laws was regarded as the confounding of an indefeasible distinction [= 'the crossing of an uncrossable line']. But when the celestial motions were accurately ascertained and the deductions performed, showing that their laws corresponded with the laws of terrestrial gravity, those celestial observations became a set of instances that precisely eliminated the circumstance of **proximity to the earth**. This proved that in the case of terrestrial objects the cause of the downward motion or pressure was not

- the earth as such, but
- the presence of some great body within certain limits of distance,

this being the circumstance common to the terrestrial and the celestial instances.

§6. There are, then, three ways of *explaining* laws of causation, i.e. *resolving them into other laws*. **(1)** The law of an effect of combined causes is resolved into the separate laws of the causes together with the fact of their combination. **(2)** The law that connects two links (not immediate neighbours) in a chain of causation is resolved into the laws that connect each with the intermediate links. **(3)** After a law has been shown to hold good in several classes of cases we decide that what is true in each of these is true under some more general supposition consisting of what all those classes of cases have in common. The first two involve resolving one law into two or more; the third resolves two or more into one. . . .

In all three processes, laws are resolved into laws more general than themselves—laws extending to all the cases that the former extended to, and others besides. In the first two

they are also resolved into laws that are more certain—i.e. more universally true—than themselves. They aren't proved to be •laws of nature, because that would require them to be universally true; what they are proved to be is •results of laws of nature. With that status, they're only conditionally true, usually true. Not so with the third process, because here the partial laws are in fact the very same law as the general one, so any exception to them would be an exception to it too. . . .

I'm using 'explanation' in its philosophical sense. Explaining one law of nature by another is merely substituting one mystery for another, and does nothing to make the general course of nature other than mysterious. We can no more assign a *why* for the more extensive laws than for the partial ones. In ordinary talk about these matters, an 'explanation' replaces a mystery that is still strange by one that has become familiar and come to seem not mysterious. But the process I have been discussing here often does the exact opposite: it resolves •a phenomenon that we are familiar with into •one of which we previously knew little or nothing; as when the common fact that heavy bodies fall was resolved into the tendency of all particles of matter toward one another. Don't forget this: in science those who speak

of 'explaining' a phenomenon mean (or should mean) to be pointing out not some more *familiar* phenomenon but merely

- some more *general* phenomenon of which it is a partial exemplification, or
- some laws of causation that produce it by their combined action, and from which its conditions can therefore be determined deductively.

Every such operation brings us a step nearer toward answering the question that I said on page 157 includes the whole problem of the investigation of nature, namely: What are the fewest and simplest assumptions, which being granted the whole existing order of nature would result? What are the fewest general propositions from which all the uniformities existing in nature could be deduced?

. . . .In minds that aren't used to accurate thinking there is often a confused notion that the general laws are the causes of the partial ones, e.g. that the law of general gravitation causes the phenomenon of the fall of bodies to the earth. But that's a misuse of the word 'cause'; terrestrial gravity isn't an •effect of general gravitation but a •case of it, i.e. one kind of the particular instances in which that general law obtains. . . .