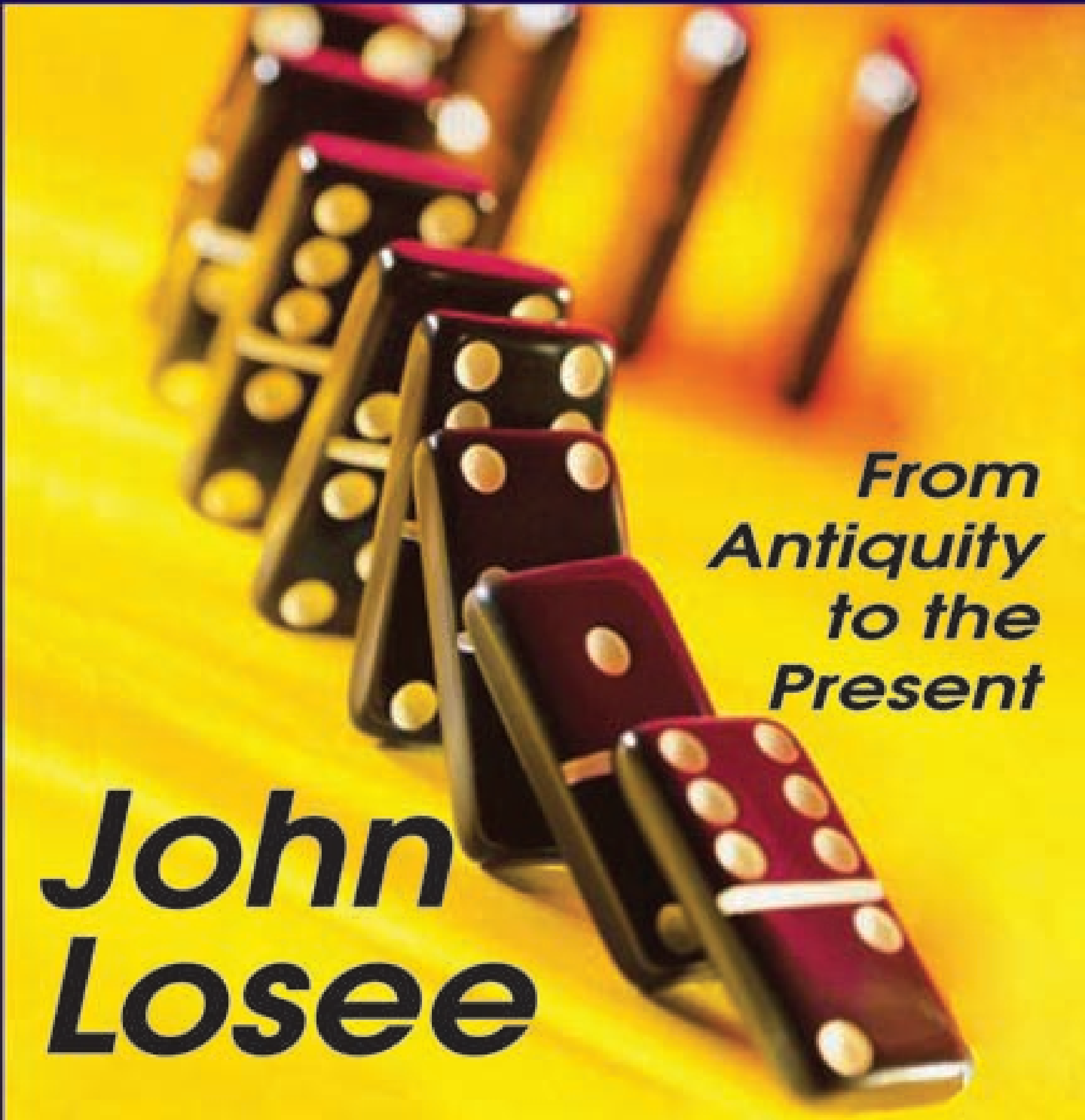
THEORIES of CAUSALITY



From Antiquity to the Present

John Losee

THEORIES of CAUSALITY

THEORIES of CAUSALITY

From Antiquity to the Present

John Losee

Transaction Publishers
New Brunswick (U.S.A.) and London (U.K.)

This book is printed on acid-free paper that meets the American National Standard for Permanence of Paper for Printed Library Materials.

Library of Congress Catalog Number: 2011010279
ISBN: 978-1-4128-1832-2
Printed in the United States of America

Library of Congress Cataloging-in-Publication Data

Losee, John.
Theories of causality : from antiquity to the present / John Losee.
p. cm.
ISBN 978-1-4128-1832-2 1. Causation. I. Title.
BD531.L67 2011
122—dc22

2011010279

Contents

III Quantum Mechanics and the Regularity-between-States View

IV Protests against the Identification of Causality and Regularity

Introduction

1. What types of entities qualify as "causes" and "effects"?
2. What is the relationship between cause and effect?
3. How are causal claims to be assessed?

This volume is a selective history of answers that have been given to the above three questions, augmented by occasional evaluative comments by the author. The first question is about the structure of the world. Answers given to this question assert that things in the world stand thus and so, and not otherwise. To ascribe "causation" to states of affairs is to make an ontological claim. For instance, one may claim that event e_1 causes event e_2 in a specific case or that all events of type *c* cause events of type *e*.

The second question is about theories that interpret causal relatedness. Some theories of causality take the relationship to be logical: *c* is a necessary condition of *e*, or *c* is a sufficient condition of *e*, or *c* is an INUS condition of *e* (*c* is an *i*nsufficient but *n*ecessary member of a set of conditions, which set is *u*nnecessary but *s*ufficient for *e*). Other theories take the relationship to be empirical: *c* and *e* are members of classes of events that display constant sequential conjunction, there is energy or momentum transfer between *c* and *e*, or there is exchange of a conserved quantity from *c* to *e*.

The third question is about proper procedure in science and everyday life. The usual procedure is to demonstrate that the claim in question satisfies the requirements of one or more theory of causality. In addition, one may support a claim by showing that it is consistent with established scientific laws and that it is superior to other possible claimants to causal status.

I have sought to present theories of causality within a historical survey that emphasizes the interrelationship between these theories and developments in science. I hope that my analysis displays the strengths and the weaknesses of these theories in such a way as to contribute to our present understanding of causal relatedness.

Some disclaimers are in order. The subject matter of this study is *physical causation*, and the history of theories advanced to interpret it. Psychical states may cause, and be caused by, external events. However, mental causation is not a principal focus here.

I have relied on present ideas about causal relatedness in the selection of subject matter for this volume. This reliance leads to some exclusions, despite the fact that the excluded theories were regarded as important at one time. In particular, I have little to say about appeals to God as cause of events, miraculous or otherwise. And I ignore the once influential preoccupations with causation by astral forces or alchemical forces.

I

Classical Sources: Theories of Causality Prior to 1900

1

Aristotle on the Four Aspects of Causation

Aristotle inaugurated discussion about the nature of causation. He selected "processes" as subject matter for causal inquiry. A process, for Aristotle, is a transition in which that which is potential is transformed into that which is actual.

According to Aristotle, there are four aspects of such transitions. A complete causal analysis specifies

1. the form of the process (*formal* cause),
2. the matter transformed (*material* cause),
3. the interaction between the transforming agent and that which is transformed (*efficient* cause), and
4. the *telos*, or purpose, of the process (*final* cause).

He wrote that

> it is the job of the natural scientist . . . to understand all four of these causes; if he refers the question "Why?" to this set of four causes—matter, form, source of change, purpose—he will be explaining things as a natural scientist should.[1]

To explain the construction of a house one needs to specify

1. a plan or blueprint,
2. construction materials such as bricks, mortar, etc.,
3. the activities of bricklayers, roofers, etc., and
4. the goal of providing a suitable shelter from the elements.

To explain the emission of inky fluid by a cuttlefish, one needs to specify

1. a correlation between movement in the vicinity of the cuttlefish and the emission of ink,
2. the material composition of the ink (Aristotle believed that it contained "earthy matter . . . like the white deposit on the excrement of birds"[2]),

3. the triggering of the ink-release response by external motions *qua* perceived by the cuttlefish, and
4. the purpose of self-defense and preservation.[3]

Aristotle extended this emphasis on final causes to inanimate motions that are "natural" to the objects in motion. For instance, heavy bodies fall to the earth in order to reach their "natural place," and fire rises in order to reach its "natural place," the region just inside the sphere of the moon's orbit. Aristotle's view of the universe is irreducibly teleological, and this is reflected in his insistence on the importance of specifying final causes for all processes.

Aristotle was critical of Pythagorean and Atomist views of causation. He complained that Pythagoreans were preoccupied with the formal aspect of actualizing potentiality and that Atomists were preoccupied with material and efficient causes. Aristotle insisted that it is important that all four aspects of causation be addressed, but he allowed that in some cases different aspects of causation coincide.[4] In his own scientific investigations, he often specified just one or two of these four aspects of causation.

Aristotle acknowledged that diverse types of entities qualify as "causes." Formal causes are patterns that correlate properties or events. For instance, chunks of ice placed in water float. The formal cause of this process is the pattern that correlates immersion and flotation. Material causes are substances. Efficient causes are activities of "agents." They involve production. An efficient cause produces its effect. Aristotle insisted that a transition from potentiality to actuality is always inaugurated by a prior actuality. The "agent" may be external (e.g., the horse whose movement propels a cart) or internal (e.g., the successive states of a tadpole that lead to its actualization as a frog). Final causes are future states of affairs (e.g., the mature frog is the final cause of the development of one tadpole stage to another).

In general, the *telos* of a process does not involve conscious deliberation. Only in the case of human volitional activity is such deliberation present. Nevertheless, the *telos* of a process is a future state that somehow determines present developments. How can this be justified? Aristotle believed that the uniform circular motions of the stars and planets are divine. He believed, as well, that there is a terrestrial process that imitates these divine motions. This process is the cycle of growth and reproduction that underlies the continuity of a species. Individuals perish. But their participation in the cycle of growth and reproduction guarantees the eternal status of their species.

Given this commitment to the pseudo-divinity of species, it seems plausible to maintain that the final cause of individual development is reproduction that contributes to perpetuation of the species. On this point of view, a future state is a cause of present development.

Aristotle stressed the interdependence of causal analysis and scientific explanation. He held that scientific explanation is achieved when there is a gain in knowledge from knowledge of a fact to knowledge of the reason why the fact is what it is. Aristotle required that the fact be expressed as the conclusion of a sound deductive argument whose premises state the cause of the attribution made in the conclusion.

Consider the fact that all oxen are animals with missing upper incisor teeth. The explanation that states the cause of this fact is

> All ruminants with four-chambered stomachs are animals with missing upper incisor teeth.
>
> All oxen are ruminants with four-chambered stomachs.
>
> ---
>
> ∴ All oxen are animals with missing upper incisor teeth.

Ruminants can store partially digested food in one stomach chamber and return it to the mouth for further chewing. Consequently, they do not need, and do not have, upper incisor teeth. Moreover, the bony material that otherwise would have constituted the missing teeth instead has formed horns useful for self-defense.[5]

The requirement that the premises of a scientific explanation state the cause of the attribution asserted in its conclusion is important. Aristotle noted that otherwise arguments that cite merely accidental correlations would have explanatory force.[6] Consider the following alternative "explanation" of the missing incisor teeth:

> All ruminants with cloven hoofs are animals with missing upper incisor teeth.
>
> All oxen are ruminants with cloven hoofs.
>
> ---
>
> ∴ All oxen are animals with missing upper incisor teeth.

Since there is no causal dependence of food processing on hoof shape, this argument fails to qualify as a scientific explanation.

Aristotle's preferred "causal" explanation of the missing upper incisor teeth is not convincing. But he deserves credit for calling attention to the problem of distinguishing causal correlations from merely accidental correlations.

Unfortunately, he was unable to stipulate a criterion to distinguish causal correlations from accidental correlations. The best he could do was to insist that causal correlations be true of every member of the subject class and that the predicate term be "essential" to being a member of the subject class. Aristotle provided examples of "essential" correlations—being an animal is essential to being a man, and slitting the throat of an animal is essentially related to its death.[7] But citing examples is not what is called for. What is needed is a criterion to distinguish essential from nonessential correlations.

Notes

1. Aristotle, *Physics II*, 7 a 21.
2. Aristotle, *Parts of Animals IV*, 679 a 20.
3. Ibid., 679 a 30.
4. Aristotle, *Physics II*, 7 a 24.
5. Aristotle, *Parts of Animals III*, 664 a.
6. Aristotle, *Posterior Analytics I*, 13, 78 a–b.
7. Ibid., 73 a 25–73 b 15.

2

Medieval Science and the Discovery of Causes

After the fall of the Roman Empire, Aristotle's writings on causal explanation were unavailable to Western European scholars until translations became available in the twelfth century. Aristotle's works had been preserved by scholars in the Arab World, either in the original Greek or in translations into Arabic. They became known to Western scholars in the twelfth century as a consequence of Arabic penetration into Spain and Sicily.[1] The impact of this newly rediscovered knowledge was immense. For the next two centuries, most writings on the sciences took the form of exposition and commentary on Aristotle's works.

One original development during the medieval period was the recommendation of procedures designed to uncover causes. Aristotle had championed two inductive procedures—simple enumeration and "intuitive induction."

Simple enumeration is straightforward generalization. One examines supposedly similar instances. From premises about what is observed to be true of individuals, one concludes that what is observed also is true of the species (or set) to which the individuals belong. For example, several stones are observed to sink when placed in water. One concludes by simple enumeration that all stones sink when placed in water. The generalization, if true, increases our knowledge of a causal process.

Conclusions reached by simple enumeration are at the mercy of a single exception. It is false that all swans are white, even though application of simple enumeration to swans observed in Europe prior to 1750 supported that generalization.

Intuitive induction is more creative. It really is not a procedure at all. It is a matter of achieving insight into what is "essential" in a situation. Aristotle provided an excellent example. An observer notices the changing shadow covering the moon during a lunar eclipse and concludes that the earth is a sphere.[2]

Aristotle was an accomplished taxonomist. He clearly distinguished cetaceans (whales, dolphins, etc.) from fish and described some 540 species, including a species of placental dogfish that was reconfirmed only in the nineteenth century. The scientist who achieves intuitive insight resembles the trained naturalist who "sees" the characteristics of an individual that are essential for its classification.

Medieval theorists agreed with Aristotle that causal relations are discovered by induction. They advanced some methodologically sophisticated additional inductive procedures. Unfortunately, they did not increase our knowledge of causal relations by means of specific applications of these procedures.

Robert Grosseteste (ca. 1168–1253) asked how it could be determined whether a particular herb caused purgation. Causal status could be established, he suggested, if one administered the herb in numerous cases for which no other purgative agents are present, and purgation occurred in each case.[3] Grosseteste's suggestion, taken as an inductive principle, is an anticipation of what came to be known as the joint method of agreement and difference. Let *h* be the herb in question, *p* be purgation, and *A*, *B*, *C*, . . . be other circumstances present when *h* is administered. Grosseteste's procedure then may be represented as

Instance	**Circumstances**	**Phenomena**
1	$h\ A\ B$	P
2	$h\ B\ C$	P
3	$h\ D\ E$	P

and

$h\ A\ B - p$	$h\ B\ C - p$	$h\ D\ E - p$
$A\ B$— —	$B\ C$— —	$D\ E$— —

Grosseteste did not claim to have conducted the trials that would constitute evidence for the claim that *h* causes *p*.

John Duns Scotus (1265–1308) subsequently formulated a version of the inductive method of agreement. Interestingly, he took the conclusions of arguments that satisfy this inductive schema to assert only the *possibility* of causal dependence.[4] His schema for the method of agreement is

Instance	**Circumstances**	**Phenomena**
1	*A B C D*	*p*
2	*A B E*	*p*
3	*A C E F*	*p*
4	*A D F*	*p*

∴ *p can be* the effect of *A*.

Applied to Grosseteste's example, the appropriate conclusion is that purgation *can be* the effect of administering herb *A*.

Duns Scotus was concerned to emphasize the omnipotence of God. He believed that God could accomplish anything that does not involve self-contradiction. Since God at any time could prevent *A* from exercising its purgative effect, possibility is the most that can be established by an instantiation of the agreement schema. Duns Scotus developed a useful inductive procedure for gaining knowledge of causal relations while concurrently insisting that such relations exist only by God's forbearance. The scientist's motivation to discover causal relations is undercut by this emphasis on divine omnipotence.

William of Ockham (1280–1349) shared Duns Scotus's concern to emphasize the omnipotence of God. He formulated an inductive method of difference, the conclusion of which asserts only the possibility of causal relatedness, viz.[5]:

Instance	**Circumstances**	**Phenomena**
1	*A B C*	—
2	*B C*	—

∴ *A can be* the cause of *p*.

Ockham called attention to the requirement that the schema list every circumstance that is causally relevant to the occurrence of *p*. If there is some circumstance *X* that is present in instance 1 and absent in instance 2, it, and not *A*, may be the cause of *p*. But it is not the possibility that some unnoticed circumstance *X* exists that discourages empirical inquiry. After all, one can multiply cases to minimize the chance that some circumstance other than *A* is the cause of *p*. What does discourage empirical inquiry is preoccupation with the notion that God at any time can suspend a causal relation by preventing *p* despite the occurrence of *A*.

Notes

1. See, for instance, A. C. Crombie, *Medieval and Early Modern Science* (Garden City, NY: Doubleday Anchor, 1959), vol. 1, 33–64.
2. Aristotle, *De Caelo II*, 14, 297 b.
3. A. C. Crombie, *Robert Grosseteste and the Origins of Experimental Science* (Oxford: Clarendon Press, 1953), 73–74.
4. Duns Scotus, *Philosophical Writings*, trans. and ed. Allan Wolter (Edinburgh: Thomas Nelson, 1962), 109.
5. William of Ockham, *Libros Sententiarum II*, Q. 1 D.

3

Francis Bacon on the Exclusion of Final Causes

Aristotle's concept of "final cause" came under attack in the sixteenth century. Francis Bacon (1561–1626) led the attack against this notion.[1] Indeed, Bacon was the lead propagandist against many aspects of the Aristotelian methodology for science.

According to Bacon, the scientist ought to approach the study of nature as would Adam at his creation. The scientist ought to be unencumbered by predispositions and prejudices. Bacon identified four classes of "idols" that cloud men's minds: idols of the tribe (e.g., reading more regularity into nature than is present), idols of the cave (e.g., glossing over differences in order to protect generalizations), idols of the market place (e.g., blurring of the meanings of scientific terms in common usage), and idols of the theatre (e.g., distortions introduced upon acceptance of philosophical dogmas, in particular Aristotelianism). The assumption that there is some purpose, or *telos*, that guides physical processes is an especially pernicious predisposition. Bacon emphasized that progress in science depends on setting aside such conceits.

Bacon also criticized Aristotle's position on induction. He pointed out that the method of simple enumeration is ineffective since a single exception suffices to falsify a conclusion reached by applying this method. Bacon recommended that simple enumeration be replaced by a tabulation of instances of presence, absence, and degree of entities that may be causally related. Inspection of such tables supposedly is effective in excluding those accidental correlations that fail to hold in all cases.

Bacon declared that

> the first task of true induction (as far as discovering forms is concerned) is the *rejection or exclusion* of all the separate natures that are not found in any instance where the given nature is present, or

> are found in any instance where the given nature is absent, or are found to increase in any instance when the given nature decreases, or to decrease when the given nature increases.[2]

He applied this procedure to see what is correlated with the phenomenon of heat. Heat, he concluded, is "an expansive motion, checked, and exerting itself through the smaller parts of bodies."[3] In his presentation of this supposedly new method of "rejection or exclusion," Bacon failed to acknowledge the contributions of Grosseteste, Duns Scotus, and William of Ockham.

According to Bacon, Aristotle and his followers also were guilty of the methodological sin of hasty generalization. They accumulated a few observations and then leaped to the most general conclusions. The remedy, Bacon maintained, is to make a stepwise progressive ascent from low-level correlations to the most inclusive generalizations. He referred to those generalizations at the top of the inductive pyramid as "forms." He declared that

> after the rejection and exclusion has been properly made . . . there will remain . . . the affirmative form, solid, true, and well-defined.[4]

Just what Bacon meant by "forms" is difficult to judge. At one point, he identified "form" and "law." He maintained that a form is what "determines the natures of substances."[5] Forms are applicable to diverse types of entities. He spoke of "forms" of yellowness, weight, and ductility. Bacon was lured by the promise of alchemical transmutation. He held that if one could achieve knowledge of the appropriate forms, one could transform a base metal into gold. Evidently, to understand a form is to understand causal agency. Causal agency, he consistently maintained, includes formal, material, and efficient aspects, but not purposeful development.

In Book II of *Novum Organum*, Bacon introduced a set of "prerogative instances," which reveal causal correlations directly in the absence of tables of presence, absence, and degree. The most famous is the "Instance of the Fingerpost" or "crucial experiment." An experiment is crucial if there are just two possible theories of a process and the experimental result is consistent with one but not the other.

Bacon outlined an experiment that supposedly is crucial in favor of

hypothesis #1—the free fall of a body is caused by attraction exerted by the earth and fatal to hypothesis #2—the free fall of a body is caused by an intrinsic tendency to reach its "natural place," the center of the earth. The experiment is to compare the rates of a weight-driven clock and a spring-driven clock at different heights from the earth's center.

An experiment can be "crucial" only if there are just two possible competing explanatory theories. Bacon was aware of this. He declared that

> it must be that heavy and weighty objects either tend by their own nature towards the centre of the earth . . . or they are attracted by the bodily mass of the earth itself.[6]

Given that this is correct,

> if the power of the weights [relative to the spring-driven clock] is found to be diminished when the clock is high up and increased when it is underground, then we may take it that attraction by the bodily mass of the earth is the cause of weight.[7]

Bacon did not claim to have performed an experiment of this kind.

Notes

1. Francis Bacon, *Novum Organum*, Bk. II, Aph. 2.
2. Ibid., Bk. II, Aph. 16.
3. Ibid., Bk. II, Aph. 20.
4. Ibid., Bk. II, Aph. 16.
5. Ibid., Bk. II, Aph. 4.
6. Ibid., Bk. II, Aph. 36, No. 5.
7. Ibid., Bk. II, Aph. 36, No. 5.

4

The Revival of Atomism

Aristotle had criticized the atomism of Democritus and Leucippus for focusing exclusively on material and efficient causes. He also might have complained that it was a mere "picture preference" to be superimposed on phenomena after the fact. Given that salt dissolves in water and sand does not, one can superimpose a picture of salt atoms fitting between water atoms, and sand atoms failing to do so. But given an untried substance, the theorist has no way to predict accurately whether it will dissolve.

Nevertheless, there were features of classical atomism that made it an attractive alternative to the Aristotelian position in the sixteenth and seventeenth centuries. The atomist position is that microphenomena cause macrophenomena. The microphenomenal realm contains atoms, their combinations, and interactions among them. Atoms differ with respect to size, shape, and motion. They supposedly are indivisible and impenetrable. Implicit in atomic theory is a distinction between primary qualities—size, shape, and motion—and secondary qualities—colors, odors, and tastes. The atomist position is that secondary qualities are caused by primary qualities. Qualitative changes at the macrophenomenal level are to be explained by citing quantitative changes at the microphenomenal level. The atomic theory presents a picture of a "real world" that differs markedly from the world we experience.

There is no role for final causes in classical atomic theory. The motions, association, and dissociation of atoms are the sole causes of the processes we experience. To many sixteenth- and seventeenth-century thinkers, the study of efficient causation is the proper focus of a scientific study of nature.

Pierre Gassendi (1592–1655) was an influential champion of the atomic theory and its emphasis on efficient causes. He superimposed on various processes pictures of atoms moving and combining—sometimes endowing atoms with hooks and eyes to facilitate union.

Gassendi realized that the uncompromising materialism of atomic theory rendered it anathema within Christian theological circles. From the dominant religious perspective of the time, it was not only absurd to reduce spiritual values to the concourse of atoms, it also was heretical. Gassendi tried to make atomism theologically respectable by emphasizing that atoms were created by God, and hence are not eternal, and that they received motions initially from God at creation.

Gassendi managed to promote atomism without running afoul of religious authorities. Galileo (1564–1642) was not so fortunate. In *Assayer* Galileo hinted that light and heat are composed of atom-like particles.[1] In *Two New Sciences* he put forward a theory that attributes the internal cohesion of a solid to the presence of interstitial gaps between its particles. The particles, or atoms, of the solid exhibit a "horror of the vacuum," which ensures that they remain bonded together as a solid object.[2]

On occasion, Galileo seemed to support the classical atomic theory of indivisible minima. But he also discussed "infinitesimal" particles subject to indefinite subdivision, separated by infinitesimal vacuums. On either interpretation, it is the motion of these "particles" that gives rise to the observed properties of the objects of our experience.

Pietro Redondi has argued that Galileo's difficulties with the inquisition arose not only because of his support for a heliocentric theory of the universe, but also because of his presumed commitment to some form of the atomic theory.[3] The atomic theory was believed by those in power to have consequences inconsistent with Catholic teaching on the Eucharist. According to the Council of Trent, there remain in the host only the accidents of shape, texture, color, and taste. Its substance has been transformed miraculously into the body of Christ. On the atomic theory, however, the above accidents are the effects of atomic motions and have no existence apart from these motions. The substance-accidents view of matter fits the transubstantiation claimed for the Eucharist. The atomic theory does not.

Galileo sometimes has been accused of shifting attention from "why" questions involving causation to "how" questions. In part, the accusation is based on extrapolation from his successful theory of falling bodies. Galileo restricted his theory to the kinematics of the motion, viz., relations between distance, velocity, and acceleration. He established that free fall conforms to a "law of odd numbers."[4] In successive equal temporal intervals, a falling body covers 1, 3, 5, 7, . . . intervals of space. In modern terminology, the relation is expressed as $s = \frac{1}{2}gt^2$. The law of odd numbers "saves the appearances" of free fall without reference to any cause of the motion. Galileo achieved this success in *Dialogue Concerning the Two Chief World Systems* after abandoning his earlier dynamical theories based on the causal impact of Archimedean buoyant forces.[5]

Galileo did address the question about the nature of causation, however. He suggested that a cause is

> that which is always present when the effect is seen, and in whose absence the effect does not take place.[6]

He thus maintained that a cause is a condition both sufficient and necessary for the occurrence of an effect. It is a sufficient condition because whenever an instance of the effect occurs, an instance of the cause is present, and it is a necessary condition because if no instance of the cause is present, no instance of the effect occurs. Galileo sought such causes in his scientific research and was willing to speculate about possible causes in cases for which he had no good evidence to support such universal correlations.

Notes

1. Galileo, "The Assayer," trans. S. Drake, in *The Controversy on the Comets of 1618* (Philadelphia: University of Pennsylvania Press, 1960), 312–13.
2. Galileo, *Dialogues Concerning Two New Sciences*, trans. H. Crew and A. De Salvio (New York: Dover Publications, 1914), 19.
3. Pietro Redondi, *Galileo Heretic* (Princeton, NJ: Princeton University Press, 1987).
4. Galileo, *Dialogue Concerning the Two Chief World Systems*, trans. S. Drake (Berkeley: University of California Press, 1953), 221–22.
5. Galileo, *De Motu* (1590), trans. I. E. Drabkin, in Galileo Galilei, *On Motion and On Mechanics* (Madison: University of Wisconsin Press, 1960), 26–38.
6. Galileo, *Bodies that Stay Atop Water, or Move in It*, trans. S. Drake, in Stillman Drake, *Cause, Experiment and Science* (Chicago: University of Chicago Press, 1981), 71.

5

Causes, Impacts, and Action-at-a-Distance

Descartes on Causation-by–Contact

In the seventeenth century, René Descartes (1596–1650) became the acknowledged leader of the attack on final causes. He insisted that all causation is efficient causation. According to Descartes, appeals to "final causes" are unscientific, a throwback to an earlier age of superstition.

Descartes maintained, in addition, that an efficient cause can produce an effect only by contact action. Paradigm cases, familiar from everyday experience, are the collisions of billiard balls and the closing of doors. A commitment to this position led many Cartesian natural philosophers to downplay the importance of Newton's theory of universal gravitational attraction. They held that the Newtonian theory is a useful mathematical device that can be applied to "save the appearances" of the motions of falling bodies, pendulums, and the planets, but that it fails to state the cause of these motions. From the Cartesian perspective, Newtonian accounts in terms of gravitational attraction violate the requirement that all efficient causes produce their effects through action-by-contact.

Descartes sought to derive general laws applicable to the universe from certain indubitable truths about God and man. He reasoned that a Perfect Being who created the universe "all-at-once" would ensure that the motion he injected into the universe would be conserved indefinitely. Unfortunately, Descartes assumed that the "motion" thus conserved is measured by the product of the size and speed of bodies. He derived seven "rules of impact" from this conservation principle and assumptions about the relation between force and resistance. These rules of impact are all incorrect. Rule 4, for example, states that a body in motion cannot move a stationary second body upon impact if that second body has a greater size.

Descartes's rules of impact soon were corrected by Huygens. Huygens argued that the appropriate conservation principle refers to momentum (mass times velocity) rather than "motion."[1] Thus amended, the Cartesian view of motion won many adherents in the late seventeenth and early eighteenth centuries. Natural philosophers could accept Descartes's basic principles—the conservation of motion (momentum), the impossibility of a vacuum, the denial of action-at-a-distance, etc.—without committing themselves to Descartes's supposed derivation of these principles from "self-evident truths" about God and man.

Descartes's conviction that the cause of motion is always the impact or pressure exerted by a contiguous body (or bodies) was widely shared. However, there are some motions that appear to take place without the benefit of impact or pressure. Important examples are magnetic attraction and the motions of the planets. Descartes observed that

> the principles that we have already discovered, are so vast that many more things follow from them than we see included in this visible universe.[2]

The basic physical principles only place general constraints on what can happen in the universe. A specific motion, for instance, must obey the conservation law, the requirement of action-by-contact, the prohibition against the occurrence of a vacuum, and other requirements. Descartes noted that in order to account for a *specific* motion it often is necessary to formulate a hypothesis. The hypothesis must be consistent with the general laws of nature and, together with these laws and statements about relevant conditions, must imply a statement about the phenomenon in question.

Descartes himself offered an imaginative hypothesis to account for magnetic attraction. He hypothesized that a magnet exudes a stream of tiny, invisible, screw-shaped particles from one of its poles. These particles move in a closed loop around to the other pole. A magnetically susceptible object, such as an iron nail, placed in the vicinity of the magnet moves toward the magnet. The cause of this motion is the lateral pressure produced by the passage of the screw-shaped particles through screw-shaped channels in the nail.

Descartes's causal explanation is fanciful, but it is consistent with his basic principles. On these principles, the alternative explanation in terms of an attractive force exerted by the magnet is unscientific. Such a force would be on a par with the "astral influences" invented by astrologers. Forces cannot act at a distance. Motion is communicated solely by impact or pressure.

The motions of the planets were a major puzzle for Descartes and his followers. All motion must be initiated by contact, but there are no visible bodies pushing on the planets. Descartes hypothesized that a swirling vortex of invisible aether particles carries the planets in their orbits around the sun. The solar system resembles a collection of corks carried at various distances around the center of a whirlpool.

Isaac Newton subsequently proved that a single, rotating vortex would impart motions to the planets that violate Kepler's third law. In response, Leibniz, Malebranche, de Molieres, and others developed complex theories of the solar system that employ multiple vortices.[3] Leibniz, for instance, suggested that the orbits of the planets are caused by the combined action of two intersecting vortices—a "solar vortex" and a "harmonic vortex." The vortices were required to act by contact on the planets to produce elliptical orbits around the sun. At the same time, the vortices were assumed to pass through one another without any interaction.[4]

An even more extreme illustration of the strength of the Cartesian belief that causal action can be achieved only by contact action is Christiaan Huygens's explanation of free fall. To account for the effects of terrestrial gravity, Huygens postulated the existence of a separate vortex around the earth (in addition to the solar vortex). He calculated that if the particles of this vortex rotate around the earth's center with a velocity seventeen times greater than the rate of rotation of the earth itself, then the observed rate of fall of bodies (32 ft./sec^2) will be produced. The fall of a body is caused by contact action with particles of the vortex. The centrifugal tendency of a vortex particle to fly off on a tangent to its motion increases with increasing distance from the earth. Thus, a particle in contact with the upper surface of a falling body has a greater centrifugal tendency than does a particle in contact with its lower surface. Huygens took these centrifugal tendencies to be "forces." He maintained that there is a net downward force on the body. As it descends toward the earth, the space formerly occupied by the body is filled by equivalent volumes of aether of progressively decreasing centrifugal tendency.

So far, so good. But bodies fall at all points on the earth's surface—at the poles as well as at the equator. Huygens recognized that a single vortex rotating around the earth can account for free fall only in one plane through the earth's center. He therefore hypothesized that there are indefinitely many vortices rotating in all planes that pass through the center of the earth. Huygens required that the rotations of these rings of aethereal particles take place without interfering with one another.

Huygens's theory may be ugly, but it does attribute the cause of free fall to contact action. Moreover, Huygens could point to a small-scale example of vortex action in a single plane. This was the famous "Spanish-wax experiment."[5] Huygens placed bits of Spanish wax, a substance more dense than water, in a rotating cylinder of water. The wax bits soon moved to the outer surface of the cylinder. Huygens then stopped the rotation of the cylinder. The water continued to rotate for a time within the vessel. The wax bits spiraled in toward the axis of the cylinder. Huygens cited this result as an indication of how a rotating fluid vortex could produce a centripetal acceleration on heavier bodies contained within it.

For the committed Cartesian, the Newtonian alternative that postulates a "gravitational force" operating at a distance fails to provide a cause for the motion of a falling body. To say that "a body is accelerated toward the earth because of a gravitational force" is to say no more than "a body is accelerated toward the earth."

Newton on Gravitational Attraction and Causation

In *Queries* appended to *Opticks*, Newton declared that

> To tell us that every Species of Things is endow'd with an occult specifick Quality by which it acts and produces manifest Effects, is to tell us nothing: But to derive two or three general Principles of Motion from Phaenomena, and afterwards to tell us how the Properties and Actions of all corporeal Things follow from these manifest Principles, would be a very great step in Philosophy, though the Cause of those Principles were not yet discover'd.[6]

Newton himself had provided a most persuasive application of this program. He accounted for motions in the solar system by attributing to each body a $1/R^2$ attractive force directed toward the center of mass of the system.

In the above passage, Newton clearly distinguished principles of motion, such as universal gravitational attraction, from the underlying causes of motion. When questioned about these underlying causes, Newton sometimes suggested that there may exist an all-pervasive aether that is responsible for gravitational attraction. In *Queries* he wrote

> Is not this Medium much rarer within the dense Bodies of the Sun, Stars, Planets and Comets, than in the empty celestial Spaces between them? And in passing from them to great distances, doth it not grow denser and denser perpetually, and thereby cause the gravity of those great Bodies towards one another, and of their parts towards the Bodies; every Body endeavouring to go from the denser parts of the Medium towards the rarer?[7]

Newton conceded that such a medium could cause gravitational attraction only if it were both extremely rare and extremely elastic, qualities that appear to be antithetical.

Newton did not express a firm commitment to the existence of an extremely elastic aether of negligible density. Rather, he presented the notion as a "query" to stimulate further inquiry. But if no such aether exists, what are the alternatives?

One alternative is to affirm action-at-a-distance. Newton declared that this is not a viable option:

> that gravity be innate, inherent, and essential to matter, so that one body may act upon another at a distance through a vacuum, without the mediation of any thing else, by and through which their action and force may be conveyed from one to another, is to me so great an absurdity, that I believe no man, who has in philosophical matters a competent faculty of thinking, can ever fall into it.[8]

A second alternative is to postulate an immaterial, spiritual cause of gravitational attraction. Newton was tempted by this alternative, but when pressed, usually elected to defend a third alternative—agnosticism about the cause of gravitational attraction.

It is not surprising that Newton's supporters were uncertain about his position. Richard Bentley, for example, suggested that Newton took gravitational attraction to be an inherent property of matter. Newton replied

> Pray do not ascribe that Notion to me; for the Cause of Gravity is what I do not pretend to know.[9]

Bentley accepted Newton's correction. It emboldened him to put forward an argument for a spiritual cause of gravitational attraction:[10]

1. One body can act upon another body only by contact (as Descartes and his followers recognized).
2. It is a matter of fact that bodies do act upon one another at a distance.

> Therefore there must be a spiritual cause of the action-at-a-distance that is gravitational attraction.

Despite his protestations of agnosticism, Newton was favorably disposed toward this argument. He spoke in *Principia* of

> a certain most subtle spirit which pervades and lies hid in all gross bodies; by the force and action of which spirit the particles of bodies attract one another.[11]

The LaPlacian Program

Pierre-Simon de LaPlace, by contrast, excluded spiritual forces from science. Neither was he willing to restrict causes to impacts and pressures. Given that the motion of bodies is not due to an all-pervasive aether and not due to spiritual forces, the distinction between these motions and their underlying causes collapses. LaPlace, who denied causal status to the aether and spiritual forces, assigned that status to force functions. For instance, he held that the cause of the mutual motions of earth and moon is the $1/r^2$ force function between the two masses (to a first approximation, of course). LaPlace presumably also would maintain that the cause of a nail's motion toward a magnet is a force exerted by the magnet and that the cause of the formation of salt is the attractive force between particles of sodium and particles of chlorine.

LaPlace implemented the program Newton had outlined in *Queries* by extending the species of particles subject to central-force interactions to include particles of "imponderable matter." LaPlace included among species comprised of particles of imponderable matter heat, light, electricity, and magnetism.[12] The particles of species of imponderable matter were held to be mutually repulsive, but to be linked by short-range attractive forces to particles of ordinary matter. LaPlace sought to account for optical, magnetic, and electrical phenomena by hypothesizing central forces that act between particles of ponderable matter and particles of imponderable matter.

Newton had accounted for gravitational effects by assigning the value $n = 2$ to a $1/R^n$ force law. The magnitude of gravitational attraction decreases relatively slowly with increasing distance. LaPlace noted that as the value of n increases, the effective range of a force decreases. He postulated short-range attractive forces between material particles and “particles of light” as the cause of refraction.

LaPlace’s colleague C. L. Berthollet took a similar approach to the causal analysis of chemical reactions. He attributed the chemical affinity of substance *A* for substance *B* to the combination of two factors: short-range central forces between *A* and *B* and the density of *A* particles in proximity to those of substance *B*.

The LaPlacian Program was short-lived. In 1820, H. C. Oersted demonstrated that the magnetic force created by a current-carrying wire is not a central force.[13] Oersted placed a compass needle beneath a straight length of wire. He observed that when a current passes through the wire, the compass needle takes a position at right angles to the wire. Moreover, the needle rotates 180° when the current direction is reversed. Oersted concluded that the magnetic force is exerted in a plane perpendicular to the flow of current. It is not a central force that acts along a line joining the centers of the bodies affected. This result was a blow to the LaPlacian Program. To accommodate electromagnetic forces, it was necessary to abandon the thesis that all naturally occurring forces are central forces of attraction or repulsion.

A second difficulty for the LaPlacian Program was the success achieved by Fourier’s theory of heat conduction. Fourier formulated a mathematical equation for heat flow that makes no mention of forces between imponderable particles of heat.[14] Of course, a supporter of the LaPlacian Program could accept Fourier’s theory as a preliminary step on the way to a “deeper” theory in terms of central-force interactions.

Perhaps the most important difficulty for the LaPlacian Program was the series of successes achieved by the wave theory of light in the period around 1820. One widely publicized success was the experimental demonstration that there is a bright spot at the center of the shadow cast by an opaque disk perpendicular to a point source of light (1819).[15]

S. D. Poisson had shown that this counterintuitive result was required by Fresnel's wave theory. The "Poisson bright spot" is an anomaly, however, for the corpuscular theory of light embraced by followers of the LaPlacian Program. If light really is a wave phenomenon and not a collection of moving particles of imponderable matter, pursuit of the LaPlacian Program would be imprudent. Interest in the LaPlacian Program declined throughout the 1820s.

The Cartesian and LaPlacian Programs share two assumptions about causation. Both programs take events observed at the macro-level to be effects of processes that take place at a micro-level. For the Cartesian Program, these processes are collisions. For the LaPlacian Program, these processes are motions produced by attractive or repulsive central forces. In addition, both programs take the meaning of "cause" to include the notion of "productive efficacy." A cause is something that produces an effect. We are familiar with this notion of "productive efficacy" from everyday experience. For example, we are aware that our muscular exertions succeed in opening doors, tying shoelaces, and removing corks from wine bottles.

Notes

1. Christiaan Huygens, *Journal des scavans* 2 (1669): 531–36; *Philosophical Transactions* 46 (1669): 925–28.
2. René Descartes, *Principles of Philosophy*, 1644, trans. V. R. Miller and R. P. Miller (Dordrecht: Reidel, 1983), 86.
3. See, for instance, E. J. Aiton, *The Vortex Theory of Planetary Motions* (New York: American Elsevier, 1972).
4. Ibid., 132–36.
5. Huygens, *Discours sur le cause de la pesanteur* (Leiden, 1690). Quoted in Alexandre Koyre, *Newtonian Studies* (Cambridge, MA: Harvard University Press, 1965), 119–21.
6. Isaac Newton, *Opticks*, 1730 (New York: Dover, 1952), 401–02.
7. Ibid., 350.
8. Newton, *Mathematical Principles of Natural Philosophy*, trans. A. Motte, revised by F. Cajori (Berkeley: University of California Press, 1962), vol. 2, 634.
9. Newton, "Letter II to Richard Bentley," in *Isaac Newton's Papers and Letters on Natural Philosophy*, ed. I. Bernard Cohen (Cambridge: Harvard University Press, 1958), 298.
10. Richard Bentley, "A Confutation of Atheism from the Origin and Frame of the World," in *Isaac Newton's Papers and Letters on Natural Philosophy*, 340–44.
11. Newton, *Mathematical Principles of Natural Philosophy*, vol. 2, 547.
12. Pierre-Simon de LaPlace, *Traité de Mecanique Celeste (1799–1825)*, trans. Nathaniel Bowditch (Boston: Hilliard, Gray, Little and Wilkins, 1839), vol. IV, Book 10, Supplement, 1000–07.

13. H. C. Oersted, "Experiments on the Effect of a Current of Electricity on the Magnetic Needle" (1820), in *Source Book in Physics*, ed. by W. F. Magee (Cambridge, MA: Harvard University Press, 1963), 438–39.
14. $\lambda(\mathrm{d}^2\theta/\mathrm{d}x^2 + \mathrm{d}^2\theta/\mathrm{d}y^2 + \mathrm{d}^2\theta/\mathrm{d}z^2) = \rho c(\mathrm{d}\theta/\mathrm{d}t)$, where θ is absolute temperature, x, y, z are spatial coordinates of a point in an infinitely long slab of material, t is time, λ is the thermal conductivity of the material, ρ is its density, and c is its specific heat.
15. See, for instance, John Worrall, "Fresnel, Poisson and the White Spot: The Role of Successful Predictions in the Acceptance of Scientific Theories," in *The Uses of Experiment*, ed. D. Gooding, T. Pinch, and S. Schaffer (Cambridge: Cambridge University Press, 1989), 142–46.

6

David Hume on Causality

David Hume (1711–1776) was critical of any theory that holds that causes somehow "produce" their effects. He maintained that if we need to understand "productive efficacy" before we can understand "causation," then we must remain ignorant of causes and effects.

We see the cutting of the apple's stem and the subsequent fall of the apple. We do not see any power or "glue" that ties one event to the other. What we can know about causation is that instances of one type of event have been followed by instances of a second type of event. Sugar added to coffee dissolves; sand added to water does not. Hume suggested that "cause" be defined as follows:

> we may define a cause to be an object, followed by another, and where all the objects similar to the first are followed by objects similar to the second.[1]

This definition interprets causes to be "objects." However, it is clear from the context in which the definition is given that Hume's "objects" are events. Hume's illustrations of causal relations are the vibrations of a string that produce a particular sound and the collisions of billiard balls.[2]

Hume's "regularity view" of causal relatedness requires that each event of type *A* is followed by an event of type *B*. Since Hume elsewhere denied that we can know that "All *A*s are *B*s" on the basis of our knowledge that "Some *A*s are *B*s," the only knowledge of a causal relation that can be achieved is that each event of type *A* hitherto observed has been followed by an event of type *B*. If knowledge of causal relations is possible (for cases in which not every event of type *A* has been examined), then "are followed by" in the definition must be interpreted to mean "have been followed by." On this reading, Hume defines "causal relation" to be a "*de facto* constant sequential conjunction." This position will be referred to henceforth as Hume's "official position" on causal relations.

Hume's "official position" fails as a theory of causal relatedness. It is false that every *de facto* constant sequential conjunction is a causal relation. Consider two pendulum clocks of identical length placed side by side. If the pendulums are arranged to swing 90° out of phase, then each tick of clock 1 is observed to be followed by a tick of clock 2. On Hume's "official position," there is a causal relation between the two sets of events. But this violates our intuitions about causes and effects. If we stop the first pendulum, the second pendulum continues to swing. And if we stop the second pendulum, the continuing swings of the first pendulum fail to restart it.

John Stuart Mill noted that the sequence day–night is a constant sequential conjunction of events, but that day is not the cause of night (nor vice versa).[3] According to Mill, both day and night are effects of a further cause—a set of conditions that include the axial rotation of the earth, its relative rates of rotation and revolution, and the energy production of the sun. C. J. Ducasse subsequently observed that, on the regularity view, "the growth of hair on babies is the cause of their growth of teeth."[4]

A more esoteric, but historically important, example is the neutrino-detection experiment. In this experiment, a neutrino interacts with a hydrogen nucleus of a water molecule. The products are a positron and a neutron. The positron is quickly annihilated upon collision with an electron. The neutron travels a short distance and is absorbed by a cadmium nucleus, a process accompanied by the emission of γ-rays. The positron–electron collision is not the cause of the γ-ray emission, even though a constant sequential conjunction is displayed. Both are effects of the neutrino–hydrogen interaction. However, the positron–electron-collision–γ-ray-emission sequence is a causal relation on Hume's "official position." In general, if two temporally ordered events regularly follow the occurrence of a cause event, then the two effect events constitute a "causal relation" on Hume's "official position," and this is counterintuitive.

It also is false that every causal relation is a *de facto* constant sequential conjunction. For instance, Bob's exposure to a burst of radiation caused his death even though only 10 percent of healthy individuals at the same distance from the source died as a result of their exposure. The regularity view, which requires that every *c* event be followed by an *e* event, cannot account for the causal significance of Bob's exposure. Since we assign causal significance to numerous statistical correlations that are excluded by the regularity view, the regularity view is not a necessary condition of causal relatedness.

Hume's Counterfactual Conditional Stipulation

Hume himself was uneasy about the regularity view. Immediately after defining "causal relation" to be "a constant sequential conjunction of the members of two classes of events," he declared (in *Enquiry*, but not *Treatise*) that

> or, in other words, if the first object had not been, the second never had existed.[5]

In addition, Hume listed in *A Treatise of Human Nature* a list of "rules by which to judge of causes and effects."[6] Application of these rules include anticipations of Mill's methods of agreement, difference, and concomitant variations.

Hume's Equivocal Use of "Causal Relation" and Its Consequences

Hume's position on causal relation is extremely complex. He assigned four distinct meanings to the phrase "causal relation." Hume maintained that a "genuine causal relation" fulfills four conditions: spatial contiguity, temporal succession, constant conjunction, and necessary connection. According to Hume, there can be no knowledge of "causal relation" in this sense. We have no sensory impression of a necessary connectedness between events, and it is not possible to deduce the existence of an effect from knowledge of its cause.

Although we cannot have knowledge of "genuine causal relations," it is possible to know "causal relations" in the sense of *de facto* constant sequential conjunctions of the members of two classes of events. Hume presented this regularity view of causal relations in two forms—objective and subjective. The objective version has been discussed above. The subjective version includes the proviso:

> and whose appearance [an event of the first class] always conveys the thought to that other [an event of the second class].[7]

Hume held that the mind, having experienced a correlation, anticipates the occurrence of an effect event upon presentation of an associated cause event. This anticipation is the source of our conviction that a necessary connectedness undergirds the correlation.

The subjective version of the regularity view stipulates that a causal relation is a three-term relation. In addition to the experience of cause events and conjoined effect events, there is an anticipation or expectation associated with the former. Anticipation is not present at the first experience of the cause event, but becomes present at some point in the experience of constant conjunction. In some cases, the anticipation may arise from an experience of some other sequence. Moreover, Hume did not restrict these experiences of regularity to a single individual. He maintained that testimony from others often is important. I may anticipate an effect event on the basis of what others have said about a correlation, even though I have no personal experience of the correlation. Whether or not I am prudent to do so is, of course, an empirical question.

Lastly, Hume unpacked "causal relation" as a contrary-to-fact conditional claim. As noted above, this usage is inconsistent with the regularity view.

That Hume assigned these several meanings to "causal relation" complicates efforts to extract a consistent statement of his philosophy. Hume invoked disparate meanings of "cause" in his discussions of philosophical issues.

In the section "On Miracles" of *An Enquiry Concerning Human Understanding*, for instance, Hume declared that

> a miracle is a violation of the laws of nature; and as a *firm and unalterable* experience has established these laws, the proof against a miracle, from the very nature of the fact, is as entire as any argument from experience can possibly be imagined.[8] (Italics added)

It is not just that our experience has been uniform in support of the correlations stated in the laws of nature (the regularity view). This experience is held to be "unalterable." But an experience that cannot be other than it is surely is a necessary experience. Since our necessary experience has established the laws of nature,[9] and this necessary experience is "a proof against an (alleged) miracle," it would appear that these laws themselves are necessary. If this is so, then the laws of nature express "genuine causal relations," viz., *necessarily connected* constant sequential conjunctions.

One page later, Hume muddied the waters by claiming that

> there must, therefore, be a uniform experience against every miraculous event, otherwise the event would not merit that

> appellation. And as a uniform experience amounts to a proof, there is here a direct and full *proof*, from the nature of the fact, against the existence of any miracle.[10]

To speak of a "uniform experience," rather than an "unalterable experience," is to implement the regularity view. But a uniform experience cannot prove a necessity such as the impossibility of the existence of a miracle. Hume elsewhere is quite clear that no premise about what has been experienced can imply that subsequent experience will conform to what has been experienced.[11]

In his discussion "Of Liberty and Necessity," by contrast, Hume consistently applied the regularity view. He maintained that

> we know nothing farther of causation of any kind than merely the *constant conjunction* of objects, and the consequent *inference* of the mind from one to another.[12]

Hume sought to defend the regularity view by means of a burden-of-proof shift. He invited his critics to show that there is an objective necessary relationship between events. An objective necessary relationship, if it could be found, would validate our habitual expectations that arise from experiences of constant sequential conjunctions. Since this objective necessity would have to be derived from sense impressions or be deduced from an analysis of the cause event, Hume was convinced that his burden-of-proof shift is successful.

Interpreters of Hume, cognizant of his diverse comments about causation, have issued different judgments on the relative importance of these comments. Norman Kemp Smith, Bertrand Russell, and John Passmore have emphasized the regularity views (objective and subjective) and the skeptical consequences of these views.[13] One can extract from Hume's writings the position that nothing can be known about natural phenomena over and above descriptions of events and their *de facto* constant sequential conjunctions. Since "genuine causal relations" are *necessarily connected* constant sequential conjunctions, we can have no "genuine" causal knowledge. Moreover, knowledge that the members of two classes of events have been conjoined fails to provide a rational justification for projection onto instances not yet encountered.

We are saved from paralysis by our human nature, which operates by reliance on expectation and custom. On this interpretation, Hume combines epistemological skepticism with confidence in our natural inclinations on the level of practice. In a much quoted passage, Hume declared that

> Custom, then, is the great guide of human life. It is that principle alone which renders our experience useful to us, and makes us expect, for the future, a similar train of events with those which have appeared in the past. Without the influence of custom, we should be entirely ignorant of every matter of fact beyond what is immediately present to the memory and senses.[14]

This passage highlights the problems that confront the interpreter of Hume. Hume evidently was anxious to find an antidote to epistemological skepticism. Because of custom, we are "not ignorant" of matters of fact that go beyond what is immediately present to us. On the other hand, the label "not ignorant of" cannot be cashed in terms of "knowledge." It seems that, according to Hume, we both have and do not have empirical knowledge that transcends *de facto* correlations.

Fred Wilson placed a quite different emphasis on Hume's comments about causation.[15] Wilson maintained that Hume subscribed to the "genuine causal relation" definition and not the "*de facto* regularity view." He held that Hume was well aware that accidental correlations qualify as causal on the objective version of the regularity view. Wilson called attention to Hume's admission in *Treatise* that

> an object may be contiguous and prior to another, without being consider'd as its cause. There is a NECESSARY CONNEXION to be taken into consideration; and that relation is of much greater importance than any of the other two above-mention'd.[16]

That Hume provided a contrary-to-fact-conditional definition (in *Enquiry* but not in *Treatise*) and a "list of rules by which to judge of causes and effects" is a further indication that he took causal relatedness to be something more than *de facto* constant sequential conjunction. Wilson maintained that Hume promoted the search for genuine causal connections (as opposed to mere *de facto* constant conjunctions) by means of empirical inquiry designed to satisfy the "rules by which to judge of causes and effects." He did so, according to Wilson, because he believed that the unavoidable habitual expectations, upon which we organize our experience, can be modified so as to be more effective.

That commentaries on Hume's position have been so diverse is a strong indication that Hume did not consistently maintain an unequivocal view of causation. The regularity view is the interpretation most discussed. Hume sometimes defended this view as the correct interpretation of "causal relation" and on other occasions appeared to deny this. The regularity view clearly is inadequate, however. Whatever be the correct interpretation of "causal relation," "*de facto* constant sequential conjunction" is not it.

Notes

1. David Hume, *An Enquiry Concerning Human Understanding* (Chicago: Open Court, 1927), 79.
2. Ibid., 79–81.
3. John Stuart Mill, *A System of Logic* (London: Longmans, Green, 1865), Book III, Ch. 5, 6.
4. Curt J. Ducasse, "Critique of Hume's Conception of Causality," *Journal of Philosophy* 63 (1966): 148.
5. Hume, *An Enquiry Concerning Human Understanding*, 79.
6. Hume, *A Treatise of Human Nature*, Selby-Bigge edition, first edition 1739 (Oxford: Clarendon Press, 1965), 173–74.
7. Hume, *An Enquiry Concerning Human Understanding*, 79.
8. Ibid., 120.
9. Hume cited as laws of nature "all men must die," "lead cannot, of itself, remain suspended in the air," and "fire consumes wood, and is extinguished by water."
10. Hume, *An Enquiry Concerning Human Understanding*, 37.
11. Ibid., 38.
12. Ibid., 96.
13. Norman Kemp Smith, *The Philosophy of David Hume* (London: MacMillan, 1941); Bertrand Russell, *History of Western Philosophy* (New York: Simon and Schuster, 1945), 659–74; John Passmore, *Hume's Intentions* (London: Duckworth, 1968).
14. Hume, *An Enquiry Concerning Human Understanding*, 45.
15. Fred Wilson, *Hume's Defence of Causal Inference* (Toronto: University of Toronto Press, 1997).
16. Hume, *A Treatise of Human Nature*, 77.

7

Kant's Response to Hume's Regularity View

Hume's contemporary Immanuel Kant (1724–1804) professed to be greatly distressed by the skeptical consequences of Hume's "official position" on causality. Kant was convinced that the basic laws of geometry and physics state necessary connections among phenomena. Euclid's axioms and Newton's laws of motion are more than statements of previously observed regularities. They state relations that hold everywhere at all times.

Kant conceded that Hume was correct to argue that *if* all empirical knowledge arises from, and is given in, sense impressions, then only *de facto* sequential regularities can be known. He pointed out that Hume's conclusion can be avoided by denying that all empirical knowledge is given in sense impressions. According to Kant, the knowing subject itself contributes empirically significant principles that establish and organize empirical knowledge.

On Kant's view, atomistic, unstructured sensations produced by the motions of external "things-in-themselves" are organized spatially and temporally by the "sensibility" and then categorized by the "understanding" (Figure 7.1).

Causality is one of twelve categories of the understanding. It superimposes upon spatiotemporally organized impressions the condition that "every event is determined by a cause according to constant laws."[1] Kant provided the following application of the category of causality:

> I see, for instance, a ship gliding down a stream. My perception of its place below follows my perception of its place higher up in the course of the stream, and it is impossible in the apprehension of this phenomenon that the ship could be perceived first below and then higher up.[2]

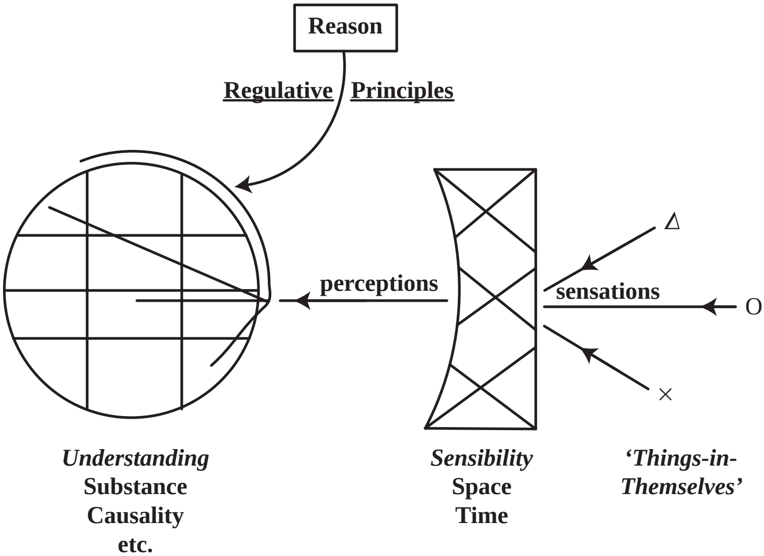


Figure 7.1 Kant's view of cognitive experience.

Contrast the visual scanning of a house:

> my perceptions could begin in the apprehension at the roof and end in the basement, or begin below and end above.[3]

It is application of the category of causality that distinguishes the perceptual experience of the ship (or, better, a log floating downstream) from the perceptual experience of the house. In the former case, but not the latter, perceptions "follow according to a rule."[4] Successive states stand in the relationship of cause to effect.

The relation of cause and effect is part of the cognitive apparatus applied by a person who acquires empirical knowledge. Kant maintained, moreover, that a principle of causality is a necessary condition of objective empirical knowledge. The principle of causality states that every state has an antecedent state from which it follows according to a rule. Within the discipline of mechanics, the principle is expressed as Newton's axiom of inertial motion. Newton's axiom stipulates that all changes of motion of a body result from forces extrinsic to the body. To acquire knowledge about a case of accelerated motion is to identify the force responsible.

Kant shifted the subject of causal relatedness from successive events to successive states of physical systems. He defended a version of the sufficient-condition interpretation of causality. One state (S_1) is the cause of a second (S_2) provided that S_1 is a sufficient condition of S_2.

In a suspected case of causal relatedness, the inquirer is faced with a choice about how to characterize "physical state."[5] If a supposed cause-state is narrowly defined by reference to just a few variables, then one can investigate whether a predicted effect-state is in fact realized. If the effect-state is realized, then causal relatedness is confirmed. On the other hand, if the supposed cause-state is so broadly defined as to approach a specification of the state of the entire universe, then the causal relation is at risk of becoming a non-falsifiable definition. Given the state of the universe *A* and causal relation "*A* causes *B*," if *A* seems to occur and *B* does not follow, one may conclude that the initial state was not really an *A*-state. The nonoccurrence of a *B*-state is taken, not to falsify the claim of causal relatedness, but to disqualify the putative *A*-state. On this approach, *A*-states are defined, in part, as states followed by *B*-states. The inquirer's choice is that the causal relation—*A* causes *B*—may be taken to be vague and factual, or precise and inapplicable.

Notes

1. Immanuel Kant, *Prolegomena to Any Future Metaphysics* (1783) (Indianapolis: Bobbs-Merrill, 1950), 43.
2. Kant, *Critique of Pure Reason* (1781), trans. F. Max Muller (London: MacMillan, 1934), 157.
3. Ibid., 157.
4. Ibid., 155.
5. See, for instance, Henry Margenau, "Meaning and Scientific Status of Causality" (1934), in *Physics and Philosophy: Selected Essays* (Dordrecht: Reidel, 1978), 39–51; Philipp Frank, *Philosophy of Science* (Englewood Cliffs, NJ: Prentice-Hall, 1957), 281–83.

8

J. S. Mill on Invariable and Unconditional Correlations

In a section on "Rules of Reasoning in Philosophy" in Book III of *Principia*, Newton spoke of "*verae causae*." He insisted that we are to cite "no more causes than are both true and sufficient to explain effects."[1] Newton did not make clear what difference, if any, there is between a cause and a "true cause." He did, however, require that a cause be a sufficient condition of its effect.

John Stuart Mill suggested subsequently that to establish that circumstance C is a sufficient condition of phenomenon p, the investigator should look for, or produce, a number of instances in which C is present along with diverse other circumstances. If C is the only circumstance present for each of the instances in which p is observed, then one may conclude that it is probable that C is a sufficient condition of effect p. Mill referred to this procedure as the "method of agreement.[2]"

Method of Agreement

Instance	Circumstances	Phenomena
1	$A\ B\ C$	p
2	$B\ C\ D\ E$	p
3	$A\ C\ F$	p
4	$C\ D\ G$	p

∴ It is probable that C is the cause (*qua* sufficient condition) of p.

Only probability can be claimed for conclusions reached by instantiating this schema. One reason for this is that the investigator may have overlooked further circumstance X, which is present in each instance and which is the real cause of p. A further reason is that there may be a plurality of causes present in a given instantiation of the schema. For instance, A may cause p in instances 1 and 3, and D may cause p in instances 2 and 4. The presence of C in each instance may be irrelevant to the occurrence of p.

Mill noted that the problem posed by the possibility of a plurality of causes does not arise in applications of the method of difference.[3] In this procedure, the investigator observes, or creates, just two instances. In the first instance, p occurs; in the second instance, it does not. If the two instances are observed to differ in just one circumstance C, then one may conclude that it is probable that C is a necessary condition of effect p.

Method of Difference

Instance	**Circumstances**	**Phenomena**
1	*A B C*	*p*
2	*A B*	—

∴ It is probable that C is the cause (*qua necessary* condition) of p.

As is the case with the method of agreement, only probability can be claimed for conclusions reached by instantiating this schema. One reason for this is that the investigator may have overlooked some further circumstance X that is present in instance 1 and absent in instance 2. It may be the case that X, and not C, is responsible for the occurrence of p. Any two instances differ in innumerable respects—spatial location, time, illumination, the motions of atoms, cosmic ray intensity, etc. Consequently, success in applying the method of difference depends on the truth of the hypothesis that unlisted circumstances are irrelevant to the occurrence of p.

Although he sometimes made extravagant claims for what can be achieved through applications of the method of difference, Mill was aware of the problem posed by relevant circumstances that are not taken into account. He discussed a case of arsenic poisoning.[4] Suppose Jones and his wife dine together one evening and Jones dies shortly afterward. Examination of the food reveals the presence of arsenious acid on Jones's plate, but not on his wife's plate.

Instance	Circumstances	Phenomena
Jones	*A B C D*	*d*
Wife	*B C D*	—

∴ It is probable that A is the cause (qua necessary condition) of d,

where *d* is death, *A* is the ingestion of arsenious acid, and *B*, *C*, and *D* are other circumstances present in the two instances.

Mill noted that it would be incorrect to generalize that ingestion of arsenious acid invariably causes death. In the presence of hydrated peroxide of iron (*I*), ingestion of arsenious acid does not result in death, viz.:

Instance	Circumstances	Phenomena
Jones	*A B C D I*	—
Wife	*B C D*	—

In addition, Mill noted that the method of difference is unavailing in situations in which circumstances reinforce or cancel one another.[5] The standard illustration is the composition of forces. Resultant force F_R is the diagonal of the parallelogram whose sides are F_1 and F_2 (Figure 8.1).

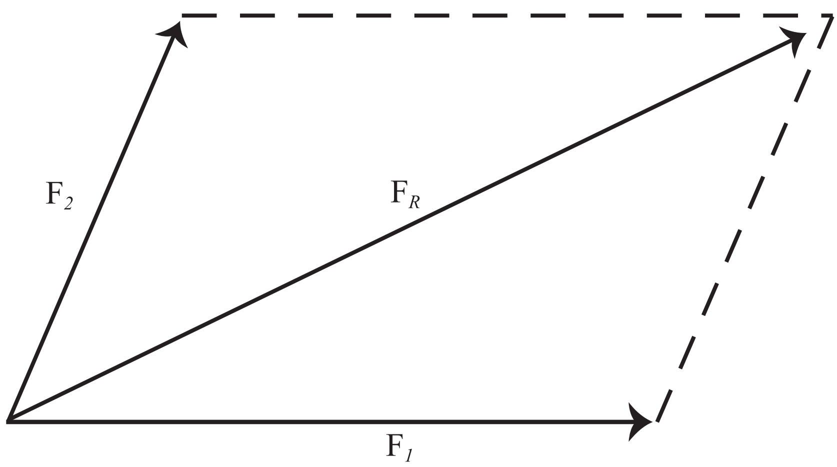


Figure 8.1 The parallelogram of forces.

Innumerably many component-force combinations can produce the resultant force F_R. Consequently, if F_R is the effect whose causes are to be discovered, one cannot establish these causes by applying the method of difference. The method of difference is of no use in situations in which multiple causes produce effects by reinforcement or cancellation.

Mill defined a cause to be a circumstance or set of circumstances both invariably and unconditionally followed by an effect of a given type.[6] Application of the schemata of agreement and difference can establish *de facto* invariability for those instances considered. It takes an inductive leap from what has been observed to what is presumed true of unobserved instances to warrant a judgment of causal relatedness. What does it mean to require that a causal relation be "unconditional"? Mill maintained that

> it is necessary to our using the word 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 things endures it always *will* be so.[7]

Mill indicated that by "the present constitution of things" he means the "ultimate laws of nature." Newton's law of universal gravitational attraction is one such ultimate law. The regular diurnal rotation of the earth is not, since it might be "terminated or altered by natural causes."[8] Mill's "unconditionality" is thus relative. A relationship is unconditional provided that it has been invariable and will continue to be so as long as the ultimate laws of nature remain unchanged. He held that we are justified to believe that the ultimate laws will always remain the same.

Mill advanced an argument designed to prove that nature is uniform, that the ultimate laws remain the same. Given a phenomenon of type p and a type of circumstance A on which it has been invariably consequent, if nature is uniform, then p is always consequent on A. Mill believed that he could establish the truth of this generalization by an appeal to experience. Numerous inductive arguments by simple enumeration have been successful. Mill conceded that such arguments are precarious. They are at the mercy of a single exception. Mill insisted, however, that every supposed exception "sufficiently open to our observation" has been traced either to the "absence of an antecedent circumstance ordinarily present or to the presence of a circumstance ordinarily absent."[9] Moreover, he claimed, successful arguments by simple enumeration are so numerous and extensive that this precarious method becomes reliable as applied to the question of the uniformity of nature as a whole. Mill declared that

> the precariousness of the method of simple enumeration is in an inverse ratio to the largeness of the generalization. The process is delusive and insufficient, exactly in proportion as the subject-matter of the observation is special and limited in extent. As the sphere widens, this unscientific method becomes less and less liable to mislead.[10]

An inductive argument by simple enumeration thus proves the uniformity of nature. Given the true premise that nature is uniform, the method of difference can be transformed into a rule of proof of causal relations, viz.,

1.

Instance	Circumstances	Phenomena
1	*A B C*	*p*
2	*B C*	—

2. *A*, *B*, and *C* are the only circumstances relevant to the occurrence of *p*.
3. If [*A B C* .... *p*
   B C .... —] is observed for instances 1 and 2,
   then [*A B C* .... *p*
   B C .... —] is the case for every set of such instances (viz., nature is uniform).

Therefore *A* is the cause (*qua* invariable and unconditional antecedent) of *p*.

Philosophers of science are in nearly unanimous agreement that Mill failed to prove his case. Premise 3 is essential to the proof. But no amount of inductive evidence—by simple enumeration or otherwise—is sufficient to establish that nature is uniform. Indeed Mill had been refuted in advance by David Hume. Hume had argued persuasively that no amount of information about what has been observed can prove what is the case in unexamined instances.

Mill's claim that the method of difference is a rule of proof of causal connection must be rejected. Nevertheless this method is a valuable investigative technique within science. The difference schema represents the ideal of a controlled experiment.

The history of science contains numerous examples of episodes in which relevant circumstances were correctly identified and controlled. Examples include Walter Reed's discovery of the cause of yellow fever and Christiaan Eijkman's discovery of the cause of beriberi.

Yellow fever is an often deadly disease that beset American troops in Cuba at the turn of the twentieth century. Initial attempts to identify the cause of this disease sought a bacillus supposedly transferred from one victim to another. Cultures created from the blood and/or body parts of infected individuals were examined and found lacking in microbes that could transmit the disease.

There was a second possibility. A physician named Carlos Finlay suggested in 1900 that yellow fever was transmitted by bites from mosquitoes that had fed on fever victims. Major Walter Reed was dispatched to Havana with orders to

> give special attention to questions relating to the cause and prevention of yellow fever.[11]

Since the disease does not affect animals, the appropriate subjects for experimental inquiry are human beings. Reed exposed volunteers to bites from mosquitoes that had fed upon yellow fever victims. Two of ten volunteers contracted the disease. Fortunately, both recovered.

Reed and his colleagues then set up a controlled experiment in which two groups of "volunteers" (some of whom were paid to participate) occupied quarters that differed with respect to the presence of yellow-fever-fed mosquitoes. This was an attempt to satisfy Mill's method of difference, viz.:

Instance	**Circumstances**	**Phenomena**
Room 1	$M\ A\ B$	y
Room 2	$A\ B$	—

∴ It is probable that M is a cause (*qua* necessary condition) of y,

where M is the presence of yellow-fever-fed mosquitoes, A and B are other circumstances common to the two instances, and y is the contraction of yellow fever.

The above instantiation of the difference schema lists *M* as the only circumstance that differs in the two instances. In practice, this cannot be arranged. The participants in room 1 differed from those in room 2 in age, weight, eye color, hair length, etc. However, if *M* is the *only relevant circumstance* that differs in the two cases, then instantiating the method of difference establishes causation (in the sense of necessary condition).

Reed next exposed members of the control group from room 2 to the vomit-laden bedding of yellow fever victims. These subjects remained free of the disease. Exposure to such bedding is not sufficient to produce yellow fever. Reed took this to be further evidence that it is mosquito bites that need to be guarded against, despite the fact that he could not *prove* that no factor other than bites from yellow-fever-fed mosquitoes is relevant to the different results in rooms 1 and 2.

One may object that the cause of yellow fever is a sequence of events—a mosquito feeds on the blood of a yellow fever patient, bites a victim, ejected material enters the bloodstream of the victim, etc.—and that emphasis on "cause-as-necessary-condition" is superfluous. This is a defensible position. However, the inquiry as actually carried out was a search for a "cause-as-necessary-condition" via instantiation of Mill's method of difference.

Cause as Absence: Beriberi and Vitamin B_1

In the course of his medical service in Java (1886), Christiaan Eijkman discovered a correlation between diet and the occurrence of beriberi, a disease that produces muscle weakness, paralysis of the limbs, and sometimes death.[12] The correlation was first noted in fed chickens.

Diet	**Incidence of Beriberi**
Polished rice (kernels only)	High
Unpolished rice (kernels + husks)	Low

Eijkman was committed to the germ theory of disease. He reasoned that there must be a bacterium present in the kernels of rice and an antidote present in the outer husks. Beriberi then results when diet is restricted to the rice kernels that remain after polishing. But beriberi does not occur when rice husks are added to the diet.

Eijkman and others sought to isolate and identify the bacterium supposedly responsible for beriberi. These efforts failed. Gerrit Grijns, another Dutch physician working in Java, drew a different conclusion from Eijkman's data. According to Grijns, the cause of beriberi is an *absence* from the diet of a constituent that is necessary for good health. This constituent, later identified as Vitamin B_1, is present in the husks of rice, but not in the kernels.

One may express the situation in either of two ways: (1) the absence of Vitamin B_1 causes beriberi or (2) the presence of Vitamin B_1 is a cause (*qua* necessary condition) of good health.

Subsequent investigations established additional correlations between vitamin deficiencies and disease:

Vitamin deficiency	**Disease**
B3	Pellagra
C	Scurvy
D	Rickets

Instantiation of the difference schema does not prove causal connection, but efforts to instantiate the schema are extremely important in science.

Notes

1. Isaac Newton, *Mathematical Principles of Natural Philosophy*, trans. A. Motte, rev. F. Cajori (Berkeley: University of California Press, 1962), vol. II, 398.
2. John Stuart Mill, *A System of Logic*, 8th ed. (London: Longman, 1970), Book III, Ch. 8.
3. Ibid., Book III, Ch. 8.
4. Ibid., Book III, Ch. 9, 267–68.
5. Ibid., Book III, Ch. 11, 299–305.
6. Ibid., Book III, Ch. 5, 222.
7. Ibid., Book III, Ch. 5, 221.
8. Ibid., Book III, Ch. 5, 221 n.
9. Ibid., Book III, Ch. 21, 374.
10. Ibid., Book III, Ch. 21, 373.
11. Quoted in Paul de Kruif, *Microbe Hunters* (New York: Harcourt, Brace & World, 1953), 288.
12. See, for instance, Kenneth J. Carpenter, *Beriberi, White Rice and Vitamin B: A Disease, a Cause, and a Cure* (Berkeley: University of California Press, 2000).

9

Options for a Theory of Causality ca. 1900

By 1900, a variety of candidates for causal status had been proposed. Among them were:

1. events (Hume),
2. states of physical systems (Kant),
3. patterns (Aristotle's formal causes),
4. matter (Aristotle's material causes),
5. future states of affairs (Aristotle's final causes),
6. impacts or pressure (Descartes),
7. forces (LaPlace),
8. necessary conditions (Bacon; Mill's method of difference), and
9. sufficient conditions (Hume's contrary-to-fact conditional view; Mill's method of agreement).

In addition, a number of interpretations of the cause–effect relation had been proposed. Among them were:

1. causes produce effects (Aristotle et al.);
2. causes and effects are related by spatial contiguity, temporal sequence, constant conjunction, and necessary connection (Hume);
3. constant sequential conjunction (Hume);
4. counterfactual dependence—if the cause event had not occurred then the effect event would not have occurred (Hume);
5. causes of a given type invariably and unconditionally are followed by effects of a given type (Mill);
6. causes are states of physical systems from which effects follow according to a rule (Kant).

II

Early Twentieth-Century Theories: Dominance of the Regularity View

10

C. S. Peirce on Causation and Causality

Charles Sanders Peirce (1839–1914) insisted that a distinction be made between entities that qualify as cause events and conceptual representations of these events that participate in the relation of causality. Cause events are found in the world: causes are facts, viz. abstractions that express in propositional form properties or relations of events.

From Peirce's standpoint, the above list of candidates for causal status conflates ontological and epistemological considerations. Among the listed candidates for causal status, impacts, forces, and states of physical systems are aspects of the natural world. To single out an impact as a cause is to make an ontological claim. The status of "necessary condition" and "sufficient condition," on the other hand, is ambiguous. On the one hand, these are logical relations. If N is a necessary condition of effect E, then the argument "$E/\therefore N$" is valid. If S is a sufficient condition of effect E, then the argument "$S/\therefore E$" is valid. But a theorist may claim that necessary conditions and/or sufficient conditions exist as well in the external world.

Peirce distinguished "events" from "causes." Certain events in the world are productive of other events. There is causation in the world. Peirce held that there are three aspects of an instance of causation. First of all, there is efficient causation. This relation is irreversible. Secondly, there is final causation. The effect event is a link in a directed sequence of events. The directed sequence of events is an instantiation of a tendency. After the occurrence of cause event c_1, events other than e_1 might have followed. But given the tendency specified by final causation, that event will be an e-type event. Thirdly, there invariably is an element of chance present in causation.

Peirce argued that universal determinism is false. The spontaneity, diversity, and emergence of new forms that we observe in the universe cannot be explained by reference to temporally symmetric laws like universal gravitational attraction and conservation of energy. Peirce maintained that chance is an objective presence in the universe and that the laws involved in causation are probabilistic laws. He declared that

> by thus admitting pure spontaneity or life as a character of the universe, acting always and everywhere though restrained within narrow bounds by law, producing infinitesimal departures from law continually, and great ones with infinite infrequency, I account for all the variety and diversity of the universe, in the only sense in which the really *sui generis* and new can be said to be accounted for.[1]

Causality, unlike causation, is a relation among facts. To ask for the cause of the fall of an apple is to ask from which fact this fact follows, given the laws of physics. Peirce noted that what is required for an instance of causality is that

1. *E* follows *C*, where "*E*" and "*C*" are facts and
2. *E* follows *C* according to a rule.[2]

In the case at hand, "that apple fell to the earth" follows from "that apple's stem was cut" according to the rule "every unsupported object near the earth falls to the earth."

Instances of causality are applications of hypothesized generalizations. Peirce insisted that "the entire fabric of science has to be built up out of surmises at truth."[3] In anticipation of the position of Karl Popper, he declared that "all that experiment can do is to tell us when we have surmised wrong."[4]

Peirce held that science progresses by the testing of hypotheses. The scientist deduces observable consequences from a hypothesis, tests to see whether these consequences are realized, and if they are not realized, modifies or abandons the hypothesis.[5] Hypotheses that survive experimental testing become laws accepted by the science of the day.

Peirce conceded that laws hold only approximately and that they are subject to exceptions. He maintained, nevertheless, that a law has objective reality apart from the minds of scientists who apply it. This objective reality is established by the confirmation of predictions deduced from it.[6]

Notes

1. Charles Sanders Peirce, "The Doctrine of Necessity" (1892), in *Values in a Universe of Chance: Selected Writings of Charles S. Peirce*, ed. Philip P. Wiener (New York: Doubleday Anchor, 1958), 175.
2. Peirce, *Collected Papers of Charles Sanders Peirce*, ed. Charles Hartshorne, Paul Weiss, and Arthur W. Burks (Cambridge, MA: Harvard University Press, 1931), vol. VII, 7.349, p. 213.
3. Ibid., vol. VII, 7.88, p. 57.
4. Ibid., vol. VII, 7.88, p. 7.
5. Ibid, vol. VII, 7.84, p. 54.
6. Ibid., vol. VII, 81.53, p. 119.

11

Karl Pearson's Version of the Regularity View

Karl Pearson defended a regularity view of causality. In agreement with Hume's "official position," he insisted that there is no necessity involved in causality. Causes do not produce their effects. There is only the experience of a repeated sequence of perceptions. Pearson declared that

> whenever a sequence of perception D, E, F, G is invariably preceded by the perception C . . . C is said to be a *cause* of D, E, F, G, which are then described as its effects.[1]

To assess whether the sequence will be repeated in the future is to pass judgment on its probability. Pearson maintained that science progresses by a

> careful and often laborious classification of facts, in the comparison of their relationships and sequences, and finally in the discovery by aid of the disciplined imagination of a brief statement or *formula*, which in a few words resumes a wide range of facts.[2]

The "formula" of which Pearson speaks is the "scientific law." Consistent with his view on causality, he insisted that a law is nothing more than an economical summary of sequences of sense perceptions.[3] Thus, Copernicus achieved "an immense advance in brevity and accuracy of description"[4] by referring our observations of planetary motions to a heliostatic system. Kepler in turn summarized an immense amount of observed data under just three laws of motion. And Newton further increased economy of description by means of the formula of universal gravitational attraction.

However, it was indisputable at the end of the nineteenth century that successful scientific theories make claims about entities not accessible to perceptual experience. Pearson conceded that concepts such as "atom" and "molecule" play a role in science despite the fact that we have no direct sense experience of entities supposedly denoted by these concepts. "Atom" and "molecule" are

> intellectual conceptions by aid of which physicists classify phenomena and formulate the relationships between their sequences.[5]

There is just one relevant test of a scientific theory—one compares the results of applying the formulae with facts established by observation. A theory is true insofar as it passes the test of universal experience.[6]

In a review of Pearson's *The Grammar of Science*, Peirce complained that, by restricting the evaluation of theories to descriptive adequacy, Pearson failed to acknowledge the important role of prediction in scientific practice. For instance, Pearson failed to account for the special value scientists accorded the successful eclipse predictions of Newton and LaPlace.[7] "It is the universal opinion of men of science," Peirce declared, "that prediction is the seal of success."[8]

Notes

1. Karl Pearson, *The Grammar of Science* (1892) (New York: Meridian Books, 1957), 130.
2. Ibid., 77.
3. Ibid., 86–87.
4. Ibid., 98.
5. Ibid., 95.
6. Ibid., 100.
7. Charles Sanders Peirce, *Collected Papers*, ed. Charles Hartshorne, Paul Weiss, and Arthur W. Burks (Cambridge, MA: Harvard University Press, 1931), vol. VII, 8.153, p. 119.
8. Ibid., vol. VII, 8.154, p. 120.

12

Bertrand Russell and N. R. Campbell On Causal Relations in Science

In an influential essay published in 1912, Bertrand Russell issued a challenge to theorists of causality. He maintained that

> the word "cause" is so inextricably bound up with misleading associations as to make its complete extrusion from the philosophical vocabulary desirable.[1]

Russell's criticism was directed at the regularity view of causality. On this view, the appropriate statement of an instance of causation is

> if the event e_1 occurs at time t, it will be followed by the event e_2.[2]

The corresponding "principle of causality" is

> given any event e_1, there is an event e_2, such that, whenever e_1 occurs, e_2 occurs later.[3]

The "events" mentioned in the principle are universals of which there are instances. Russell noted that

> what is meant by an "event" is something like striking a match. . . . If such an event is to recur, it must not be defined too narrowly: we must not state with what degree of force the match is to be struck. . . . For if such considerations were relevant, our "event" would occur at most once, and the law would cease to give information. An "event," then is a universal defined sufficiently widely to admit of many particular occurrences in time being instances of it.[4]

The more fully we characterize a putative cause event, the less likely is it that there is a repetition of the event. The limiting case, of course, is to specify the state of the entire universe. In this limiting case, the causal relation is precise but inapplicable (barring an exact repetition of a state of the universe).

We must, then, allow a measure of vagueness in the characterization of putative cause events. The causal relation then can be tested, but disputes are likely on whether or not particular events qualify as cause events.

It was Russell's thesis that causal relations between vaguely characterized events are important only in the underdeveloped areas of science. In mature areas of science such as gravitational astronomy, "the word 'cause' never occurs."[5] He conceded that

> it may be that there never will be an exception to the rule that when a stone of more than a certain mass, moving with more than a certain velocity, comes in contact with a pane of glass of less than a certain thickness, the glass breaks. I also do not deny that the observation of such regularities, even when they are not without exceptions, is useful in the infancy of science: the observation that unsupported bodies in air usually fall was a stage on the way to the law of gravitation. What I deny is that science assumes the existence of invariable uniformities of sequence of this kind, or that it aims at discovering them.[6]

According to Russell, science aims at the discovery of functional relations that describe successive states of (approximately) isolated systems. These relations typically are differential equations involving variables such as position, velocity, electric field strength, magnetic field strength, and time. The relevant equations make no reference to "causes."

Russell was preoccupied with the mathematical structure of the theories of the advanced sciences. He failed to acknowledge the important role of causal judgments in the operation of instruments utilized to develop and test theories. To utilize an instrument to make a measurement is to invoke a concept of causation. In a typical case, apparatus *A* is constructed and is subject to manipulation *M* in anticipation of result *R*. *R* is properly spoken of as the "effect" of *M* on *A*. Causal analysis is involved in the description of experimental arrangements and results. However, it may be appropriate to restrict, or even exclude, causal concepts at the theoretical level.

The physicist and philosopher of science N. R. Campbell agreed with Russell's negative assessment of the importance of causal regularities in physics. In *Foundations of Science* (1919), Campbell noted that a regularity view of causality had been defended by Whewell, Hamilton, Mill, and Jevons. According to Campbell, there are four aspects of this regularity view. Causal relations are (1) dyadic, (2) invariable, (3) temporal, and (4) asymmetrical.[7]

Campbell observed that numerous scientific laws fail to satisfy these four conditions. Consider Ohm's law. For an electrical circuit of constant resistance, current is directly proportional to voltage. There is no temporal sequence involved. It is not that an increase in voltage at t_1 is followed by an increase in current at t_2. Condition 3 is not satisfied.[8] Moreover, Ohm's law is a symmetrical relation. Changes in voltage are accompanied by changes in current, and vice versa. Condition 4 is not satisfied either. The same is true of Boyle's law $(P \propto 1/V)_{T=k}$, Charles's law $(V \propto T)_{P=k}$ and the law of temperature dependence of pressure $(P \propto T)_{V=k}$ The ideal gas law ($PV = kT$), which combines these relations, is not a causal relation either. It is a nontemporal, symmetrical, nondyadic relation.

Hooke's law, which correlates the force on a spring with its extension ($F = -k\delta x$), also fails to satisfy the conditions of causal status. The relation is dyadic, invariable, and asymmetrical, but nontemporal. The law does not specify that a force applied at t_1 is followed by lengthening (or compression) at t_2.

This is a substantial list of noncausal laws. The obvious next question is whether there are *any* scientific laws that satisfy the four conditions of causal status. Campbell's answer is "Yes." He gave as an example the explosion of a hydrogen–oxygen mixture upon passing a spark through the mixture. There is a small, but definite, time interval between passage of the spark and the explosion. The relation is invariable and asymmetrical. We cannot reverse the order so that the explosion precedes sparking.

Campbell pointed out, however, that whenever scientists have the opportunity to replace such low-level causal laws with functional relations applicable to processes, they do so.[9] In free fall, for example, a certain distance of fall in the first time interval is followed by a certain distance of fall in the second equal time interval. This relation is "causal" in the sense of the regularity view. But scientists do not discuss free fall in terms of such causal relations. Rather, they affirm the functional relation "$d^2s/dt^2 = g$." The differential equation is taken to be fundamental. The "causal" relations between temporal segments are subsumed by the equation.

A defender of the importance of causal relations in science could argue that scientists' acceptance of the differential equation governing free fall is really acceptance of a formula that summarizes numerous causal relations. Campbell pointed out that this defense will not work in the case of the law of inertia. Suppose a body is set into inertial motion by an impact at t_1 and has this inertial motion altered by a second impact at t_3. During the period from t_1 to t_3, there is continuous motion. The motion from t_2 to t_3 follows the motion from t_1 to t_2 such that the four conditions of causal status are satisfied. However, inertial motion is uncaused motion. It is motion in the absence of impressed forces. What is important is the differential equation governing this motion ($d^2s/dt^2 = 0$) and not a wrongly supposed "causal" relation between stages of the motion. Campbell's conclusion is that causal relations play a much less important role in physics than has been claimed by defenders of the regularity view.[10]

Notes

1. Bertrand Russell, "On the Notion of Cause" (1912), in *Mysticism and Logic* (New York: Doubleday Anchor, 1917), 174.
2. Ibid., 177.
3. Ibid., 177.
4. Ibid., 180.
5. Ibid., 194.
6. Ibid., 181.
7. Norman Robert Campbell, *Foundations of Science* (1919) (New York: Dover, 1957), 58.
8. Ibid., 59–60.
9. Ibid., 66.
10. Ibid., 67.

13

Causality in Physics: Revision of the Regularity View

During the early decades of the twentieth century, there was general agreement among scientists and philosophers interested in physics that

1. causal relatedness is a matter of uniformity of sequence and
2. the regularity in question is between states of (relatively isolated) systems.

Moritz Schlick pointed out that the starting point of inquiry in physics is the measurement of quantities rather than the description of events.[1] These quantities are values of the state of a physical system. If a functional relation exists linking successive states of the system such that whenever state *A* is realized it is followed by state *B*, then the states are causally related.

Suppose that a bullet is fired into the center of a block of wood in such a way that the block-with-embedded-bullet leaves the edge of a table and falls to the floor. The block and bullet may be considered to constitute a single physical system that assumes a succession of states. Two of these states are illustrated in Figure 13.1.

Physicists have found that the horizontal component of the motion of such systems can be predicted if the masses, positions, and velocities of the bodies are specified. Hence, they take mass, position, and velocity to be the state variables of the system and ignore color, chemical composition, and other factors.

A causal relation for this system is a functional relationship that describes the spatiotemporal variation of its state. The law of conservation of momentum qualifies as a causal relation on this understanding, viz.,

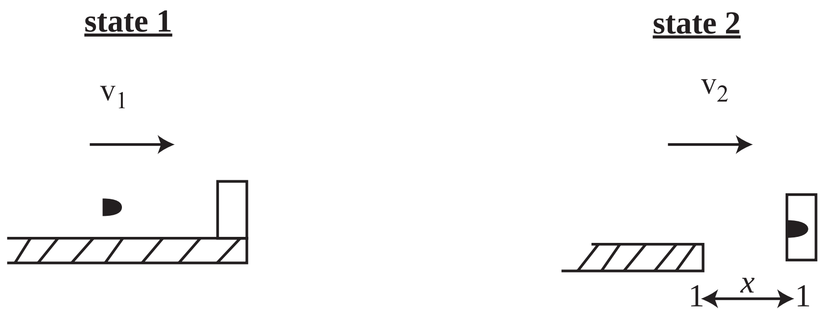


Figure 13.1 Momentum transfer as a causal process.

Momentum of system in state 1 = Momentum of system in state 2

$$m_b\, v_1 = (m_b = m_B)\, v_2$$

From information about the state of the system before impact, i.e., the values of m_b, m_B, and v_1, it is possible to predict the resultant velocity of the block-with-bullet. It also is possible to predict the way in which the velocity of the block-with-bullet varies with the incident velocity of the bullet. In addition, since the way in which the earth's gravitational attraction affects bodies is known, it is possible to predict the path followed by the block-with-bullet as it falls to the floor.

Henry Margenau pointed out that the concept "closed system" is an idealization.[2] No finite segment of the universe is completely unaffected by outside forces. Nevertheless, scientists have found that by characterizing physical systems by reference to a small number of state variables, these systems obey causal laws. This is true of the Hamiltonian and LaGrangean formulations of Newtonian mechanics, the theory of statistical mechanics, and the theory of electromagnetism.

Within these new physical theories, harmony prevailed with respect to causation, explanation, and prediction. To explain why state B_1 came about, one points out that state A_1 preceded it and that because A-states always are followed by B-states, A_1 caused B_1. Confirmed predictions of the occurrence of additional B-states reinforce convictions about putative causal and explanatory relations.

But suppose we apply a causal relation "Whenever A, then B" to a putative case of state A, and B does not follow. Does this prove that the relation is not a bona fide causal relation? Not necessarily. We may take the nonrecurrence of *B* as evidence, not against the causal relation, but against identification of the prior state as an *A*-state.

The system in question may have been incompletely isolated. Perhaps additional state variables need to be specified. In this situation, it still may be maintained that every *A* is followed by a *B*, but the putative *A*-state is disqualified from the set of *A*s because the proper state-function for the system contains additional variables. For instance, if iron bullets were used and if a very strong magnetic field was present, then the observed trajectory of a block-with-bullet would differ from the trajectory originally calculated. This result need not be taken to disqualify the momentum-conservation law from the class of causal relations, however. The original set of variables is insufficient to characterize the state of the system in this new instance. In this situation, the scientist either may seek to formulate a new causal relation for an expanded set of variables that includes magnetic field strength or may seek to reestablish conditions under which the original relation was observed to hold.

We can increase the precision with which we describe a physical system by increasing the number of variables taken into account. The limiting case would be to describe the entire state of the universe. But the more completely one specifies the state of a system and the forces acting on it, the less likely is a recurrence of a precisely similar state. Thus a gain in precision may be accompanied by a reduction in its applicability to real situations. By contrast, if we characterize a system by a "small number" of state variables, then we can identify recurrent states. However, the price paid is a certain looseness of fit between the mathematically characterized state and empirically determined magnitudes.

Notes

1. Moritz Schlick, "Causality in Everyday Life and in Recent Science," in *Causality* (Berkeley: University of California Press, 1932), 111.
2. Henry Margenau, "Causality and Modern Physics," *Monist* 31 (1931), 19.

III

Quantum Mechanics and the Regularity-between-States View

14

The Interpretation of Quantum Phenomena

The regularity-between-states view is a viable interpretation of causal relatedness for systems discussed in classical physics. Ambiguities arose, however, upon application of this view to quantum mechanical systems. The following sketch of this theory is based primarily on the work of Niels Bohr and Werner Heisenberg.

Heisenberg pointed out that applications of the quantum theory involve translations from a level of language in which observations are recorded, into a theoretical level, and back again.[1] A quantum mechanical interpretation of an experiment typically is a three-stage process:

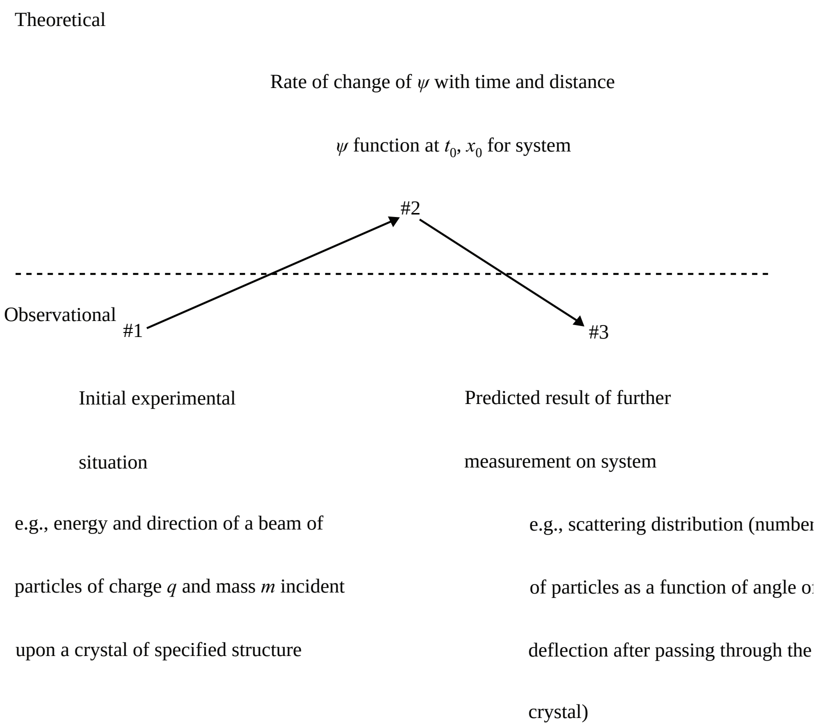
1. translation of an initial experimental situation into a probability function,
2. mathematical analysis of the variation of this function with time, and
3. translation of the results of calculation into a prediction of the results of a new measurement to be made of the system.

The Schrödinger version of the theoretical level specifies the spatiotemporal variation of state-function ψ. For one spatial dimension the state-function is

$$\psi = Ae^{2\pi i/h(Px-Et)},$$

where A is a constant, e is the base of the system of natural logarithms, $i = \sqrt{-1}$, h is Planck's constant, P is a momentum operator, and E is an energy operator.

A quantum mechanical interpretation may be represented as follows:



To interpret experimental results in this way would seem to presuppose that (1) observational and theoretical language levels can be distinguished; (2) the results of experiments can be expressed adequately without reference to the theoretical language level (Bohr and Heisenberg insisted that descriptions of experimental arrangements be given in the language of classical physics[2]); and (3) there are available semantical rules for translating statements of the observational level into the theoretical level, and conversely.

Max Born proposed the following semantical rule: if the proper function corresponding to the state of a subatomic particle is ψ, then $|\psi|^2\, dV$ is the probability that the particle is in the volume element dV.[3] The Born interpretation correlates the state of an individual particle with a probability only. But given an ensemble of particles each in state ψ, the Born interpretation correlates $|\psi|^2$ with experimentally determinable quantities.

In the case of an orbital electron, the Born interpretation correlates $|\psi|^2$ with a probability distribution of negative charge around the nucleus. Electron charge distributions calculated from the Schrödinger equation on the Born interpretation have been confirmed for numerous instances of chemical bonding in molecules. The patterns of inference that lead from values of $|\psi|^2$ to observable properties of molecules are complex. Nevertheless, these chains of inference do terminate in statements about experimentally determinable magnitudes.

In the case of scattering experiments, the Born interpretation correlates $|\psi|^2$ with the statistical distribution of an ensemble of particles after scattering. The final position of an individual particle after scattering may be determined experimentally by observing a scintillation counter or a photographic plate. Since the state of an individual particle is given by the ψ-function, it might seem that the path taken by an individual particle is calculable from the interpreted formalism. However, the Born interpretation enables a prediction to be made only of the *probability* that a particle is deflected through a specific angle. F. S. C. Northrup emphasized that a principal feature that distinguishes quantum mechanics from classical mechanics and relativity theory is that probability considerations are involved in the very definition of the state of the system in the quantum domain.[4]

In the case of transitions between electronic states in an atom, solutions of the Schrödinger equation take the form $\psi = c_1\psi_1 + c_2\psi_2 + c_3\psi_3 + \ldots$, where c_1, c_2, and c_3 are amplitudes of the eigen-functions for the respective quantum states to which a transition may lead. On the Born interpretation, the squares of the amplitudes (e.g., c_1^2) are correlated with the probability that after the transition the system is in the corresponding quantum state (e.g., state 1). These probabilities are not themselves observable. However, the Born interpretation correlates the values c_1^2, c_2^2, c_3^2, ... with the statistical distribution of resultant states in an ensemble of systems undergoing transitions from a given quantum state. This statistical distribution is an experimentally determinable magnitude.

The Born interpretation sanctions predictions about the outcome of measurements on ensembles of particles and sanctions judgments of probability about the outcome of measurements on individual particles. In either case, however, the Born interpretation is applicable only to quantum-systems-*as-observed*. It provides no description of quantum mechanical systems between observations made of them.

Niels Bohr superimposed upon the theoretical level a "pictorial level" that "depicts" the behavior of quantum mechanical systems between observations made on the system. Pictorial-level interpretations may be given in terms of the classical concepts of waves and particles. Alternatively, the interpretations may be given in terms of "paths" followed between observations and interactions in which energy and momentum are conserved. In either case, the component interpretations are mutually exclusive. For instance, the "picture" appropriate to a given experimental arrangement invokes either waves or particles, but not both. Moreover, as Bohr pointed out,

> we are presented with a choice of *either* tracing the path of a particle *or* observing interference effects.[5]

Which picture is appropriate depends on the nature of the experimental arrangement employed. Bohr labeled "complementary" the interpretations of experimental arrangements set up to measure the values of different conjugate properties of a quantum mechanical system.[6] These conjugate properties are subject to Heisenberg's uncertainty relations at the theoretical level—$\Delta p \ \Delta q \geq h/4\pi$ and $\Delta E \ \Delta t \geq h/4\pi$. De Broglie observed that if we wish to

> connect the energy and the momentum of a particle with the frequency and the wave-length which we associate with it—we are led to formulae in which Planck's constant *h* figures in an essential way; this shows that the duality of waves and corpuscles, the necessity of employing two pictures—apparently contradictory—to describe the same phenomena, is closely associated with the existence of the quantum of action.[7]

These two pictures are "apparently contradictory" because the concepts "energy" and "momentum" refer to masses localized in small regions of space and time, whereas the concepts "frequency" and "wavelength" refer to a wave that is infinitely extended in space and time. It is for this reason that Born declared that

> the imagination can scarcely conceive two ideas which appear less capable of being united than these two, which the quantum theory proposes to bring into such close connection.[8]

Within classical physics, the results of a collision process, for instance the collision of billiard balls, may be interpreted *both* by specifying the paths taken in space and time *and* by providing a causal analysis of the interaction. The causal analysis invokes the principles of conservation of energy and conservation of momentum.

Bohr maintained that spatiotemporal description and causal analysis are mutually exclusive interpretations within quantum physics. He declared that

> the impossibility of a separate control of the interaction between the atomic objects and the instruments indispensable for the definition of the experimental conditions prevents the unrestricted combination of space–time coordination and dynamical conservation laws on which the deterministic description in classical physics rests. In fact, any unambiguous use of the concepts of space and time refers to an experimental arrangement involving a transfer of momentum and energy, uncontrollable in principle, to fixed scales and clocks which are required for the definition of the reference frame. Conversely, the account of momentum and energy involves in principle a renunciation of detailed space–time coordination.[9]

C. A. Hooker and Henry Folse have argued that Bohr's primary emphasis was that spatiotemporal description and causal analysis are mutually exclusive interpretations in the quantum domain and that it was of secondary importance to Bohr that wave pictures and particle pictures are mutually exclusive as well.[10] I think this is correct. Nevertheless, Bohr frequently employed both the "spatiotemporal-versus-causal" contrast and the "wave-versus-particle" contrast in his analyses of quantum phenomena. For instance, in the case of the diffraction of electrons or photons at a single slit, Bohr noted that it is possible to set up an apparatus to measure accurately either the "path" of a particle or the diffraction pattern (actually, no measurements are made on the electron or photon between the slit and the photographic plate upon which the particle makes contact). If the diaphragm is rigidly connected to a massive frame, the transverse component of the position of the particle passing through the diaphragm can be determined by means of the position of the slit. The transverse component of the momentum of the particle then may be determined with only very low accuracy, although the *distribution* of particles striking the photographic plate can be calculated from the Schrödinger equation upon substitution of a specific distance between the diaphragm and the plate. At the pictorial language level, the wave picture is appropriate. The pattern on the photographic plate is "explained" by reference to the diffraction of a wave at the slit. Bohr pointed out that if the diaphragm instead is connected to the frame by an elastic spring,

> it should, in principle, be possible to control the momentum transfer to the diaphragm and, thus, to make more detailed predictions as to the direction of the electron path from the hole to the recording point.[11]

This increase in accuracy with which the path of a particle may be determined is achieved only at the expense of a blurring of the diffraction pattern. At the pictorial language level, the particle picture is appropriate. The electron is pictured as a particle whose path is determined by the momentum transfer at the slit.

The experimental arrangement in which an interpretation in terms of "wave language" is applied (the diaphragm is rigidly connected to the frame) does not lend itself to an interpretation in terms of "particle language," and the experimental arrangement in which an interpretation in terms of "particle language" is applied (the diaphragm is connected to the frame by an elastic spring) does not lend itself to an interpretation in terms of "wave language." Bohr placed a restriction on linguistic usage at the pictorial level—the results of a particular experiment may not be interpreted in terms of both a "particle picture" and a "wave picture." Thus, although the wave picture and the particle picture are mutually exclusive descriptions of what happens between measurements made on a system, the pictorial level of language does not contain logically incompatible claims. In the case of the passage of electrons through a single slit, for instance, the Bohr interpretation restricts the wave picture to one experimental arrangement and the particle picture to a second experimental arrangement.

Bohr insisted that individual pictorial-level interpretations are perfectly coherent. In particular, descriptions of quantum processes by reference to conservation laws are both consistent and appropriate. What is incoherent is the combining of mutually exclusive pictures of the same experimental arrangement. Although superposition of spatiotemporal description and causal analysis upon a particular experimental arrangement is a consistent interpretation within classical physics, this superposition is not a consistent interpretation within quantum mechanics.

Bohr defended a principle of complementarity which specifies that two mutually exclusive descriptions of a quantum process, in which two different experimental arrangements are employed, are complementary and combine to produce an exhaustive interpretation of the process. In the case of the diffraction of electrons by a single slit, the two mutually exclusive descriptions are (1) "electrons which have a certain position passing through a diaphragm slit produce a certain pattern of impacts on the photographic plate" and (2) "electrons which have a certain momentum in passing through a diaphragm slit transfer a certain momentum to the diaphragm." Based on the Bohr interpretation, these two descriptions combine to exhaust the range of possible knowledge about the passage of electrons through a single slit in a diaphragm.

The principle of complementarity stipulates that pictorial-level interpretations of quantum phenomena in wave language and particle language are of equivalent importance. This equivalence is based ultimately on the uncertainty relations, Einstein's equation $E = h\nu$, and De Broglie's equation $p = h\nu/c$. If the equations expressing energy and momentum in terms of frequency are substituted into the uncertainty relations $\Delta p \, \Delta q \geq h/4\pi$ and $\Delta E \, \Delta t \geq h/4\pi$, it is evident that there is a formal symmetry between the specification of spatiotemporal coordinates and the specification of wave-like properties. Bohr and Heisenberg emphasized that these symmetry properties are the foundation of quantum mechanics.

Max Jammer has called attention to von Weizsäcker's distinction between the "complementarity" championed by Pauli and the "complementarity" championed by Bohr.[12] Pauli predicated "complementarity" of concepts. Two concepts are "complementary" if the use of an experimental arrangement to measure values of one concept limits the precision with which values of the second concept can be measured. "Position" and "momentum" are "complementary" concepts because Heisenberg's uncertainty principle stipulates that $\Delta p \, \Delta q \geq h/4\pi$. In Pauli's usage, two concepts within the same classical picture may be "complementary." "Position" and "momentum" within the particle picture, and "energy" and "time" within the wave picture, are "complementary" in this sense.

Bohr agreed that these concepts are "complementary" in Pauli's sense. However, Bohr's application of the term "complementary" is to pictures of quantum mechanical systems between observations. Given a type of quantum mechanical process, its description in spatiotemporal categories and its description in causal categories are mutually exclusive and "complementary."

The wave picture and the particle picture of the process are likewise mutually exclusive and complementary. Bohr's principle of complementarity is a criterion of completeness. It stipulates that a complete interpretation of a type of quantum mechanical process includes both spatiotemporal description and causal analysis or both a wave picture and a particle picture. John Honner expressed Bohr's view as follows:

> our position as observers in a domain of experience, where unambiguous application of concepts depends essentially on the conditions of observation, permits the use of complementary descriptions, and demands them if description is to be exhaustive.[13]

Notes

1. Werner Heisenberg, *Physics and Philosophy* (New York: Harper, 1958), 46.
2. Heisenberg, *Physics and Philosophy*, 44; Niels Bohr, "Quantum Physics and Philosophy," in *Essays 1958–1962 on Atomic Physics and Human Knowledge* (New York: Interscience, 1963), 1–7.
3. Max Born, *Atomic Physics* (New York: Hafner, 1946), 92–100. See also Born, *Physics in My Generation* (New York: Pergamon Press, 1956).
4. F. S. C. Northrup, "Introduction," in Werner Heisenberg, *Physics and Philosophy*, 7.
5. Niels Bohr, *Atomic Physics and Human Knowledge* (New York: Wiley, 1958), 46.
6. Ibid., 72; Bohr, *Atomic Theory and the Description of Nature* (Cambridge: Cambridge University Press, 1961), 114.
7. Louis de Broglie, *Physics and Microphysics* (New York: Pantheon, 1955), 125.
8. Born, *Atomic Physics*, 84.
9. Bohr, *Atomic Physics and Human Knowledge*, 72.
10. C. A. Hooker, "The Nature of Quantum Mechanical Reality," in *Paradigms and Paradoxes*, ed. R. Colodny (Pittsburgh: University of Pittsburgh Press, 1972), 140, 158; Henry Folse, *The Philosophy of Niels Bohr* (Amsterdam: North-Holland, 1985), 222–57.
11. Bohr, *Atomic Physics and Human Knowledge*, 45.
12. Max Jammer, *The Conceptual Development of Quantum Mechanics* (New York: McGraw-Hill, 1966), 355.
13. John Honner, "The Transcendental Philosophy of Niels Bohr," *Studies in History and Philosophy of Science* 13 (1982): 27.

15

Causality and the Three Levels of Language in Quantum Mechanics

The observational level of language in quantum mechanics records the results of manipulating instruments. There is no new challenge to the regularity-between-states view at this level. If the regularity-between-states view is applicable to the systems of classical physics, then it is also applicable to the observational level of quantum theory.

Bohr maintained that a necessary condition of the unambiguous communication of experimental results is that these results be expressed in the language of classical physics. He declared that

> the description of the experimental arrangement and the recording of observations must be given in plain language, suitably augmented by the usual physical terminology. This is a simple logical demand, since by the word "experiment" we can only mean a procedure regarding which we are able to communicate to others what we have done and what we have learnt.[1]

Bohr sometimes claimed that the concepts of classical physics are mere refinements of the concepts of everyday language. This is a dubious claim at best. "Inertial motion" (Newton), "tangentially-acting forces" (Oersted), and "electromagnetic induction" (Faraday) are not straightforward generalizations of everyday experience. However, the thesis that operations of measurement must be described in the concepts of classical physics may be defended without presupposing that the concepts of classical physics are mere refinements of the concepts of everyday language.

In support of this position, Heisenberg maintained that the Newtonian–Kantian understanding of "space," "time," and "causality" is a necessary condition for the very possibility of science.[2]

Measuring instruments that are set up for the purpose of observing atomic events are sufficiently heavy to allow an account of their positions and momentums in the terms of classical physics. If the purpose of the experiment is to measure spatial coordinates or time, then rigid, massive measuring rods and clocks are required. If the purpose of the experiment is to measure momentums or energies, the measuring instrument must contain a freely moveable part to which the conservation laws may be applied. In either case, the experimental arrangement and the results of observation must be described in the language of classical physics. The Newtonian concepts are necessary for the description of experimental arrangements and results, but they are inappropriate concepts for defining the state of a quantum mechanical system. Heisenberg pointed out that

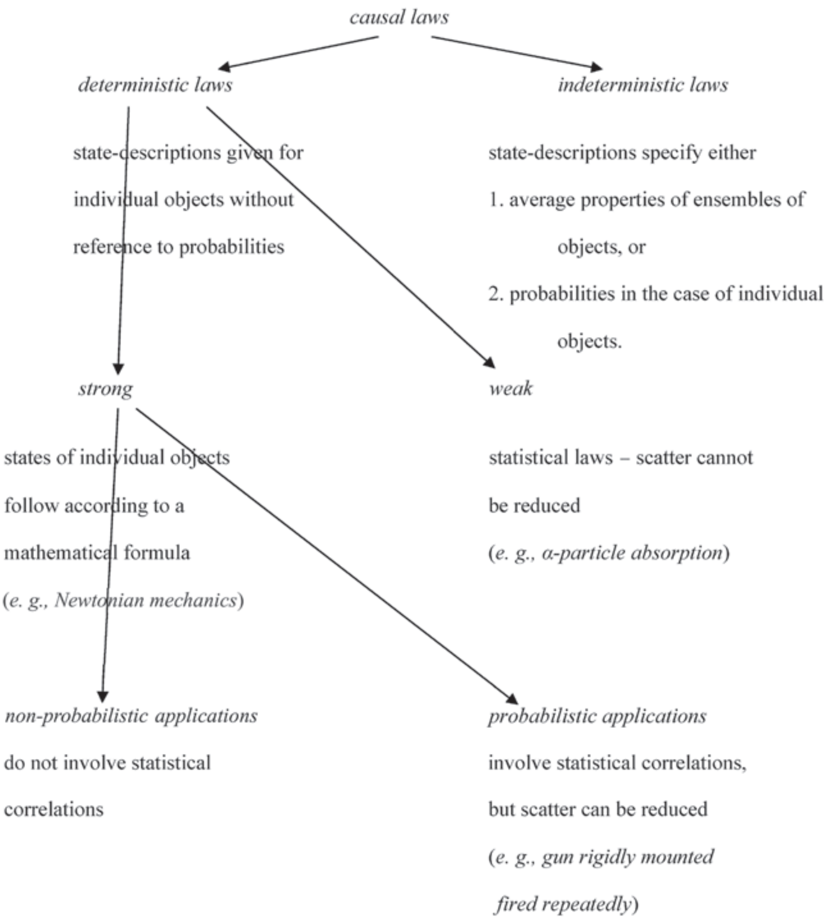
> what Kant had not foreseen was that these *a priori* concepts can be the conditions for science and at the same time can have only a limited range of applicability.[3]

It is the position of Bohr and Heisenberg that the description of experimental arrangements and results in the concepts of classical physics is a necessary condition of knowledge of the quantum domain. They acknowledged, however, that the range of applicability of the classical concepts does not extend to the theoretical level of the quantum theory. The states of quantum mechanical systems cannot be specified by reference to the concepts of classical physics.

Applications of the regularity-between-states view to statements at the theoretical level involve statistical considerations, since the state of a quantum mechanical system is defined in terms of probabilities. Nevertheless, successive states of quantum mechanical systems are governed by functional relations—instantiations of Schrödinger's equation or Heisenberg's matrices—that satisfy the regularity requirement. Statements about successive states of a quantum mechanical system are "causal" on the regularity-between-states view. The taxonomy of causal laws on page 79 is consistent with this position.

Causal laws are functional relations that, given information about the state of a system in one region of space and time, imply information about the state of the system in some other region of space or time. Causal laws may be either deterministic or indeterministic. A causal law is deterministic if its state variables may be specified without reference to probabilities. A causal law is indeterministic if its state variables are specified by reference to probabilities.

A Taxonomy of Causal Laws



The laws of Newtonian mechanics are causal laws because they satisfy the requirement of inferability. Newtonian mechanics is deterministic because its state descriptions are given for individual bodies without reference to considerations of probability. Physical systems such as billiard balls, pendulums, and planets are characterized by specifying individual positions, velocities, and masses.

There are strong and weak versions of deterministic laws. Strong deterministic laws state invariant correlations among states of physical systems. The laws of gravitational attraction and electrostatic attraction are strong deterministic laws. Weak deterministic laws state

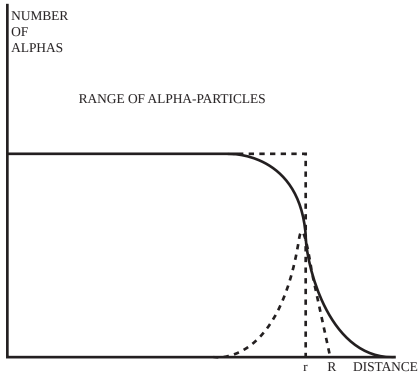


Figure 15.1 α-Particle absorption.

statistical correlations, even though the appropriate state descriptions make no reference to probabilities. An example is α-particle absorption (Figure 15.1).

α-Particles of given energy and direction, incident upon a sheet of aluminum, are absorbed at varying distances inside the sheet. The law governing this absorption is a statistical law, but since the state description is accomplished without reference to probabilities, this law also is a deterministic law. A further complication is that applications of strong deterministic laws may yield a statistical correlation of results. The laws of Newtonian mechanics, for instance, may be used to predict the distribution of hits on a target produced by firing a relatively rigidly mounted gun. For this experimental arrangement, the extent of scattering can be diminished by improving the mechanism and mounting of the gun. This case contrasts with the case of α-particle absorption. In the latter case, the extent of scattering is irreducible (for a given α-particle energy).

The laws of quantum mechanics also are causal laws because they satisfy the requirements of the regularity-between-states view. For example, given information about the energy, mass, charge, and direction of a beam of subatomic particles incident upon a crystal of specified structure, one can calculate from these laws the scattering distribution of the particles of the beam (viz., the relative number of particles as a function of angle of deflection).

The laws of quantum mechanics are indeterministic, however. State variables are specified for average properties of ensembles of particles. Alternatively, if a single particle is under consideration, state description is in terms of probabilities. Whereas complete knowledge of a classical system includes statements about the numerical values of each and every state variable, complete knowledge of a quantum mechanical system includes only probability values for the state variables. In the case of a quantum mechanical scattering experiment, one can predict for an individual particle only that there is a certain probability that it will emerge at a particular angle after scattering. Within the context of quantum theory, there is no way to replace the statement of probability by a "more exact" state description. Quantum theoretical state description is, in principle, a specification of probabilities.

The regularity-between-states view of causality has several advantages over theories that begin with distinct events and seek to elaborate conditions under which one event is a cause of a second event. The regularity-between-states view provides a unified treatment of deterministic and indeterministic generalizations. In addition, by linking causal relatedness to functional relations between states of isolated systems, one can sidestep questions about nonlocality, action-at-a-distance, and backward causation. What counts for causal relatedness is the success of inferences made about changes in the states of physical systems.

At the pictorial level, Bohr restricted causal analysis to one of the two mutually exclusive interpretations of the behavior of quantum mechanical systems. Spatiotemporal description is the other interpretation. Bohr held that the two interpretations are complementary. A complete picture requires both interpretations.

Consider once again the passage of electrons through a single slit to a photographic plate. At the theoretical level, one can predict from the results of measurement of the energy and direction of the electron beam (state 1) the distribution of impacts on the plate (state 2). At the pictorial level, there are two pictures of "what happens between" measurements. The "causal" picture depicts an exchange of energy and momentum at the diaphragm subject to the respective conservation laws. On the causal picture, the diaphragm is free to move. It participates in a momentum–energy exchange. The spatiotemporal picture specifies paths taken by electrons through the slit to the plate. On this picture, the diaphragm is taken to be rigidly anchored. The thesis of complementarity requires that both pictures be included to achieve complete knowledge of the diffraction of electrons at a single slit.

Some critics have rejected the idea that a pictorial level is required for the interpretation of quantum phenomena. Pascual Jordan directed attention to the successful predictions of quantum effects achieved by application of the Schrödinger equation. He held that since the task of physics is to describe and predict the results of observations and experiments and since pictorial-level language is not involved in this process, it may be safely eliminated.[4] According to Jordan, Bohr is wrong to claim that superimposing complementary pictures on the calculations of the theoretical level contributes to our understanding of quantum processes.

Henry Margenau also sought to dispense with the pictorial level of interpretation. He criticized Bohr for maintaining that physicists should accept both causal descriptions in terms of abstract states and complementary descriptions in terms of classical waves and particles. Bohr leaves us impaled on the horns of a dilemma—strictly speaking, classical concepts are inapplicable to quantum mechanical processes, and yet we continue to apply them. According to Margenau, a better strategy is to remain at the theoretical level of interpretation.

Margenau attributed "physical reality" to the Born interpretation of the ψ-function. He did so on the basis of the principle that physical reality includes the class of all "verifacts."[5] A "verifact" is a construct that (1) participates in circuits of confirmation that begin and terminate with observables and (2) satisfies a set of regulative principles. Among these principles is the requirement that the construct be embedded in a theory from which laws of causal form may be derived. A causal law, in turn, is a functional relation in which

> physical systems be described in terms of states which are self-unfolding in a determinate manner; that the state of affairs at time *t* be sufficient for a prediction of the state . . . at any other time *t*.[6]

The ψ-function qualifies as a verifact on Margenau's criteria, despite the fact that it includes the imaginary term "$\sqrt{-1}$." It does so on the basis of the Born interpretation that correlates $|\psi|^2$ with observables such as scattering distributions and energy-level transitions in atoms.

Notes

1. Niels Bohr, "Quantum Physics and Philosophy," in *Philosophy in the Mid-Century*, ed. R. Klibansky (Firenze: La Nuova Italia Editrice, 1958), 308.
2. Werner Heisenberg, *Physics and Philosophy* (New York: Harper, 1958), 90.

3. Ibid., 90.
4. Pascual Jordan, *Physics of the 20th Century* (New York: Philosophical Library, 1944), 46.
5. Henry Margenau, "Reality in Quantum Mechanics," *Philosophy of Science* 16 (1949): 287–302.
6. Margenau, "Advantages and Disadvantages of Various Interpretations of Quantum Theory," *Physics Today* 10 (October 1954): 13.

16

Philipp Frank on Causality and Inferability

Margenau's position implements what might be called an inferability theory of causality. Philipp Frank was an influential advocate of this theory. He maintained that a causal relation is a functional relation that, conjoined with information about a physical system in one region of space and time, permits an inference about its state in some other region of space and time.[1] Frank maintained that causal laws are differential equations of the form

$$\frac{\mathrm{d}\alpha}{\mathrm{d}t} = F_k\left(\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_k\right),$$

where $\alpha_1, \ldots, \alpha_k$ are the state variables of the system and F_k is some function of these variables.[2]

In order for a relation of this form to qualify as a causal relation, three conditions must be satisfied. We must be able to specify the state of the system with a manageably small number of variables. We must be able to determine experimentally the values of these variables for some initial state at a given time. And the function F_{k} must be of mathematical form sufficiently simple to permit calculation of further values of the state variables.

Many important scientific laws are causal relations in the above sense. Newton's laws of motion and Mendel's laws of inheritance are examples. The former permits deduction of the spatial location of planets, billiard balls, and pendulums as a function of time; the latter permits deduction of the characteristics of succeeding generations from information about the characteristics of given populations.

Boyle's law, by contrast, is not a causal law in the above sense. Boyle's law states that at constant temperature the volume of a given mass of gas is inversely proportional to its pressure. Although Boyle's law does express a functional relation between states of a system, it does not enable prediction of the way in which the state of a gas varies with spatial location or time.

The mathematical form of a causal relation is such that the temporal variation of the state of a system is determined only by the values of the state variables $\alpha_1, \ldots, \alpha_k$ at a given time. If a particular set of state variables is labeled "state *A*," and if state *A* is followed by some other state *B*, then every time *A* recurs, it will be followed by *B*. Thus, Frank's position on causal relatedness is a variant of Hume's regularity view. The inferability view connects causality and explanation by identifying causes and effects with states of physical systems, causal laws with functional relations of a certain form between these states, and causal explanations as deductive arguments from premises that include causal laws and initial state descriptions.

But suppose we apply a causal relation "Whenever *A*, then *B*" to a putative case of state *A*, and *B* does not follow. This does not prove that the relation is not a causal relation. We may take the nonoccurrence of *B* as evidence, not against the causal relation, but against identification of the prior state as an *A*-state.

Frank observed that various strategies may be pursued to protect the causal status of a relation.[3] Perhaps the values $\alpha_1, \ldots, \alpha_n$ had been incorrectly determined, so that state *A* had not been realized after all, or perhaps the system was incompletely isolated in this instance, in which case additional state variables need to be specified. On either option, it still is maintained that every *A* is followed by a *B*, but the putative *A*-state is disqualified from the set of *A*s because the proper state-function for the system contains additional variables. In the case of the previously mentioned block-with-embedded-bullet illustration, if iron bullets were used and if a very strong magnetic field were present, then the observed trajectory of a block-with-embedded bullet might differ from the trajectory originally calculated. This result need not be taken to disqualify the momentum-conservation law from the class of causal relations, however. The original set of variables is insufficient to characterize the state of the system in this new instance. In this situation, the scientist may seek to formulate a new causal relation for an expanded set of variables that includes magnetic field strength or may seek to reestablish conditions under which the original relation was observed to hold.

A third strategy to protect the causal status of a relation is to introduce sufficient complexity into the F_k function, so that the recalcitrant data fits the revised relation. This strategy was pursued by the eighteenth-century mathematician Clairaut, who dealt with recalcitrant data on the motion of the moon's line of apsides by adding a $1/R^4$ term to the law of gravitational attraction.

Frank conceded that conventionalist strategies of this sort always can be implemented. He pointed out that we have a choice between making causal relatedness precise and inapplicable or vague and factual.[4] We can increase the precision of a causal relationship by increasing the number of state variables and forces considered. The limiting case would be to describe the entire state of the universe. But the more completely one specifies the state of a system and the forces acting on it, the less likely is the recurrence of a precisely similar state. By contrast, if we characterize a system by a small number of state variables, then we can identify recurrent states. However, the price paid is a certain looseness of fit between the mathematically characterized state and empirically determined magnitudes.

Frank maintained that most scientists accept a "principle of causality" in the following form: it is possible to specify state variables for physical systems

> in such a way that a small number of such variables suffice to express the results of experiments in laws of causal form.[5]

The principle of causality, thus formulated, is vague. It is left unspecified how many variables may be included before the "small number" stipulation is exceeded. It also is unclear how closely experimental results must conform to the deductive consequences of the laws. Differences of opinion most likely will arise over the status of specific candidates for causal laws.

Frank's inferability view fails to reproduce the temporal asymmetry of causal relations. He defined a "causal law" to be a functional relation that permits inferences about a system from one region of space and time to another. Newton's laws of motion qualify as causal laws, not only because they permit inferences to future states, but also because they permit inferences to past states. However, it would be counterintuitive to claim to have explained an event as an effect of subsequent conditions. On our preanalytic understanding of causal relatedness, causes precede their effects.

Notes

1. Philipp Frank, *Philosophy of Science* (Englewood Cliffs: Prentice-Hall, 1957), 264.
2. Ibid., 266.
3. Ibid., 268–71.
4. Ibid., 274.
5. Ibid., 288.

IV

Protests against the Identification of Causality and Regularity

17

Ducasse and the Singularity Definition

Scientists have redefined "causality" to refer to "functional relations between states of (reasonably isolated) physical systems." However, this is not what the guilty motorist means when she admits to have "caused" an accident by pressing the accelerator pedal instead of the brake. In an essay published in 1926, Curt Ducasse challenged the dominant position that takes causal relations to be regularities between events or states of physical systems.[1] He maintained that Hume's preoccupation with regularities was misguided. According to Ducasse, when we seek a cause, we do not seek an invariable correlation. Suppose my car engine fails at an intersection. If I ask what caused this failure, I am not asking about a constant conjunction of which this incident is a member. Rather, I wish to know what was the *single difference* among attendant circumstances from the case in which the car was running. Ducasse insisted that causal inquiry is inquiry about single events and not regularities. He defined a cause to be a particular change in the immediate environment that occurred just before the effect in question. Consider the motion of a brick (*C*) that strikes and shatters a pane of glass (*K*). Ducasse declared that *C* is the cause of *K* if:

1. the change *C* occurred during a time and through a space terminating at the instant *I* at the surface *S*,
2. the change *K* occurred during a time and through a space beginning at the instant *I* at the surface *S*,
3. no change other than *C* occurred during the time and through the space of *C*, and no change other than *K* during the time and through the space of *K*.[2]

He provided the following informal presentation of the definition:

> More roughly, but in briefer and more easily intuited terms, we may say that the cause of the particular change *K* was such particular change *C* as alone occurred in the immediate environment of *K* immediately before.[3]

Ducasse emphasized that his definition

> defines the cause of a particular event in terms of but a single occurrence of it, and thus in no way involves the supposition that it, or one like it, ever has occurred before or ever will again. The supposition of recurrence is thus wholly irrelevant to the meaning of cause.[4]

It might seem that Ducasse's "singularity view" of causation either is inapplicable to concrete situations or involves a tacit appeal to the regularity thesis. If we take Ducasse's definition literally, then causal analysis requires the specification of *every* change in the immediate environment during the period in question. This would include the motions of every molecule, atom, electron, and wave in the (vaguely circumscribed) immediate environment. This is required in order to show that *C* and *K* are the *only* changes in this environment. In the precise form in which Ducasse formulated his definition of "cause," causal judgments cannot be made.

However, we can apply Ducasse's definition if we agree to characterize the immediate environment by specifying only a manageably small number of changes. On this "loose" interpretation, causal analysis can be performed, even though the term "cause" is a vague term. But how can one decide which changes to select for examination? Ducasse himself suggested the following complication—suppose that

> at the instant a brick strikes a window pane, the pane is struck . . . by the air waves due to the song of a canary near by.[5]

One way to single out the appropriate candidates for causal status is to examine what has happened in other contexts. Substantial impacts with bricks have shattered panes of glass; sound waves from the songs of birds have not done so. Does the successful application of the singularity view require an appeal to experienced regularities?

Ducasse thought not. He suggested that specific instances of causation are directly observed in the course of ordinary experience. According to Ducasse,

> every person has *perceived*—and I say *perceived*, not *inferred*—that, for example, a particular tree branch was *being caused to bend* by

> a particular bird's alighting on it; that a particular bottle was *being caused to break* by the fall on it of a particular rock; that a particular billiard ball was *being caused to move* by a particular other billiard ball's rolling against it . . . etc.[6]

David Hume was wrong. In some cases we observe, not just a sequence of events, but a causal relation between them.

Ducasse noted that the regularity-between-states view of Margenau and Frank requires considerable abstraction in order to establish "causal relations" between states of a physical system. The state of a physical system is a construct, the isolation of a system is an idealization, and the successive states of a system become empirically significant only through appeal to "operational 'bridges' between the conceived and the perceived."[7]

> Thus, when causation is defined as in theoretical physics, concrete cases of it are *not perceived* to be such but are *inferred* to be such.[8]

Ducasse acknowledged that scientists successfully have employed functional relations to predict subsequent states of physical systems. He insisted, however, that

> the problem of giving a "correct" definition of the causal relation is that of making analytically explicit the meaning which the term "cause" has in actual concrete phrases that our language intuition acknowledges as proper and typical cases of its use.[9]

Since "proper and typical cases of its use" include applications of such verbs as "to break," "to bend," "to heat," "to kill," "to twist," "to melt," "to prevent," "to steer," "to remind," and "to irritate,"[10] Ducasse concluded that his proposed "singularity definition" is the correct definition of the causal relation.

Notes

1. Curt J. Ducasse, "On the Nature and the Observability of the Causal Relation," *Journal of Philosophy* XXIII (1926); Reprinted in *Truth, Knowledge and Causation* (New York: Humanities Press, 1969), 1–14.
2. Ducasse, *Truth, Knowledge and Causation*, 3–4.
3. Ibid., 4.
4. Ibid., 6.
5. Ibid., 11.
6. Ducasse, "Causation: Perceivable? Or Only Inferred?" *Philosophy and Phenomenological Research* XXVI (1965); in *Truth, Knowledge and Causation*, 26.

7. Ibid., 23.
8. Ibid., 23.
9. Ducasse, "On the Nature and the Observability"; *Truth, Knowledge and Causation*, 1.
10. Ducasse, "On the Analysis of Causality," *Journal of Philosophy* LIV (1957); *Truth, Knowledge and Causality*, 15.

18

Cause as Sufficient Condition

Ducasse maintained that the cause of an event is that condition which makes the difference between its occurrence and nonoccurrence. Such a condition is a condition sufficient to produce the event (in the given context).

The same type of event can be a sufficient condition of a type of effect on one occasion and merely a necessary condition of that type of effect on another occasion. Consider a grandfather clock. Amy has stopped the swinging of its pendulum. Beth restores its timekeeping function by displacing and releasing the pendulum. Her actions qualify as a sufficient condition of the restoration. On another occasion, the clock has unwound and stopped. Beth's displacing and releasing the pendulum is at best a necessary condition in this instance. It also is necessary to wind the clock, thereby raising the weight whose slow downward motion supplies the energy that maintains the swings of the pendulum.

Some theorists have maintained that causes *always* are sufficient conditions. On this view, the restorative cause in the case of the unwound and stopped clock is the conjunction of raising the weight and displacing and releasing the pendulum. The cause-as-sufficient-condition thesis has been defended, in some passages at least, by John Stuart Mill,[1] R. B. Braithwaite,[2] Carl Hempel,[3] and Karl Popper.[4] On the "cause-as-sufficient-condition" thesis, the cause of a match's lighting is the total set of conditions required for ignition. These conditions include the presence of oxygen, the chemical composition of the match, and the act of striking the match on an appropriate surface.

In science, and in everyday experience, we do attribute causal status to necessary conditions such as the striking of matches and inoculations against anthrax. But the sufficient-condition theorist is not required to defer to this usage. She may claim that the "cause-as-sufficient-condition" thesis is an *explication* of "cause" that is put forward as a refinement that corrects ordinary usage.

If the "cause-as-sufficient-condition" thesis is correct, then the following conditional claims are true:

1. "If x is the cause of E, then x is a sufficient condition of E" and
2. "If x is a sufficient condition of E, then x is the cause of E."

The sufficient-condition theorist who accepts conditional claim 1 is unwilling to terminate causal inquiry with the identification of merely necessary conditions.

In some instances, the sufficient-condition theorist will need to stipulate that certain conditions are absent. John Stuart Mill called attention to the case of arsenic poisoning (cf. Chapter 8, section Method of Difference). According to Mill, it would be incorrect to attribute the death solely to the ingestion of arsenious acid. Swallowing the poison is not a sufficient condition of death. A description of the genuine (sufficient condition) cause includes the stipulation that there was no hydrated peroxide of iron in the victim's body. Hydrated peroxide of iron is an antidote to arsenic poisoning.

The "cause-as-sufficient-condition" thesis can be defended in such a way that it becomes vacuous. Given the occurrence of event E at time t, it would be safe, but uninformative, to select as its sufficient condition the entire state of the universe at time $t - \Delta t$ (i.e., the state of the universe as viewed from earth, and not from, e.g., α Centauri). Since this state of the universe (U) did precede E, the cause of E was realized. The only way to disprove the causal attribution would be to show that states similar to U were not followed by events similar to E.

But how could this be achieved? The more detailed the description of state U, the less likely there is a recurrence of that state. By selecting U to be the sufficient condition of E, we shield the causal attribution from the possibility of falsification. The price paid, however, is to render causal analysis ineffective. In response to the question "What was the cause of E at t?" the answer given is "Whatever was the prior state of the universe at $t - \Delta t$."

On the sufficient-condition interpretation, causal analysis can be effective only if it is restricted to a manageably small number of relevant conditions. The sufficient-condition theorist has a choice. She can render the interpretation of "sufficient condition" precise, thereby making causal analysis ineffective. Or she can accept a lack of precision in the characterization of "sufficient condition," thereby rendering causal analysis manageable and falsifiable.

There are good reasons to believe that conditional claim 2 is false. Some sufficient conditions of effects are not causes of these effects. The standard countercase in the literature is the flagpole case. The length of a vertically mounted flagpole's shadow and the position of the sun comprise a sufficient condition of the flagpole's length (Figure 18.1).

Given length of shadow s and angle α, length of pole l follows from the axioms of plane geometry. However, this sufficient condition of the flagpole's length clearly is not the cause of this length. Rather, the cause is the manufacturing process that created the pole.

Other countercases may be advanced. Consider a flame test performed on a sample containing the barium ion (Ba^{++}). A green color of a specific wavelength is produced. A sufficient condition of this effect is the presence of Ba^{++} in the sample and its exposure to the flame. Scientists accept this sufficient condition as the cause of the green color. But scientists also use flame tests for the purpose of identification. Given that a sample of unknown composition is placed in a flame and yields this green color, they conclude that Ba^{++} is present in the sample. The placing of that sample in a flame such that this green color is produced is a sufficient condition of the presence of Ba^{++} in the sample. But this sufficient condition clearly is not the cause of the presence of Ba^{++} in the sample. Conditional claim 2 is false.

Causes as INUS Conditions

J. L. Mackie has developed a modified version of the sufficient-condition approach. Mackie directs our attention to the investigation

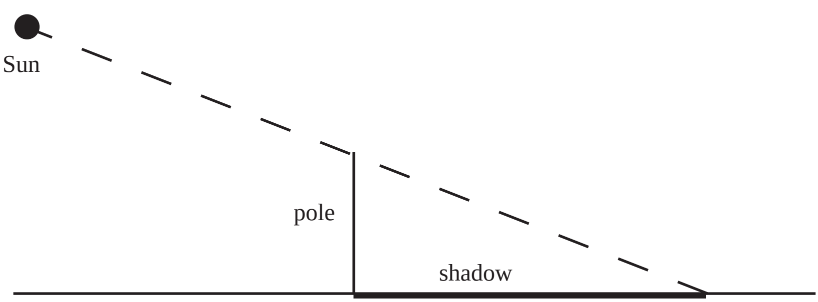


Figure 18.1 A flagpole and its shadow.

of a fire. The investigator concludes that a short circuit in the electrical wiring caused the fire. This is a paradigm case of causal inquiry, but the short circuit is neither a necessary condition nor a sufficient condition of the fire. Nevertheless, it is appropriate to refer to the short circuit as the "cause" of the fire. Mackie suggested that the short circuit is an INUS condition of the fire, viz., an *insufficient* but *necessary* part of a set of conditions that is itself *unnecessary* but *sufficient* for the fire. Mackie defined an INUS condition as follows:

> A is an INUS condition of a result P if and only if, for some X and for some Y, (AX or Y) is a necessary and sufficient condition of P, but A is not a sufficient condition of P and X is not a sufficient condition of P.[5]

In this definition, X is the conjunction of conditions that, together with A, constitutes one sufficient condition of P, and Y is the disjunction of other sets of conditions that also are sufficient conditions of P (e.g., conditions that include a lightning strike, an overturned kerosene stove, a carelessly thrown lighted match, etc.).

Given this definition of an INUS condition, a condition is "at least an INUS condition" provided that it has one of the following forms—(AX or Y), (A or Y), AX, or A. Mackie maintained that four conditions must be satisfied for A to be the cause of P:

1. A is at least an INUS condition of P;
2. A is present on the occasion of P;
3. the conditions represented by X, if any, in the necessary and sufficient condition were present on this occasion; and
4. every disjunct in Y that does not contain A as a conjunct was absent. That is, no other sufficient condition of P was present on this occasion.[6]

The investigator is correct that a short circuit caused the fire provided that

1. a short circuit is at least an INUS condition,
2. a short circuit did take place,
3. the other conditions that conjoined with the short circuit constitute a sufficient condition of fire were present, and
4. no other sufficient condition of fire was present.

Mackie's analysis is not subject to the flagpole countercase. The length of the flagpole is "at least an INUS condition," and also the cause, of the length of its shadow, viz., let

> *A* be the length of the vertically mounted flagpole *l*; *X* be the position of the sun, absence of light-absorbing material between flagpole and shadow; *Y* be the [length l^*, position of sun S^*, absence of absorbing material] or [length $l^\#$, position $S^\#$, absence of absorbing material] or . . . ; and *P* be the shadow length *s*.

(*AX* or *Y*) is "at least an INUS condition" of *P* because
(*AX* or *Y*) is a sufficient condition of *P*,
(*AX* or *Y*) is a necessary condition of *P* [given a full elaboration of those (theoretically infinite) correlations of flagpole lengths and positions of the sun associated with shadow length *s*],
A is not a sufficient condition of *P*, and
X is not a sufficient condition of *P*.

Moreover, *A* is the cause of *P* because

1. *A* is at least an INUS condition of *P*,
2. *A* is present on the occasion of *P*,
3. the factors represented by *X* were present on the occasion of *P*, and
4. every disjunct of *Y* was absent on the occasion of *P*.

However, the shadow length does not qualify as the cause of the flagpole's length, viz., let A^* be the shadow length *s*, X^* be the position of sun, vertical placement of flagpole, Y^* be the details of the flagpole-manufacturing process, and P^* be the length of the flagpole.

A^* does qualify as "at least an INUS condition" of P^*, but A^* is *not* the cause of P^* because it is false that Y^* was absent on the occasion of P^*.

The INUS interpretation is useful in many causal contexts. However, this modified sufficient-condition analysis fails to account for causal relations in processes governed by statistical laws. Consider the absorption of an α-particle by a sheet of aluminum. It has been determined experimentally that the following type of absorption curve results upon the incidence of α-particles of given energy upon a sheet of aluminum (Figure 18.2).[7]

Suppose a single α-particle of energy *E* is emitted and is absorbed at distance R within the plate. The absorption at R is caused by placing the aluminum in the path of the α-particle. But the impact of the α-particle on the plate does not qualify as "at least an INUS condition" of the absorption: viz., let *A* be the α-particle of energy *E* that impacts an aluminum plate, *X* the 90° orientation of plate-to-particle path, absence of strong electromagnetic fields, and *P* be the particle absorbed at distance R along with 15 percent of the total.

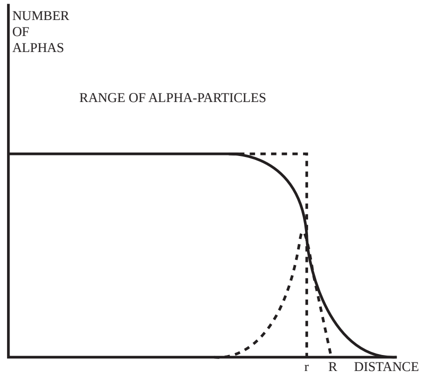


Figure 18.2 α-Particle absorption in aluminum.

AX is not a sufficient condition of *P* since only 15 percent of α-particles of energy *E* are absorbed at distance R.

There are good reasons for retaining causal language in the discussion of statistical correlations. This issue is addressed in Chapter 19. The INUS-condition interpretation is not effective in the context of events governed by statistical laws. However, it does provide a perspicuous account of causal relations in many contexts. Within such contexts, the adequacy of an attribution of causal relatedness made by INUS-condition analysis is contingent upon the correct elaboration and exclusion of the sufficient conditions incorporated under the term *Y*.

If the above analysis is cogent, then "*x* is a cause of *y*" is equivalent to neither "*x* is a necessary condition of *y*" nor "*x* is a sufficient condition of *y*." A weaker claim is plausible, however. *Some* causes are necessary conditions of their effects, and *some* causes are sufficient causes of their effects. This suffices for the purposes of science. Scientists persevere in a search for causes despite the failure of philosophers to prove either that all causes are necessary conditions or that all causes are sufficient conditions.

It remains to be specified what counts as a "condition." Of course a sufficient condition is some "arrangement" such that the effect in question must occur in its presence. But what sorts of "constituents" are relevant to the arrangement?

Clearly, properties qualify. A list of constituents sufficient for the lighting of a match includes the dryness of the match and the roughness of the surface scratched. And a list of constituents sufficient for the motion of a body includes its mass and charge. In addition, the presence of a substance in some region of space during a particular time interval may qualify. For instance, the presence of oxygen in a warehouse may be part of the arrangement of constituents that constitute the sufficient condition of a fire.

The account that a questioner accepts as an explanation of a phenomenon depends on the type of information sought. Sometimes a recitation of properties is preferred to the description of an "event." Consider a simple case. A block of iron dropped into water sinks. An observer poses the question "Why did the block sink?" Suppose the explanation proposed is "The block sank because it was released above the surface of the water." The questioner may reject this proposal on the grounds that a similar type of explanation would account for the floating of an ice cube in the water. "What I want to know," she states, "is what makes the difference between releasing the iron block above the water and releasing the ice cube." A new account in terms of relative densities and Archimedes's law may be accepted as providing the desired explanation.

Consistent with this approach, the questioner may reject as a candidate for the cause of the block's sinking the act of releasing it and accept as cause the relative densities of block and water. This would be to assign as cause a relation between the properties of the respective objects.

Relations too may be constituents of causal attributions. For instance, attractive forces that specify the rate of change of position over time often are cited as causes of the changes of motion of bodies. A sufficient condition of the fall of an apple includes the relation of gravitational attraction. This is the case even when reference to a force falls short of providing an acceptable explanation of an effect. For example, the gravitational attraction between the earth and a leaf is correctly labeled a "cause" (*qua* necessary condition) of the fall of the leaf from the top of a tall building. However, there is no detailed explanation available for the path of the falling leaf.

Philosophers of science have emphasized the interdependence of the concepts "cause" and "law." Laws of nature often are cited in the identification of causes. Relations that merely "save appearances" do not qualify. The standard example of saving appearances is the epicycle-deferent model of planetary motions championed by Ptolemy. Such models were used successfully to predict the positions of planets against the background stars. However, no one argued that the motion of Mars through Capricorn is caused by the motion of an epicycle circle as it revolves around a deferent circle (Figure 18.3).

More controversial is causal attribution in the case of processes whose growth conforms to the requirements of the Fibonacci series—1, 1, 2, 3, 5, 8, 13, 21, ... Is this series a cause of the axial distribution of leaves as they emerge in certain monocotyledonous plants species? Is it a cause of the observed rate of crystal growth from initially supersaturated solutions? Is it a cause of population growth in successive generations of haploidally reproducing insect societies (Figure 18.4)?

I suspect that most scientists would deny that processes that conform to the Fibonacci series qualify as causal—the series merely "saves the appearances" of certain naturally occurring growth patterns. On the other hand, a few committed Pythagoreans may accord lawful status to the series on the grounds that it must be significant that it fits such diverse natural phenomena. Whatever the status of the Fibonacci

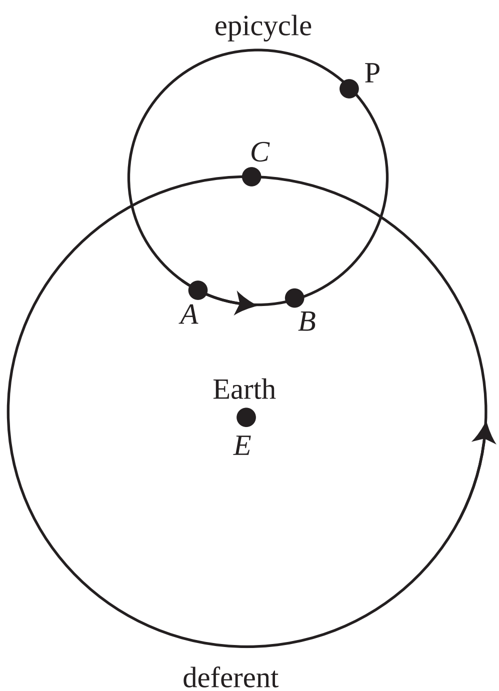


Figure 18.3 An epicycle-deferent planetary model.

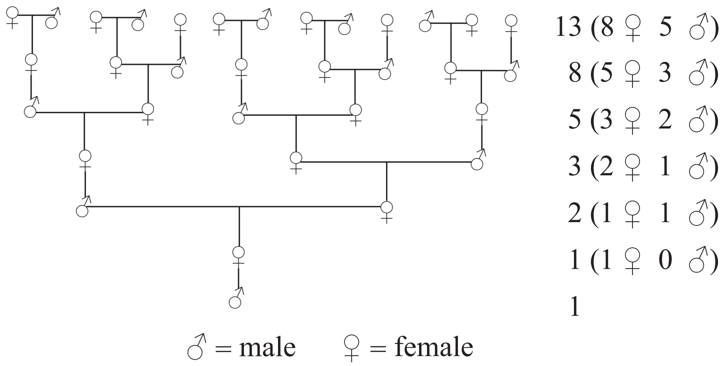


Figure 18.4 The genealogical tree of the male bee.

series, relations that are accepted as laws clearly are candidates for inclusion among conditions sufficient for causal status.

Notes

1. John Stuart Mill, *A System of Logic* (London: Longman, 1970), Book III, Chapt. 5.
2. R. B. Braithwaite, *Scientific Explanation* (Cambridge: Cambridge University Press, 1953), 311–12.
3. Carl Hempel, *Aspects of Scientific Explanation* (New York: Free Press, 1965), 347–54.
4. Karl Popper, *The Logic of Scientific Discovery* (New York: Basic Books, 1959), 59–62; *Objective Knowledge: An Evolutionary Approach* (Oxford: Clarendon Press, 1972), 91.
5. John L. Mackie, "Causes and Conditions," *American Philosophical Quarterly* 2 (1965): 246.
6. Ibid., 247.
7. Gerhart Friedlander and Joseph W. Kennedy, *Nuclear and Radiochemistry* (New York: Wiley, 1949), 187.

19

Probabilistic Causality

Ducasse's attempt to shift focus from regularities to individual events found little support. However, the inferability view received criticism from another direction. Problems arose about the causal status of statistical correlations. Consider the radioactive decay of Rn^{222}. The laws governing radioactive decay are indeterministic because state description in such cases involves reference to probabilities. Given a "large number" of atoms, the inferability criterion is satisfied. One can predict with high accuracy that 50 percent of the original atoms will have decayed (by emitting an α-particle) within 3.8 days (the half-life of the isotope). But if there is just one atom, no reliable inference is available about its state after that same period of time. One can infer only a probability that it will have decayed (50 percent).

The critic may object that on the "causation-as-inferability" view, there must be some number of atoms above which radioactive decay becomes causal. We cannot predict accurately the number of atoms that decay during 3.8 days from an initial group of two or twenty atoms. However, we can predict with very high accuracy the number of atoms that decay during this period from a gram atomic weight of Radon (222 gm, which contains 6.0 H 10^{23} atoms). The critic may urge that it is counterintuitive to hold that the radioactive decay of a few atoms is noncausal but that the decay of many atoms is causal. The appropriate response to such criticism is that this may be counterintuitive, but that we have learned that the results of extending scientific inquiry to the subatomic realm do not conform to our prior intuitions.

Wesley Salmon delivered a more important criticism of the inferability view. On the inferability view, high probability is required for causal relatedness. Only if *B*-states follow *A*-states with high probability is the relationship linking the states a "causal relation." If, on the other hand, the probability is, say, 5 percent, then the appropriate inference from the occurrence of an *A*-state is to "probably *not—B*."

But clearly there are causal sequences of states that follow with low probability. Consider the correlation between exposure to radiation and the subsequent development of leukemia. Salmon maintained that it is appropriate to speak of a "causal relation" in such cases, even if only 5 percent of persons exposed to a certain level of radiation develop leukemia within a specified period of time.[1] What counts is not the value of probability *per se*, but the statistical relevance of the probability value. A 5 percent probability value is causally significant if this is a higher value for the incidence of leukemia than is found in the absence of exposure to radiation.

In some cases, therefore, the inferability view fails to account for statistically significant correlations. Some theorists sought to develop a probabilistic view of causal dependence, on the grounds that causes raise the probabilities of the occurrence of their effects—at least to some degree. Theorists of probabilistic causality include I. J. Good,[2] Patrick Suppes,[3] Brian Skyrms,[4] Michael Tooley,[5] Peter Menzies,[6] and Ellery Eells.[7]

If a satisfactory interpretation can be formulated, then it may be possible to take "cause as sufficient condition" to be the limiting case of "probabilistic causality" achieved when the probability = 1. The central idea of this approach is that a cause is an event, or set of conditions, that increases the probability that a particular effect will occur, viz.,

$$\mathrm{P}\ (\mathrm{E/C}) > \mathrm{P}\ (\mathrm{E})$$

where $P(E/C)$ is the probability that effect E occurs, given that C has occurred and $P(E)$ is the unconditional probability that the effect occurs. If this relationship holds, then, of course,

$$\mathrm{P}\ (\mathrm{E/C}) > \mathrm{P}\ (\mathrm{E/{\sim}C})$$

where "$\sim C$" states the absence of C.

For example, scientists have argued that the probability that a heavy cigarette smoker develops lung cancer before age seventy is greater than the probability that a nonsmoker does so. If this is true, then on the probabilistic interpretation, smoking is a cause of lung cancer. And presumably Jones is correct to conclude that if he commences to smoke as a teenager then he will decrease the likelihood that he celebrate his seventieth birthday.

Some amendments are required, however.

Logically Necessary Conditions

Participating in a marriage ceremony increases the probability that a couple subsequently is involved in a divorce. But we do not wish to say that marriage is the cause of divorce. Marriage is a logically necessary condition of divorce, just as possession of a raffle ticket is a logically necessary condition of winning the raffle prize. We need to exclude logically necessary conditions from the ranks of putative causes. This presupposes that a distinction can be enforced between logically necessary conditions and contingent empirical generalizations.[8]

Causes that Decrease the Probability that Its Usual Effect Occurs

Questions have been raised about the fundamental principle of probabilistic causality

$$P\ (E/C) > P\ (E)$$

It is possible that the acknowledged cause of an event actually decreases the probability that the effect occurs. If so, then probability-raising is not a necessary condition of causal relatedness. Brian Skyrms suggested the following empirically possible scenario. He noted that there is much evidence to support a causal relationship between cigarette smoking and lung cancer. If this is the case, then

$$P\ (L/S) > P\ (L)$$

> where *L* is the development of lung cancer and *S* is cigarette smoking. But one can imagine circumstances in which
>
> $$P\ (L/S) < P\ (L)$$
>
> Suppose that air pollution got so bad that most people in the cities refrained from smoking out of sheer terror of putting their lungs in double jeopardy, while many people in areas of the countryside with relatively little pollution felt that they could allow themselves the luxury of smoking.[9]

Then,

$$P\ (L/S\ \&\ U) > P\ (L/{\sim}U)$$

$$P\ (L/S\ \&\ U) > P\ (L/{\sim}S\ \&\ {\sim}U)$$

where *U* is the recognition of, and living in, the presence of high levels of air pollution. Under such circumstances,

$$P\ (L/S) < P\ (L)$$

Skyrms labeled the supposed relationship in which smoking lowers the incidence of lung cancer a "spurious correlation." He maintained that correlations of this type may be identified as spurious by noting causally relevant background conditions. In the example cited, these conditions include the urban-versus-rural setting, the respective pollution levels in these areas, and the attitudes of residents within these two settings. Ellery Eells concurred that to specify causally relevant background conditions is to identify appropriate subpopulations.[10] Skyrms emphasized that the hypothetical smoking scenario is not an objection to the fundamental principle of probabilistic causality.[11]

One need not rest the case on hypothetical scenarios, however. There are instances in which the presence of the acknowledged cause of a type of effect actually does decrease the probability that the effect occurs. Germund Hesslow noted that the use of contraceptive pills (*C*) raises the probability that thrombosis take place (*T*),[12] viz.,

$$P(T/C) > P(T)$$

Nevertheless, use of contraceptive pills by women of child-bearing age may actually reduce the probability of thrombosis for such women. This is because the incidence of pregnancy (*G*) also raises the probability of thrombosis, viz.,

$$P\ (T/G) > P\ (T)$$

Given the magnitudes of the respective probabilities—the probability of thrombosis, given pregnancy, is greater than the probability of thrombosis, given the use of contraceptive pills—pill use may actually decrease the likelihood of thrombosis.

The impact of this case may be reduced by relativizing probabilities to specific populations of women. Infertile women who take contraceptive pills, not realizing their infertility, increase the probability of thrombosis. Within this group, taking contraceptive pills is a cause of thrombosis. Fertile women who take contraceptive pills decrease the probability of thrombosis, since the pills prevent pregnancy. Within this latter group, taking contraceptive pills is not a cause of thrombosis. Attention to the relevant subject class removes the paradox that a type of event can be both a cause of a type of effect and not a cause of the same type of effect. This case also is consistent with affirmation of the fundamental principle:

$$P\ (E/C) > P\ (E)$$

Nevertheless, probability-raising is not a necessary condition of causal relatedness. In some cases, where a cause lowers the probability of its effect in a context where reference to background conditions is unhelpful and relativization to a different subject class is not available. Peter Menzies suggested the example in Figure 19.1,[13] where *a*, *b*, *c*, . . . represent neurons and the arrows represent axons connecting neurons "→" represents stimulation and "⤙"represents inhibition).

Suppose that the probability that neuron *e* fires via the upper path from neuron *a* is high, the probability that *e* fires via the lower path from neuron *b* is low, and the probability that neuron *c* be inhibited from firing by an axon from *b* is moderate. Suppose further that neurons *a* and *b* fire simultaneously, that the path from *a* to *e* is blocked by an inhibiting axon from *b* to *c*, and that *e* fires nevertheless. It would be correct to say that neuron *b*'s firing is the cause of neuron *e*'s firing, despite the fact that it lowers the probability that this effect occurs.

Wesley Salmon presented another fictitious, but plausible, example in which the occurrence of a causal relation decreases the probability that its effect occur. Consider an excited atom in its fourth energy level that decays to its ground state (level 0). Suppose that some decay

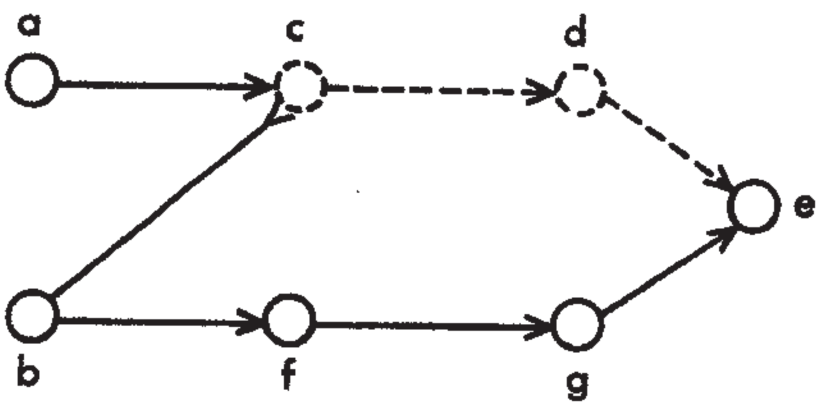


Figure 19.1 A neuron-firing network.

paths involve occupation of level 1 and that the following probability values govern the transitions $4 \to 3 \to 1$ and $4 \to 2 \to 1$:

$$P(4 \to 3) = \frac{3}{4}, \quad P(3 \to 1) = \frac{3}{4},$$
$$P(4 \to 2) = \frac{1}{4}, \quad P(2 \to 1) = \frac{1}{4}.$$

The probability that the atom occupies level 1 on its way to level 0 is 10/16. If the atom stops at level 2, its probability of occupying level 1 is 1/4. Passing through level 2 decreases the probability that level 1 be occupied. But suppose that, in fact, the decay is $4 \to 2 \to 1 \to 0$.

Salmon noted that this sequence

> constitutes a causal chain, in spite of the negative statistical relevance of the intermediate stage.[14]

Elliott Sober provided a further example.[15] A golfer's putt is intercepted by a charging squirrel, who kicks the ball aside such that it ricochets off a tree and rolls into the cup. It seems correct to say that the squirrel's kick caused the ball to drop into the cup, even though the squirrel's interference lowered the probability of holing the putt. There is a continuous sequence of events—putt-kick-ricochet-drop into cup—that constitutes a causal sequence (however low the probability that the squirrel's kick produces the required ricochet).

It seems clear, from the examples presented by Menzies, Salmon, and Sober, that probability-raising is not a necessary condition of causal relatedness. There remains the question whether probability-raising is a sufficient condition of causation. Jonathan Schaffer argued that cases of "overlapping" show that probability-raising is not a sufficient condition of causation. In cases of overlapping, a situation is altered such that the probability that effect *E* occurs is increased, but this alteration is not a cause of *E*.

For example, Merlin casts a spell that has a 50 percent chance of turning the king and the prince into frogs before midnight. Simultaneously, Morgana casts a spell that has a 50 percent chance of turning the queen and the prince into frogs before midnight. The addition of Morgana's spell increases the probability that the prince becomes a frog before midnight to 75 percent. At midnight, we find that the king and the prince are frogs. Only Merlin's spell is a cause. Morgana's spell failed. It raised the probability that the prince becomes a frog, but did not cause the transformation.[16]

This is a fanciful example to be sure. Schaffer provided a second example of overlapping from physics. Suppose an atom of U^{238} is placed in a box. There is a specific probability that it will decay into a Th^{234} atom plus an α-particle ($_2He^4$) before time *t*. Adding an atom of α-emitter Ra^{226} to the box increases the probability that an α-particle is produced before *t*. Suppose the U^{238} atom decays before *t*. Adding the Ra^{226} atom increases the probability that an α-particle be produced before *t*, but this addition is not a cause of the emission that did take place.[17]

There is an obvious objection to this example. Why take the effect to be the emission of an α-particle, regardless of its energy? If the effect is taken to be the emission of an α-particle of energy 4.25 MeV from a U^{238} atom (and not the emission of an α-particle of energy 4.86 mev from a Ra^{226} atom), then the addition of the Ra^{226} atom does not raise the probability of *this* effect, and the example collapses.

Schaffer's example may be reinstated by replacing the Ra^{226} atom by a second U^{238} atom. Adding a second U^{238} atom to the box doubles the probability of emission of a 4.25-mev α-particle. If the first atom decays before *t* and the second atom does not, this is an instance where increasing the probability of an effect is causally irrelevant to its occurrence.

In the essay "Probabilistic Causality" Salmon declared that it is a mistake to

> attempt to carry out the construction of causal relations on the basis of probabilistic relations among discrete events without taking account of the physical connections among them.[18]

What is needed, according to Salmon, is attention to those physical processes that transmit causal influence from one region of space–time to another. He sought to develop a theory of causality that takes account of both statistical relevance relations and the propagation of causal influence.

Menzies, on the other hand, was more optimistic about the prospects for probabilistic theories of causation. He declared that

> in practically all cases the simple probabilistic theory in terms of straightforward increase in probability is close enough to the truth. It is only the special cases such as those involving the flukey coincidences of pre-emption that the idea that a cause raises the probability of its effect is liable to lead us astray. Thus, if one has good reason to believe that some event increases the probability of

another event occurring and one also has good reason for thinking that the situation does not involve pre-emption, then one is justified in believing the corresponding causal statement. Increase in probability is a reliable, but not infallible guide to causation.[19]

Notes

1. Wesley Salmon, "Why Ask 'Why'? An Enquiry Concerning Scientific Explanation," *Proceedings of the American Philosophical Society* 6 (1978): 689. Reprinted in *Scientific Knowledge*, ed. Janet Kourany (Belmont: Wadsworth, 1987), 56.
2. I. J. Good, "A Causal Calculus I–II," *British Journal of Philosophy of Science* 11, 305–18; 12, 43–51; *Good Thinking* (Minneapolis: University of Minnesota Press, 1983).
3. Patrick Suppes, *A Probabilistic Theory of Causality* (Amsterdam: North Holland, 1970); *Probabilistic Metaphysics* (Oxford: Blackwell, 1984).
4. Brian Skyrms, *Causal Necessity* (New Haven, CT: Yale University Press, 1980).
5. Michael Tooley, *Causation: A Realist Approach* (Oxford: Oxford University Press, 1987).
6. Peter Menzies, "Probabilistic Causation and Causal Processes: A Critique of Lewis," *Philosophy of Science* 56 (1989): 642–63.
7. Ellery Eells, *Probabilistic Causality* (Cambridge: Cambridge University Press, 1991).
8. The distinction between logically necessary statements and contingent empirical statements was challenged by W. V. O. Quine, "Two Dogmas of Empiricism," in *From a Logical Point of View* (Cambridge: Harvard University Press, 1953), and defended by H. P. Grice and P. F. Strawson, "In Defense of a Dogma," *Philosophical Review* 65 (1956): 141–58, and Hilary Putnam, "The Analytic and the Synthetic," in *Minnesota Studies in the Philosophy of Science, Vol. III*, ed. H. Feigl and G. Maxwell (Minneapolis: University of Minnesota Press, 1962), 358–97.
9. Brian Skyrms, *Causal Necessity* (New Haven, CT: Yale University Press, 1980), 106.
10. Ellery Eells, *Probabilistic Causality*, 3–4.
11. Skyrms, *Causal Necessity*, 106–07.
12. Germund Hesslow, "Two Notes on the Probabilistic Approach to Causality," *Philosophy of Science* 43 (1976): 290–92.
13. Peter Menzies, "Probabilistic Causation and Causal Processes: A Critique of Lewis," *Philosophy of Science* 56 (1989): 645–47.
14. Wesley C. Salmon, *Causality and Explanation* (Oxford: Oxford University Press, 1998), 222–23.
15. Elliott Sober, "Two Concepts of Cause," *PSA 1984, Vol. II*, 406–07.
16. Jonathan Schaffer, "Overlappings: Probability-Raising without Causation," *Australasian Journal of Philosophy* 78 (2000): 40–41.
17. Ibid., 41.
18. Wesley Salmon, "Probabilistic Causality," in *Causality and Explanation* (Oxford: Oxford University Press, 1998), 224.
19. Menzies, "Probabilistic Causation and Causal Processes," 662.

20

Wesley Salmon on Processes and Interactions

In his discussion of the atomic explosion–leukemia example, Salmon declared that there are two requirements for a causal explanation. The first requirement is inferability. In the case at hand, it is inferability to a higher incidence of leukemia than otherwise would be expected. The second requirement is a specification of appropriate underlying mechanisms. In the example under consideration, these mechanisms include the production of γ-rays upon nuclear fission, the modification of cellular structure by γ-rays, and the differential response of modified and unmodified cells to attack by the leukemia virus. On this view, a "cause" is an event that triggers a mechanism by which some structure is produced and propagated.

Salmon maintained that the entities of which "cause" is properly predicated are "processes" and "interactions." Salmon maintained that "causal processes"

> are the means by which structure and order are *propagated* or transmitted from one space–time region to other times and places.[1]

And "causal interactions" are

> the means by which *modifications* of structure (which are propagated by causal processes) are *produced*.[2]

Salmon adopted Bertrand Russell's concept of "process" as the persistence throughout time of some entity, quality, or structure. The extent of persistence required is unclear. The motion of a positron emitted from a nucleus upon bombardment presumably qualifies as a process, even though its annihilation by an electron takes place within a minute time interval after emission. It is unclear, however, whether the persistence of an apple for that same minute period of time qualifies as a process. It also is unclear how much alteration is permitted within which a body (or wave) retains self-identity. A billiard ball that gains electrons or water vapor on its journey along a table presumably qualifies as a process. The growth of a sand dune may be controversial. Despite the attendant vagueness, there may be agreement on a sufficiently large number of cases to make the concept of "process" useful as a starting point for the analysis of causation.

Salmon noted that not every process is a causal process. Causal processes transmit modifications of structure or "marks" imposed upon them but "pseudo-processes" do not. Salmon's illustration of this distinction is the illumination of the walls of a darkened room by a rotating searchlight beam.[3] The bright path on the wall may be marked by placing a red filter on the wall. The light path becomes red for a time, but this mark is not transmitted as the searchlight continues to rotate. The light path on the wall is a "pseudo-process." By contrast, the searchlight beam itself (an electromagnetic wave) is a causal process. It may be marked by placing a red filter on the lens or on an extension tube attached to the lens. This mark is transmitted in subsequent rotations of the searchlight.

Structures are propagated by causal processes. It remains to specify how structures are produced. According to Salmon, structures are produced when causal processes intersect in such a way that the respective processes are modified. Normally, such intersections generate modifications of structure. An exception is the intersection of two light rays that takes place without any photon–photon collisions (Figure 20.1).[4, 5]

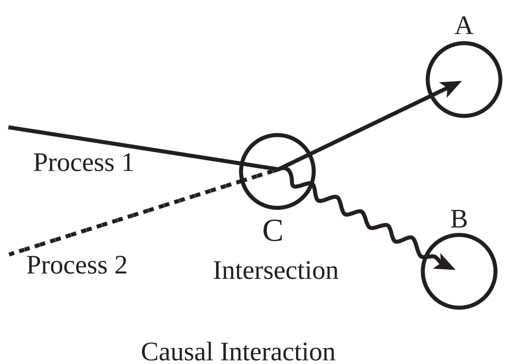


Figure 20.1 Intersection of causal processes.

One type of causal production is the "conjunctive fork" (Figure 20.2[6]). In a conjunctive fork, the instances of production of effects are independent of one another. Salmon's frequently chosen example is the production of cases of leukemia by the explosion of an atomic bomb.

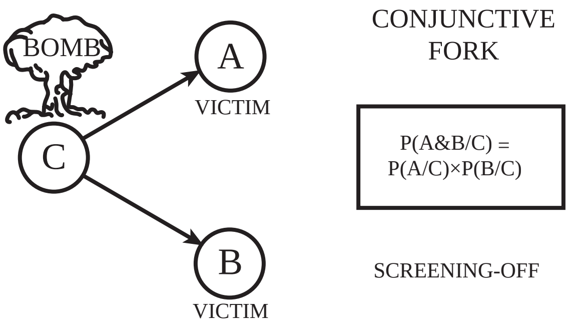


Figure 20.2 The conjunctive fork.

As noted earlier, the causal mechanisms involved include the production of γ-rays upon nuclear fission, the passage of these rays through the air, the interaction of the rays with cells of the human body, and the differential responses of modified and unmodified cells to the presence of the leukemia virus. Given these causal mechanisms, some individuals in the vicinity of the explosion develop leukemia.

The probability that an individual one mile from the epicenter develops leukemia within five years is independent of the probability that other individuals at the same distance develop the disease. Given that *C* is the explosion and *A* and *B* are individual leukemia cases, then

$$\mathrm{P}\ [(\mathrm{A\&B})/\mathrm{C}] = [\mathrm{P}(\mathrm{A}/\mathrm{C}) \times \mathrm{P}\ (\mathrm{B}/\mathrm{C})]$$

Salmon observed that the following four conditions are fulfilled in the case of a conjunctive fork:

1. $P[(A \,\&\, B)/C] = [P(A/C \times P(B/C)]$
2. $P[(A \,\&\, B)/{\sim}C] = [P(A/{\sim}C) \times P\ (B/{\sim}C)]$
3. $P(A/C) > P(A/{\sim}C)$
4. $P(B/C) > P(B/{\sim}C)$

Hans Reichenbach had proved that these four conditions imply[7]

$$\mathrm{P}\ (\mathrm{A} \,\&\, \mathrm{B}) > [\mathrm{P}\ (\mathrm{A}) \times \mathrm{P}\ (\mathrm{B})]$$

Thus the probability that both *A* and *B* occur is greater than the product of the individual probabilities that *A* occur and *B* occur. The greater probability of the joint occurrence of *A* and *B* is due to the occurrence of the prior event *C*. *C* is the "common cause" of these two events.

Common cause *C* "screens off" effects *A* and *B* from one another. It renders their individual occurrences statistically independent, such that

$$P\,(A/C) = P\,(A/C\;\&\;B)$$
$$P\,(B/C) = P\,(B/C\;\&\;A)$$

Salmon noted that one way to establish the causal status of a process is to show that it is related to independent events that occur with a higher-than-expected frequency. For example, if 50 percent of the students who dined in Smith Hall Thursday night fell victim to food poisoning (presumably an unexpectedly high percentage for a meal at Smith Hall), then it is prudent to search for a common cause among the items consumed at dinner. On Salmon's theory, the production of an effect is an individual interaction, but the warrant for causal status involves appeal to statistical generalizations.[8]

A second type of causal production is the "interactive fork." Salmon called attention to Compton scattering (Figure 20.3[9]), in which a photon of energy *E* collides with an orbital electron, dissociating it from its atom and imparting to it kinetic energy E_1. The collision also creates a second photon of energy E_2.

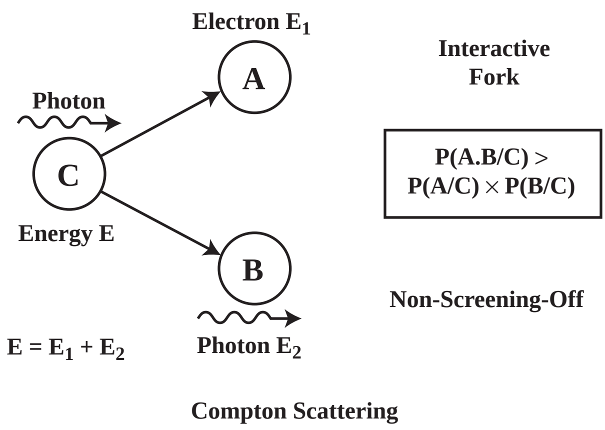


Figure 20.3 Compton scattering.

The effects of the photon–electron collision are not independent of one another. Energy is conserved in the interaction $E = E_1 + E_2$. (The orbital energy possessed by the electron before impact has been neglected here. Its value is very small compared to E_1.)

No common cause is present to screen off the effects from one another. The two effects are statistically dependent on one another, viz.,

$$P [A/C \& B] > P (A/C),$$
$$P [B/C \& A] > P (B/C).$$

Salmon noted that, for this interaction

$$P [A \& B/C] > [P (A/C) \times P (B/C)]$$

Suppose the probability that an electron of kinetic energy E_1 is created by the photon–electron collision is 0.1, and the probability that a photon of energy E_2 also is created is 0.1. The joint probability of these two effects is *not* the product of the individual probabilities. For this hypothetical case, the joint probability of *A* and *B*, given *C*, is the same as the individual probability of *A*, given *C*, i.e., 0.1.

Phil Dowe called attention to an obvious limitation of the statistical relevance approach to the production of causal structure. The conjunctive-fork pattern is useful when two or more effects arise from a common cause, but it is inapplicable to cases in which a single effect is to be analyzed. Suppose that just one student falls victim to food poisoning. This is a causal relation—new structure has been produced—despite the absence of statistical data.[10]

Causal Interaction and the Exchange of Conserved Quantities

Critics soon developed objections to the mark-transmission criterion of causal process. Nancy Cartwright suggested a modification of Salmon's rotating-beam illustration.[11] Suppose that a red filter is placed on the lens at the same time that a second red filter is placed over the formerly white spot on the wall. The spot on the wall is marked by red color, and this mark *is* transmitted along the path traced on the wall. The light path on the wall no longer qualifies as a "pseudo-process." Salmon conceded that the path on the wall can be relegated to the status of a "pseudo-process" only upon appeal to a counterfactual claim. If no filter had been placed on the lens, then the moving spot on the wall would not have remained red. Salmon regarded an appeal to a counterfactual condition in the application of a criterion to be unsatisfactory. He therefore abandoned the mark-transmission criterion in favor of an alternative criterion recommended by Phil Dowe.[12] Dowe suggested that causal processes, unlike pseudo-processes, display the conservation of some quantity.[13]

Scientists link the concepts "conservation" and "closed system." Mario Bunge observed that

> the hypothesis of isolation, or, conversely, of a noninterfering background, is, then, a methodological requirement of the sciences dealing with the material world; hence, the fiction of the isolated "causal chain" will work to the extent to which such an isolation takes place.[14]

Roughly speaking, a system is closed with respect to quantity Q if, and only if, there is no net flow of Q into or out of the system. Within such systems interactions occur in which momentum, energy, charge, spin, lepton number, etc., remain unchanged. On Dowe's theory, such interactions are "causal."

To avoid circularity, "net flow" must be explicated without reference to the concept of causation. Sungho Choi suggested that this is achieved by the following definition:

> a system is closed with respect to a [scalar] physical quantity Q at a time t if, and only if,
>
> (a) $dQ_{in}/dt = dQ_{out}/dt = 0$ at t, or
>
> (b) $dQ_{in}/dt \neq \sim dQ_{out}/dt$ at t,[15]

where Q_{in} is the amount of Q inside the system and dQ_{in}/dt is the rate at which Q changes within the system. On the other hand, if Q_{in} does increase (decrease) while Q_{out} decreases (increases) at the same rate, then the system is not closed with respect to Q at t.

New structures emerge whenever there is an interaction in which a quantity conserved within a closed system is exchanged. Interactions in which energy, momentum, charge, spin, etc., are exchanged qualify as genuine causes. Salmon declared that

> a causal interaction is an intersection of world-lines that involves exchange of a conserved quantity.[16]

Salmon's theory, modified to emphasize conserved quantities rather than mark-transmission, assigns causal status to the intersection of world lines of cue ball and 8-ball upon collision. A conserved quantity exchanged is linear momentum, initially possessed exclusively by the cue ball, but divided after impact between the cue ball and the 8-ball. This picture is an idealization of course. It ignores frictional losses and departure from perfect elasticity upon impact. Such idealization is common in explanations given in the physical sciences.

Notes

1. Wesley Salmon, "Causality: Production and Propagation," in *PSA 1980 Vol. 2*, ed. P. Asquith and R. Giere (East Lansing, MI: Philosophy of Science Association, 1981), 64. Reprinted in *Causation*, ed. E. Sosa and M. Tooley (Oxford: Oxford University Press, 1993), 169.
2. Ibid., 64.
3. Salmon, *Scientific Explanation and the Causal Structure of the World* (Princeton, NJ: Princeton University Press, 1984), 141–43.
4. Salmon, "Why Ask 'Why'? An Inquiry Concerning Scientific Explanation," *Proceedings and Addresses of the American Philosophical Association* 51 (1978); Reprinted in *Causality and Explanation* (Oxford: Oxford University Press), 135.
5. Salmon, "Causality: Production and Propagation," 60.
6. Ibid., 134.
7. Hans Reichenbach, *The Direction of Time* (Berkeley: University of California Press, 1956), 160–64.
8. Salmon, "Causality: Production and Propagation," 66.
9. Salmon, *Causality and Explanation*, 133.
10. Phil Dowe, *Physical Causation* (Cambridge: Cambridge University Press, 2000), 19.
11. Cartwright's criticism is cited by Salmon, "Causality without Counterfactuals," *Philosophy of Science* 61 (1994): 302.
12. Salmon, "Causality without Counterfactuals," 302–03.
13. Phil Dowe, "Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory," *Philosophy of Science* 59 (1992): 195–216.
14. Mario Bunge, *Causality and Modern Science* (New York: Dover, 1979), 130.
15. Sungho Choi, "The Conserved Quantity Theory of Causation and Closed Systems," *Philosophy of Science* 70 (2003): 519.
16. Salmon, "Causality and Explanation," *Philosophy of Science* 64 (1997): 468.

21

Phil Dowe and the Conserved Quantity Theory

Dowe gave high marks to Salmon for analyzing causality by reference to processes and interactions. He sought to improve upon Salmon's theory by replacing the criterion of mark transmission by the requirement that causal processes possess a conserved quantity. What counts as a conserved quantity is determined by the structure of current scientific theories. "Mass-energy," "linear momentum," "charge," and "spin" are quantities that appear in currently accepted conservation laws, and hence qualify as conserved quantities.

Given a set of conserved quantities sanctioned by science, Dowe redefined "causal process" and "causal interaction":

> A *causal process* is a world line of an object that possesses a conserved quantity.
>
> A *causal interaction* is an intersection of world lines that involves exchange of a conserved quantity.[1]

To apply these concepts, one must trace individual objects through time. This, in turn, presupposes that "identity-through-time" can be predicated correctly of an object. Dowe was content to leave "identity-through-time" as a primitive term within his conserved quantity theory. He noted however, that scientists often achieve success in tracing individual objects through time.

Consider the impact of two moving billiard balls *a* and *b*. The impact fits the causal interaction pattern presented in Figure 21.1[2] and not the intersection pattern in which *a* and *b* pass through one another.

In the realm of classical physics, at least, scientists have no difficulty distinguishing causal interactions from intersections.

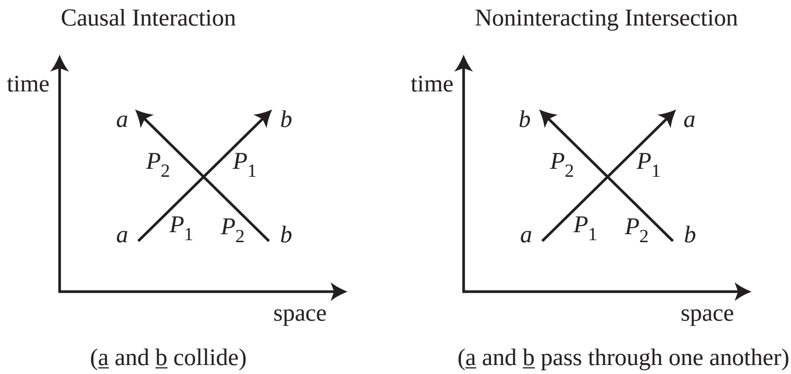


Figure 21.1 Causal interaction and noninteracting intersection.

Possession versus Transmission of Conserved Quantities

Salmon stressed that Dowe's identification of a causal process with the *possession* of a conserved quantity over time requires the identity-over-time of the object that possesses the quantity. Salmon maintained that Dowe's conserved quantity theory would be improved if explicit commitment to the self-identity of objects could be avoided. He sought to achieve this by identifying a causal process, not with the possession by an object of a conserved quantity, but with the continuous transmission of that quantity. He declared that

> a process transmits a conserved quantity between *A* and *B* ($A \neq B$) if and only if it possesses [a fixed amount of] this quantity at *A* and at *B* and at every stage of the process between *A* and *B* without any interactions in the open interval (*A*, *B*) that involve an exchange of that particular conserved quantity.[3]

Salmon restricted causal processes to the transmission of a *fixed amount* of a conserved quantity within a closed system. Unless the "fixed amount" requirement is relaxed, Salmon's characterization will fit very few processes involving the transfer of energy and linear momentum. There is a loss of energy from systems that are incompletely isolated (closed) from their environments. For example, the collision of a cue ball and an 8-ball on a pool table, strictly speaking, would not be susceptible to causal analysis. The world line of the moving cue ball does not display a *fixed* amount of linear momentum.

There are frictional losses during the motion. Taking note of this, Dowe declared that

> the Conserved Quantity theory does not require that a constant amount of the relevant quantity over the entire history of the process.[4]

Dowe defended "possession" in this loose sense of "approximate conservation" and not "transmission." According to Dowe, Salmon's "transmission" concept of causal process fails to discriminate between the "causal interaction" and "intersection" diagrams of figure 21.1. Salmon had rejected an explicit commitment to "identity-through-time" in favor of the concept "transmission of conserved quantity *Q* at every space–time point of a process." But without assumptions about self-identity, Salmon has no way to tell whether *a* and *b* collide with an exchange of *Q* or *a* and *b* intersect without interacting. Dowe asked the reader to consider intersections involving neutrinos:

> a stray neutrino passes through my body. Am I still myself, unaffected by the event, or am I now the thing that used to be the neutrino, having been radically transformed by the experience?[5]

According to Dowe, the identification of a causal process with the transmission of a conserved quantity fails to discriminate between the interaction and the intersection of processes.

Salmon conceded that this is a difficulty for the transmission theory. However, he maintained that there is an additional consideration in its favor. The transmission theory enables one to avoid the problem posed by "gerrymandered objects." Consider a rotating searchlight located on the axis of a cylindrical room. Its beam traces a line on the wall of the room. Salmon argued that the occasionally illuminated segments of the wall surface possess the conserved quantity energy but fail to qualify as a causal process.[6] If this is correct, then it is wrong to identify causal processes with the possession of a conserved quantity. Salmon took this to support his own position that it is the transmission of a conserved quantity, and not just its possession, that distinguishes a causal process from a pseudo-process.

Dowe replied that the successively illuminated wall-surface segments do not qualify as a genuine object. Instead this is a "gerrymandered aggregate." He maintained that

> it is intuitively clear that the temporal stages of certain timewise gerrymanders are not temporal parts of a single object.[7]

This is the case even in the absence of a satisfactory theory about identity-through-time.

Dowe observed that we sometimes cite as causes events that prevent the occurrence of further events ("preventions") and events whose nonoccurrence is followed by the occurrence of further events ("omissions"). He called attention to debates in philosophy of law about whether neglect is a cause and debates in moral philosophy about whether letting die is a cause of death. The conserved quantity theory does not account for such usage directly. There is in such cases no set of causal processes involving exchange of a conserved quantity that lead to the effect in question. Dowe appealed to counterfactual conditional claims to account for preventions and omissions.

Consider a typical case of preventing suggested by Brian Ellis—"Pulling down the windowshade caused the room not to be light."[8] Dowe restated this causal claim about prevention as the conjunction of a causal claim about the world (1) and a counterfactual conditional claim about a "similar" alternative world (2), viz.:

1. "Pulling down the windowshade caused light to be reflected from the shade" and
2. "If light were not reflected from the shade, then the room would be light."

Claim 1 fits the conserved quantity model but does not account for the prevention. Claim 2 needs to be appended. The conjunction of 1 and 2 provides a satisfactory account of this instance of causation-by-prevention.

Dowe provided a similar complex analysis for cases of causation-by-omission. He maintained that to say that "the father's inattention was the cause of the child's accident" is to say that although the accident occurred, "it was possible for him to have prevented it."[9] To unpack the claim that it was possible for the father to have prevented the accident, it is necessary to make a counterfactual judgment about an alternative world in which the father's actions were different.

So although causation-by-prevention and causation-by-omission are not subject to a direct interpretation in terms of interacting causal processes, a hybrid interpretation can be given. The hybrid interpretation involves both actual and counterfactual claims.

Of course, this hybrid interpretation is a bit artificial. If causation requires an *actual* exchange of a conserved quantity, "causation-by-prevention" and "causation-by-omission" are not instances of causation at all. Dowe preferred instead to incorporate preventions and omissions by introducing "causation*," an augmented "causation" that includes counterfactual claims. However, he conceded that

> causation* is not genuine causation but a counterfactual truth about causation, in certain circumstances.[10]

The conserved quantity theory is put forward as an analysis of the structure of causal processes and their interactions. It remains to be specified what sorts of entities are related as causes and effects.

Dowe, following David Armstrong[11] and Ludwig Wittgenstein,[12] maintained that the world is a collection of individual states of affairs. Basic, or "atomic," states of affairs include the possession of properties by particulars and relations among particulars (represented in the usual logical notation by "Pa" and "Rab").

The constituent parts of states of affairs are events and facts. Dowe characterized an event as "a change in a property of an object at a time"[13] and a fact as "an object having a property at a time or over a time period."[14] He suggested that

> the things we call causes and effects (facts, events, or whatever they are) at base involve the possession of, or change in the value of, some conserved quantity.[15]

Application to typical events on a pool table is straightforward. Suppose a cue ball is struck by a cue stick, travels a short distance, strikes a stationary 8-ball, and stops, as the 8-ball travels to and enters the far right-hand corner pocket. We identify as cause the interaction of cue stick and cue ball, and as effect the interaction of cue ball and 8-ball. The quantity "linear momentum" is exchanged in the cause, possessed by the causal process that is the moving cue ball, and exchanged in the effect.

Dowe applied the conserved quantity theory to the photon-induced decay of an atom. Suppose atom X is struck by photon p at time t_1. Subsequently, atom X transmutes into atom Y, with the release of energy at time t_2. The incident photon is linked to the existence of atom Y by one interaction that involves an exchange of energy and a second interaction that involves an exchange of both energy and charge. Both energy and charge are conserved quantities in the decay process. Dowe noted that if atom *X* does not decay, it would be incorrect to cite the photon impact as a cause, since the effect involves an exchange of charge, and there is no such exchange in this instance.[16]

One might argue that not every process in which a conserved quantity is conserved qualifies as a causal interaction. The spontaneous decay of a radioactive atom is a case in point. Suppose an atom of Ra^{226} emits an α-particle at time t (Figure 21.2).[17] Mass-energy and charge are exchanged but there is no causal interaction.

Dowe conceded that the exchange is not accomplished by an interaction. Nevertheless, he interpreted radioactive decay to be a causal process. It shares causal status with inertial motion. Dowe declared that taking radioactive decay to be a causal process may "rattle everyday intuitions."[18]

But what is the "object" that possesses conserved quantities? In the case of inertial motion, it is clear that it is the moving body (given that only approximations to inertial motion are possible). In the case of the Ra^{226} decay process, at one point in time it is the Ra^{226} nucleus, but subsequently it is an Rn^{222} nucleus plus an α-particle. Dowe elected to include radioactive decay processes among causal processes despite questions about the identity-through-time of the "object" possessing the conserved quantity.

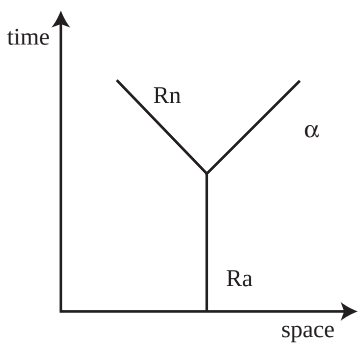


Figure 21.2 Decay of radium^{226}.

Notes

1. Phil Dowe, *Physical Causation* (Cambridge: Cambridge University Press, 2000), 90.
2. Ibid., 108.
3. Wesley Salmon, "Causality and Explanation," *Philosophy of Science* 64 (1997): 462.
4. Dowe, *Physical Causation*, 118.
5. Ibid., 254–55.
6. Salmon, *Scientific Explanation and the Causal Structure of the World* (Princeton, NJ: Princeton University Press, 1984), 145–45.
7. Dowe, *Physical Causation*, 101.
8. Dowe, "Good Connections: Causation and Causal Processes," in *Causation and Laws of Nature*, ed. H. Sankey (Dordrecht: Kluwer, 1999), 257–58.
9. Ibid., 261.
10. Dowe, *Physical Causation*, 145.
11. David Armstrong, *A World of States of Affairs* (Cambridge: Cambridge University Press, 1997).
12. Ludwig Wittgenstein, *Tractatus Logico-Philosophicus* (1922) (New York: Humanities Press, 1961).
13. Dowe, *Physical Causation*, 169.
14. Ibid., 170.
15. Ibid., 171.
16. Ibid., 173.
17. Ibid., 90.
18. Ibid., 23.

22

Causality and the Transfer of Energy or Momentum

The conserved quantities most often cited in transition from one state to another are energy and momentum. A plausible option for causal theorists is to identify causal relatedness and the transfer of energy or momentum. On this theory, state *A* is a cause of state *B* only if there is a transfer of energy or momentum from *A* to *B*. A mere succession of states is not causally significant.

David Fair defended this theory. It is instructive to compare his position with that of William Whewell (1794–1866). Whewell believed that causality was a "fundamental idea" applicable, like the ideas of space and time, within every science. He maintained that three axioms stipulate the meaning of this idea: (1) nothing takes place without a cause, (2) the magnitude of an effect is proportional to the magnitude of its cause, and (3) to every action there is an equal and opposed reaction.[1]

According to Whewell, these axioms specify the formal structure of causality. He held that the content of the idea was first specified by Isaac Newton. Newton showed that (1) bodies possess no intrinsic internal source of acceleration, (2) every acceleration is directly proportional to a force, and (3) there are certain appropriate ways to define "action" and "reaction." Thus the empirical significance of the causal relation was discovered within the history of science.[2]

David Fair defended a similar position. He observed that physicists have established an identity between temperature and kinetic energy. Statements of temperature values are reducible to statements about mean kinetic energy. Fair proposed a comparable reduction of causal relatedness to a transfer of energy or momentum from cause to effect. He insisted that

> physical science *has* discovered the nature of the causal relation for a large class of cases. As a first approximation, it is a physically-specifiable relation of energy–momentum flow from the objects comprising cause to those comprising effect. The causal relation of ordinary language seems likely to be *reducible* to that relation.[3]

Fair noted that causal relatedness usually involves a transfer from cause to effect of both energy and momentum. However, momentum transfer alone may suffice to establish causal relatedness. This is the case where objects exert forces on other objects without an exchange of energy. For instance, a satellite's motion around its planet involves a continuous transfer of momentum without a transfer of energy. Of course, the concept of "motion subject to a $1/R^2$ force directed from a stationary point center" is an idealization. Gravitational attraction between real bodies is mutual. The planet also moves in response to the force exerted by the satellite. Nevertheless, it is appropriate to speak of a momentum transfer from a "cause object" (the planet) to an "effect object" (the satellite). There also is a momentum transfer from a magnetic field to an electron moving at constant velocity in a plane at right angles to it. The curved path of the electron indicates changing momentum, but not changing energy. It is appropriate, as well, to speak of this momentum transfer as an instance of causal relatedness.

Fair insisted, moreover, that transitions from potential energy to other forms of energy qualify as causal. If I am careful, I do not impart energy to an apple when I cut its stem. But I do release the potential energy it possessed in the gravitational field of the earth. This release of potential energy is a transfer that should count as a cause.

A release of potential energy requires that restraining forces be overcome. My cutting the stem accomplishes this. In more esoteric contexts, the restraining forces are not so obvious. For example, ${}_6C^{12}$ atoms have a high-potential energy barrier against positron emission. ${}_7N^{13}$ atoms have a much lower barrier against positron emission. Bombarding ${}_6C^{12}$ atoms with protons transmutes them to ${}_7N^{13}$ atoms, thereby lowering the potential energy barrier to positron emission. The proton bombardment causes positron emission by lowering restraining forces.[4]

With the advance of science, new ways are discovered to reduce barriers to the release of potential energy. The reduction of causal relations to energy–momentum transfers will remain incomplete until an exhaustive list of such restraint removers is available. This is one reason why Fair's reduction thesis has the status of a program to be achieved.

Since energy and momentum are quantities conserved (approximately) in many structure-modifying interactions, Fair's proposal might seem to be just another way of stating the Salmon–Dowe theory of causation. Consider Tiger's last drive. On the Salmon–Dowe interpretation, the world lines of two causal processes—club and ball—intersected so as to create a modification of structure. On the Fair interpretation, energy and momentum were transferred from the club to the ball. But energy and momentum are just the conserved quantities responsible for the observed modification of structure.

There is a difference, however, in what is taken to be fundamental. Salmon and Dowe took causal processes to be fundamental. Fair sought to remain neutral on the ontological status of the relata of the causal relation. He referred to these objects an "A-objects" and "B-objects" and noted that

> to state the physical conditions of energy–momentum transference . . . it is necessary to have physical descriptions of the energy–momentum properties of what I call "A-" and "B-objects."[5]

What remains to be shown is how our ordinary usage of causal language, or at least most of it, can be reduced to instances of energy–momentum transfer. Ordinary usage predicates "cause" not only of "events" (the impact of cue ball and 8 ball causes the latter to enter the left side pocket), but also of "fields" (the presence of the magnet caused the motion of the nail), "properties" (the arrangement and spacing of water molecules in the ice cube caused it to float in liquid water), "facts" (that Andy's prints were on the gun beside the body caused the jury to convict him), and "mental states" (Vance's vertigo caused him to cry out).

What are needed are redescriptions of the above "causes" and "effects" as objects manifesting energy or momentum such that there is a transference thereof from "A-objects" to "B-objects." Fair conceded that, for many cases of causal relatedness, the needed redescriptions are not available. In that respect as well, his theory of causation is a program for further research.[6]

There is a further problem for the energy transfer theory of causation. Given candidates for "A-objects" and "B-objects," it remains to be shown that the difference in their energy values is an aspect of a *bona fide* process. Recall Salmon's rotating-beam thought experiment. It might seem that successive illuminated segments of the wall surface qualify as causally related on the energy transfer theory. Object A (a particular patch of surface) loses energy as the beam moves on and object B (an adjacent patch) gains energy (a minute amount to be sure). There is a need to distinguish "genuine" from "pseudo" energy transfer. The energy transfer theory is subject to the "individuation-of-a-process" problem that was the subject of a debate between Salmon and Dowe (see above, pp. 121–24).

In some ordinary-language causal attributions, the relevant energy flow is from "effect" to "cause." For instance, we speak of the addition of ice cubes as a cause of the cooling of water in a glass, despite the fact that, in terms of energy flow, it is the water that supplies heat to the ice cubes, causing them to melt. Fair suggested that ordinary usage be accommodated by expanding the class of causes to include both energy donors and energy recipients.[7]

Causation-by-omission is an obvious problem for the "causation-as-energy-transfer" position. Consider an example cited by Fair. There is no energy transfer between a flagpole and its shadow. And yet in common parlance the pole *causes* its shadow. As Fair put it

> the failure of the flagpole to transmit incident light causes the failure of light to reach the shadow region.[8]

One option for Fair would be to deny that such "causation-by-omission" is genuine causation. Instead he sought to validate cases of causation-by-omission as genuinely causal by considering what would take place in a possible world alternative to our own. In this alternative world that resembles our own, the flagpole is not present, and light from the sun illumines the region that corresponds to the shadow in the actual world. A "genuine" causal relation exists in this alternative world. This "genuine" causal relation is a transfer of energy to the area that corresponds to the actual-world shadow. Fair suggested that since there is a "genuine" causal relation in the alternative world that closely resembles our own, we may accept the flagpole–shadow relation as an example of "indirect causation."

In another frequently discussed example, Gardener Gus neglects to water a rose bush in his care and the bush dies. Fair sought to accommodate the demise of the bush as an instance of "indirect" causation. In a world that closely resembles our own, Gus faithfully waters the bush and it flourishes. In this alternative world, there is energy transfer from the addition of the water that helps maintain the essential metabolic processes of the bush.

Fair recommended accepting such cases as "causal," with the understanding that, unlike "direct" causal relations that involve energy transfer, they depend on what is believed to be true under counterfactual conditions. For a case of omission to qualify as "causal," there must be a large measure of agreement among individuals' beliefs about the relevant counterfactual analysis.[9] For instance, President Obama also failed to water the bush in question, but we do not believe that it is correct to call this omission a "cause." By including instances of causation-by-omission in this way, Fair protects the status of energy or momentum transfer as necessary for causal relatedness. However, the energy transfer in question takes place in a world other than our own.

Notes

1. William Whewell, *Philosophy of the Inductive Sciences*, 2nd ed. (1847) (London: Cass, 1967), Part I, Book III, 177–85.
2. Ibid., Part I, Book III, 248.
3. David Fair, "Causation and the Flow of Energy," *Erkenntnis* 14 (1979): 220.
4. Ibid., 245.
5. Ibid., 233.
6. Ibid., 236.
7. Ibid., 243.
8. Ibid., 248.
9. Ibid., 247.

23

Causality and Powers

Hume had drawn a distinction between regularities and the forces and powers that supposedly are responsible for them. His "official position" is that we may achieve knowledge of regularities in nature, but that we can achieve no knowledge of any underlying powers that may determine them. To speak of "causally efficacious powers" is to engage in prescientific discourse. A causal relation is nothing over and above a constant sequential conjunction of the members of two classes of events.

Since Hume's time, most theories of causality have been given in the categories "events," "states of physical systems," "necessary or sufficient conditions," or "probabilities." References to "powers," "capacities," or "productive efficacy" have been few and far between.

Harré and Madden on Causal Powers

Rom Harré sought to redress this imbalance. In collaboration with Edward H. Madden, Harré called attention to the fact that theories of causal relatedness are set within a specific worldview.[1] In particular, the Humean regularity view is set within an event ontology according to which:

1. The world is a succession of discrete, self-contained events, each of which is independent of the rest. Moreover, coexisting properties are independent of one another, as are the predicates that refer to them.
2. Causal relations are constant sequential conjunctions of the members of two types of events.

Harré and Madden recommended the replacement of the Humean event ontology by a new set of categories, viz.,

Competing worldviews		
	Hume	**Harré and Madden**
Basic entity	Events	Powerful particulars
Basic relation	Succession	Grouping into natural kinds
Laws	Statements of constant sequential conjunctions	Power ascriptions
Cognitive status of laws	Contingent	Exhibit natural necessity
Theories	Deductively subsume laws	Specify generative mechanisms

Harré and Madden sought to invert the ontological dependence of things upon events. They declared

> on our view, "event" is to be understood in terms of an ontology of enduring things, while on the Humean view enduring things are conceived to be constructions of events.[2]

According to Harré and Madden, there exist in the world "powerful particulars" which have the power to produce certain effects in certain circumstances. On their view, causal laws ascribe powers to particulars. The form of a power ascription is:

> *X* will (or can) do *A*, in the appropriate conditions, *in virtue of its intrinsic nature.*[3]

A power may be ascribed correctly even when there is little knowledge of the nature of the particular. For example, the power of preventing scurvy was ascribed correctly to fresh lime juice long before it was known that lime juice contains vitamin C.[4] Power ascriptions may be defended in two ways. One may show that a particular does act in the manner predicted. For instance, a sodium lump exposed to chlorine does produce salt. Or one may hypothesize a causal mechanism that is responsible for the action. In a subsequent essay, Harré declared that

> causal mechanisms are simply interacting assemblages of powerful particulars. When stimulated by an event (the cause in the old ontology) they behave in such a way as to transform themselves so that an event of the effect type can be observed, detected, or inferred to occur.[5]

Harré included among the types of events that trigger causal mechanisms the impact of a neutron upon an atomic nucleus, the completion of an electric circuit, and the mixing of chemicals.[6]

The "causes-as-powers" position conflicts with much usage of causal language by scientists. It requires us to say that the impact, the movement of the switch handle, and the mixing of chemicals are "occasioning conditions" or "triggering events," but not, strictly speaking, the "causes" of the relevant effects. The real causes are the "powers" of the nucleus, the closed circuit, and the chemical reactants to produce the effects in question.

Perhaps scientists could be persuaded to accept a modification of this usage. If the term "cause" is reintroduced as shorthand for "occasioning condition" or "triggering event," then one may speak of "causes" as individual entities and "causal powers" as capacities or dispositions that entail generalizations. Harré's position then could be restated as the demand that causal analysis go beyond the specification of "causes" to develop theories about the underlying "causal processes."

This modification clearly is necessary for legal discourse and moral discourse about causation. Appeals to the "causes-as-powers" theory normally are not acceptable within discourse assessing human responsibility. Appeals to the powers of arsenic or wood are not even *prima facie* defenses against charges of murder or arson.

To formulate a mechanism that, if realized, would explain why particulars manifest the causal powers they do is an achievement of theory. Successful theories often explain causal powers in terms of structures and processes at a deeper level of organization. For instance, the causal power of sodium to react with fluorine, chlorine, and bromine is explained, in part, by reference to the arrangement of orbital electrons in the sodium atom.

A power ascription states more than a Humean *de facto* association. The association is claimed to hold in virtue of the very *nature* of the particular. Harré and Madden maintained that our fundamental causal laws express natural necessities. They declared that

> given some general theory specifying the fundamental causal powers and thereby laying down the general lineaments of a world, the necessity of certain effects can be inferred. Such effects are "hypothetically necessary" in the sense that, given the specification of the causal powers of the things and substances of the world, the denial of statements describing these effects of those powers, when the environment allows them to be exercised, would be inconsistent

> with the nature of those things ascribed to them on the basis of the theory.[7]

Natural necessities thus are relative to the assumed truth of theories. Given the truth of gravitational theory, for instance, it is necessary that a stone released near the earth falls to the earth. It is not logically necessary that the stone fall, but in *this* world—a world whose structure is determined (in part) by gravitational attraction—its fall is naturally necessary.

The Harré–Madden view of science includes a hierarchy of explanatory levels:

A dynamicist worldview

	Example
Causal powers of particulars$_1$	Powers of sodium samples to react with chlorine
Explained by	
Nature of particulars$_1$	Sodium's electropositive valence
Explained by	
Causal powers of particulars$_2$	Powers of atoms of sodium
Explained by	
Nature of particulars$_2$	Orbital electron structure of sodium atoms
Explained by	
Causal powers of particulars$_3$	Powers of orbital electrons

Harré and Madden noted that at any given stage in the development of science certain causal powers are taken to be ultimate. They are taken to be "ultimate" in the sense that these powers are not explained by reference to the nature of some underlying particulars. An example is gravitational attraction in Newtonian physics. (Newton himself hedged his bets. At times he confessed ignorance of the nature of gravity. But at other times he speculated that gravity is caused by the operations of an aethereal fluid.)[8]

An interesting question is whether there are any entities that intrinsically qualify as ultimate. The nature of such an entity would be identical with its powers. Harré and Madden suggested that physical fields satisfy this requirement of identity. A field has no nature apart from its powers. An electromagnetic field, for instance, is nothing more than a distribution of potentials throughout space. Subtract from the field the power to affect magnets and electric charges, and there is nothing left.

Of course, it would be imprudent to claim that a specific field concept is ultimate. A theory one day may be available that explains the distribution of potentials by reference to the nature of some underlying particulars. Nevertheless,

> if only fields exhibit an identity between the powers and nature of a particular, then the ultimate level, wherever it may be, must be described in field-like concepts.[9]

Cartwright on Capacities and Causal Laws

Nancy Cartwright observed that

> the term "cause" is highly unspecific. It commits us to nothing about the kind of causality involved nor about how the causes work.[10]

Consider a few examples:

1. "Addition of the seed crystal caused precipitation in the supersaturated solution." In this context, the "cause" is a triggering event that actualizes the potential present within a situation.
2. "Cutting the apple's stem caused it to fall to the ground." In this context, the "cause" is an event that removes an impediment that, up to that moment, prevented the occurrence of an effect.
3. "The atomic explosion caused an increased incidence of leukemia." In this context, the "cause" is an event that increases the probability that other events occur.
4. "Regular use of contraceptive pills by fertile women cause a reduced incidence of pregnancy." In this context, the "cause" is events that decrease the probability that other events occur.

In each case, to make a causal claim is to ascribe a capacity to an object (or event) within a particular type of situation. Cartwright maintained that

> the generic causal claims of science are not reports of regularities but rather ascriptions of capacities.[11]

To assign a causal role to a force, for example, is to specify a *capacity* to change the state of motion of a body, given appropriate circumstances. Gravitational attraction is the *capacity* of a body, in virtue of its mass, to affect the motions of other bodies. Electrostatic repulsion is the *capacity* of a body, in virtue of its charge, to affect the motion of other like-charged bodies.

Cartwright noted that to explain the motions of two charged bodies, we presuppose

> first, that there is nothing that inhibit either object from exerting both its Coulomb and its gravitational force on the other; second, no other forces are exerted on either body; and third, everything that happens to the two bodies that can affect their motions can be represented as a force.[12]

She emphasized that the above *caveats* refer to capacities and their exercise.

A charged body retains its capacities to affect the motions of other bodies in virtue of its mass and charge, even under conditions that prevent the exercise of these capacities. In chemical interactions, by contrast, powers may be destroyed upon combination. A sodium atom, for instance, has the capacity to form the molecule NaCl in the presence of chlorine. Within the compound, this power of sodium is lost. Sodium, *qua* combined with chlorine, no longer has the capacity to react with chlorine.

Cartwright maintained that the retention of a capacity is the normal state of affairs, and that the loss of a capacity requires an explanation.[13] In the case where sodium combines with chlorine to form salt, one may attribute the lost capacity to the fact that sodium *qua* element is not the same chemical species as sodium *qua* bound to chlorine (a positively-charged ion).

In a universe in which capacities are basic, laws of nature summarize

> the repeated operation of a system of components with stable capacities in particular fortunate circumstances.[14]

Causal laws state generic claims. A causal law may be unpacked as: "*C* causes *E* within population *P* under circumstances *K*," where *C* and *E* are general terms and not names of particulars. Singular causal reports about named particulars provide raw material for the formulation of generic causal claims and may be cited as evidence in support of causal laws.

Cartwright suggested that capacities, laws, and regularities are related in the following hierarchy:

Laws and capacities[15]	
	Ascriptions of capacities
	e.g., "Charged bodies have the capacity to attract oppositely charged bodies."
	Causal laws
Levels of modality	e. g., "Sodium samples react with chlorine"; "Aspirin relieves headaches."
	Functional laws
	e.g., "Boyle's law"; "The function of lungs in homo sapiens is to aerate the blood."
	Probabilistic laws
	e.g., "The probability that an atom of radon ($_{86}Rn^{222}$) emits an α-particle within 3.8 days is 50 percent."
	Occurrent regularities
Non-modal relations	e.g., "All ravens are black."
	"Presidential elections in the United States are preceded by party nominating conventions."

A critic might complain that the separation into "levels of modality" is artificial. After all, "capacity language" may be used at each rung of the "modal hierarchy," e.g., "Sodium samples have the capacity to react with chlorine," "The lungs of *homo sapiens* have the capacity to deliver oxygen to the bloodstream," and "A mass of $_{86}Rn^{222}$ has the capacity to emit α-particles such that, after 3.8 days, 50 percent of that mass remains as $_{86}Rn^{222}$." It would be difficult to base an ordering of levels on some scale of "degree of modality."

Cartwright recognized this. She based the ordering of levels within the hierarchy, not on some scale of "degree of modality," but on an ontological commitment to the Aristotelian concept of "natures."[16] But a universe of entities with distinct "natures" is not yet ordered with respect to levels of modality. Granted that it is the "nature" of sodium to react with chlorine, it also is the "nature" of $_{86}Rn^{222}$ to decay at a specified rate. And it is the "nature" of a confined gas to expand with a diminution of pressure upon it. Thus there are probabilistic relations and functional relations, as well as causal laws, that exist because of the exercise of capacities that issue from the natures of the relevant entities.

What is needed is an ordering of natures with respect to importance. It is the high-level theories of the day that single out the most fundamental entities and capacities. The fundamental entities at present include quarks, leptons, photons, and four basic forces. The ascription of capacities to these fundamental entities belongs at the apex of a modal hierarchy. Judgments about the relative importance of other "natures," and hence their place in a modal hierarchy, remains problematic.

Notes

1. Rom Harré and Edward H. Madden, *Causal Powers* (Oxford: Blackwell, 1975), 105–12.
2. Ibid., 109.
3. Ibid., 86.
4. Ibid., 93.
5. Harré, "Causality and Reality," in *Nature and Scientific Method*, ed. D. O. Dahlstrom (Washington, DC: Catholic University Press, 1991), 5.
6. Ibid., 5.
7. Harré and Madden, *Causal Powers*, 15.
8. Newton's position on gravitational attraction is discussed by I. Bernard Cohen, *Franklin and Newton* (Cambridge, MA: Harvard University Press, 1966), 91–149; A. Koyré, *Newtonian Studies* (Cambridge: Harvard University Press, 1965), 139–63; R. S. Westfall, *Force in Newton's Physics* (New York: American Elsevier, 1971), 323–423.
9. Harré and Madden, *Causal Powers*, 156.
10. Nancy Cartwright, *The Dappled World* (Cambridge: Cambridge University Press, 1999), 105.
11. Cartwright, *Nature's Capacities and Their Measurement* (Oxford: Clarendon Press, 1989), 2–3.
12. Cartwright, *The Dappled World*, 67.
13. Cartwright, *Nature's Capacities and Their Measurement*, 163.
14. Cartwright, *The Dappled World*, 49.
15. Cartwright, *Nature's Capacities and Their Measurement*, 160.
16. Cartwright, *The Dappled World*, 72.

24

Manipulability and Causality

A further challenge to the inferability view arose from consideration of the asymmetries that exist in certain functional relations. For instance, the period of a pendulum is inversely dependent on the square root of its length ($T = k\sqrt{l}$). We can infer its period from the knowledge of its length, and we can infer its length from knowledge of its period. These inferences are symmetrical, but our ability to intervene to change one variable by changing the other is not. Some theorists took the ability to change the value of one variable by manipulating the value of a second variable as a mark of causation. We can *cause* a change of the period of a pendulum by changing its length. We cannot change the length of a pendulum by changing its period. Since the inferability view does not acknowledge this distinction; a manipulability view of causal relatedness is a more appropriate position.

Von Wright on Causal Status

Georg Henrik von Wright championed the manipulability approach to causal analysis in works published in the 1970s. He suggested that

> what makes *p* a cause-factor relative to the effect-factor *q* is . . . the fact that by *manipulating p* . . . we could bring about changes in *q*.[1]

Von Wright pointed out that the concept of "causation" is "secondary" to the concept of a "human action."[2]

Causal relations have objective status in the world independently of human awareness. However, knowledge of causal relations depends on our ability to contemplate counterfactual states of affairs. For instance, *p* is a cause *qua* necessary condition of *q* if, and only if, it *would be* the case that preventing *p* from occurring *would* prevent *q* from occurring. And *p* is a cause *qua* sufficient condition of *q* if, and only if, it *would be* the case that if p had been made to occur then q would have occurred also. Causal analysis cannot succeed solely by inspecting observed regularities. *P*s may be followed by *q*s by accident or because *p*s and *q*s are joint effects of some common cause *c*.

Often we can disqualify accidental correlations from causal status by appropriate interventions. Consider the case of two nearby pendulum clocks of identical length swinging 180° out of phase. Ticks of clock #1 can be shown not to cause ticks of clock #2. Were we to arrest the pendulum of clock #1, there would be no effect on the motions of clock #2. And if both pendulums were made initially stationary in the vertical plane, displacing and releasing the pendulum of clock #1 would not inaugurate motion of the pendulum of clock #2.

Von Wright maintained that we often may exclude from causal status those regularities that result from a common cause. What is required is to manipulate the world so that p takes place in the absence of c. If q follows, we can eliminate the hypothesis that the sequence $p \rightarrow q$ is the result of c. This elimination of the common-cause hypothesis is based on the belief that the world *would have* continued in the *non-p* state in the absence of our manipulation.[3]

Most philosophers of science deny causal status to functional laws such as those of Boyle, Snel, and Ohm. It is clear, however, that these laws pass the manipulability test. In the case of Boyle's law, for example, we can manipulate the pressure of a confined gas and thereby alter its volume.

Von Wright suggested that functional relations such as the above express "causal relationships" rather than "causal laws." Although both "causal laws" and "causal relationships" satisfy the manipulability criterion, "causal laws" state relations between successive states of physical systems, but "causal relationships" do not. Apart from that distinction, nothing much hangs on the decision whether to include among "causal laws" the laws of Boyle, Snel, and Ohm.

Menzies and Price on Causation and Agency

Peter Menzies and Huw Price, in an essay published in 1993, promoted an "agency theory" of causation as an improved version of the manipulability view. They declared that the central thesis of an agency theory is that

> an event *A* is a cause of a distinct event *B* just in case bringing about the occurrence of *A* would be an effective means by which a free agent could bring about the occurrence of *B*.[4]

The agency theory characterization of causation involves a counterfactual claim. To establish that $A \rightarrow B$ is a causal relation, one needs to determine whether an agent could have produced B by producing A. Menzies and Price addressed several of the obvious objections to the agency view.

Critics have complained that the agency view is circular. To characterize causation as a "bringing about" of an effect is to use causal language to characterize causation. Menzies and Price sought to defuse the circularity objection by drawing an analogy to the dispositional theory of color. Just as a person may learn the use of "red" by being shown samples of red surfaces, so also a person may learn the use of "cause" and "effect" by doing one thing and achieving another. Menzies and Price suggested that

> the notion of causation thus arises not, as Hume has it, from our experience of mere succession, but rather from our experience of success.[5]

If we learn causal language ostensibly by applying the terms "cause" and "effect" to successful instances of accomplishing our ends by acting in certain ways, then to speak of "causation" as a "bringing about" is harmless.

A second problem for an agency theory is that there are bona fide causes in nature that are beyond the power of an agent to manipulate. Menzies and Price conceded that although the friction of continental plates caused the 1989 San Francisco earthquake, there is no way that an agent could manipulate the motions of these plates. They maintained that such causes can be accommodated within a modified agency theory that allows analogical inferences from similar means-end relations that are subject to manipulation.[6] Menzies and Price noted that seismologists simulate plate interactions in the laboratory in order to study their effects. This is a case covered by the agency theory. We can extrapolate analogically to the earthquake case, thereby qualifying the plate-tectonics account as a causal explanation, despite the inability of an agent to manipulate continental plates.

On this modified agency theory, the "bringing about" achieved by human agency is analogous to the "bringing about" in cases that cannot be accomplished by human agency. Success at accommodating "unmanipulable causes," for instance, in the domain of astrophysics, depends on the latitude permitted to analogical extrapolation. Consider a case of gravitational lensing. A scientist might state that "the cause of the double image of that quasar is the presence of a strong gravitational field between the quasar and the observer." Granted that this is a genuine causal relation, it is not obvious what the analogical extrapolation from manipulable cases would be in this instance. Of course, one could point to our ability to create a refraction of images by introducing an interface of media, but it is doubtful that Snel's law refraction is a genuine analogy to the gravitational lens case.

Donald Gillies recommended an "action-related theory" of causation to incorporate and supercede the "agency theory" of Menzies and Price. Gillies noted that causal status may be attributed to both "productive actions" and "avoidable actions."[7] He emphasized the importance of avoidance actions in scientific and medical contexts. In an avoidance action, agent x seeks to prevent the occurrence of B by ensuring that A does not occur.

If x succeeds in preventing the occurrence of A, B may not be realized. However, there may be causes of B other than A, such that B occurs despite prevention of the occurrence of A. In addition, background conditions must be taken into account. Gillies observed that, for most cases, the proper statement of a causal law is "A causes B under normal circumstances."[8] Consequently, there are two types of avoidance action to prevent the occurrence of B—x may prevent the occurrence of A directly, or x may alter the circumstances normally present when A causes B.[9]

Gillies was sensitive to the problem of unmanipulable causes such as earthquakes and asteroid–earth collisions. There is no way for an agent to prevent the occurrence of earthquakes or asteroid–earth collisions. Gillies insisted, however, that although unmanipulable causes cannot be preempted, agents nevertheless may perform avoidance actions. In the case of an earthquake, an agent may elect to be far away, and in the case of an asteroid–earth collision, an agent may elect to participate in the network of telescopic observers who seek to alert the public about possible collisions.[10]

On an "action-related theory," these unmanipulable event sequences still qualify as "causal" despite the futility of agents' attempts to intervene. The agents' actions do not qualify as preventions. On the "action-related theory," they then must qualify as altering the normal circumstances under which the earthquake and asteroid–earth collision sequences occur.

An obvious objection is that the agents' actions are "causally irrelevant" to the sequences in question. The fact that an agent takes avoidance action by fleeing does not alter appreciably the "normal circumstances" under which earthquakes are followed by alterations of landscapes, collapses of buildings, loss of life, etc. And the fact that an agent participates in an asteroid-observing program does not alter appreciably the "normal circumstances" under which asteroid–earth collisions are followed by ecological changes, extinctions, etc. (The sequence would not be unmanipulable if foreknowledge of an impending collision enabled scientists to deflect an incoming asteroid by sending suitable explosive devices into space.) Since any action alters the circumstances under which a correlation takes place, we need some way to distinguish "causally relevant" avoidance actions from "causally irrelevant" avoidance actions. The "action-related theory" fails to provide criteria to enforce this distinction.

Pearl on Causal Models

Judea Pearl has created a mathematically sophisticated framework to uncover causal relations in complex cases. Pearl's starting point is a qualitative concept of causation—"*c* is an immediate cause of *e*." This concept is treated as a primitive from which a framework is developed for the estimation of quantitative measures of causation. The framework utilizes functional causal models that consist of sets of equations:

$$x_i = f_i\left(\mathrm{pa}_i, u_i\right), \qquad i = 1, \ldots, n,$$

where pa_i (connoting parents) stands for the set of variables judged to be immediate causes of X_i and U represents errors (or "disturbances") due to omitted factors.[11]

Pearl displayed these models as network diagrams in which arrows stand for immediate causes. An example given by Pearl focuses on a section of pavement (Figure 24.1[12]).

Pearl called attention to the interventions that may be performed within causal networks. Given antecedent probability estimates for the immediate causal relations in the network, one then can assess the quantitative changes that result from an intervention. For example, suppose someone intervenes to turn on the sprinkler. The intervention is introduced into the diagram by deleting the link $x_1 \rightarrow x_3$ (Figure 24.2[13]). The probabilities within the network then are recalculated with x_3 set to "on."

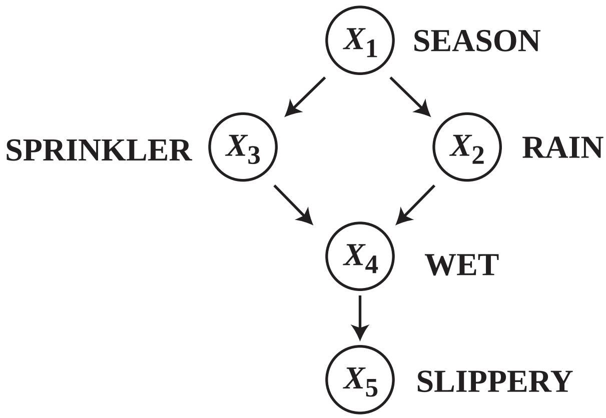


Figure 24.1 The slippery pavement network.

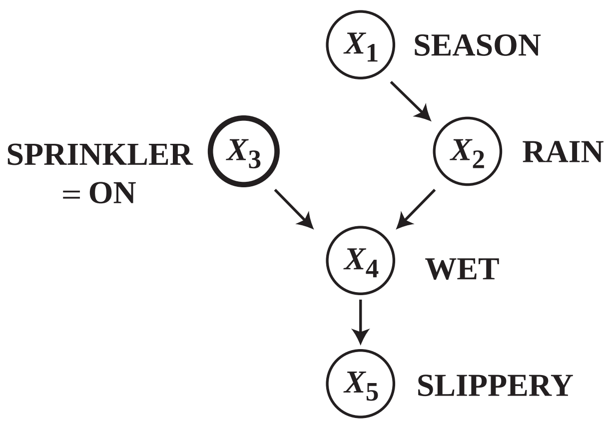


Figure 24.2 The sprinkler turned "on."

Pearl introduced the concept of a "do-operator" that represents an intervention. He referred to interventions as "surgery over equations."[14]

Pearl maintained that his causal models can provide answers to questions about

1. *predictions* (e.g., would the pavement be slippery if we *find* the sprinkler off?),
2. *interventions* (e.g., would the pavement be slippery if we *make sure* that the sprinkler is off?), and

3. *counterfactuals* (e.g., would the pavement be slippery *had* the sprinkler been off, given that the pavement is in fact not slippery and the sprinkler is on?).[15]

The causal modeling program has achieved certain successes. Given a set of immediate causal relations and the pertinent antecedent probability estimates for these relations, one can perform various calculations by making changes ("interventions") among these relations. James Woodward pointed out, however, that Pearl leaves undetermined the immediate, direct, causal relations, and that a theory of causation ought to seek to uncover the "meaning and content" of these relations.[16]

Woodward's Version of the Manipulability Theory

James Woodward recently has developed a version of the manipulability view of causation. He conceded at the outset that his version of the theory is a non-reductive interpretation.[17] He did not propose to explicate "causation" by reference to *prima facie* noncausal concepts such as "constant sequential conjunction" or "possession (or transfer) of a conserved quantity."

According to Woodward, the mark of a causal relation is that it furnishes information that is "potentially relevant to manipulation and control."[18] Included among causal relations, thus understood, are the collisions of billiard balls, the expansion of heated balloons, and the eventual convergence of initially diverse sex ratios to the ratio 1:1.[19]

Woodward unpacked the concept "cause" by reference to the concepts "phenomena," "information," "intervention," "counterfactuality," and "invariance." He declared that

> the notion of information that is relevant to causally explaining an outcome involves the identification of factors and relationships such that *if* (perhaps contrary to fact) manipulation of those factors *were* possible, this would be a way of manipulating or altering the phenomenon in question.[20]

Woodward noted that we often assess causal status by performing interventions. For instance, the relationship between falling barometer readings and subsequent storms is a candidate for causal status. We judge this relationship to be noncausal because interventions that change barometer readings do not alter the likelihood of storms. Conversely, if an intervention that changes the value of some variable X under specified background conditions does change the value of some other variable *Y*, then we judge that *X* and *Y* are causally related. For instance, changing the temperature of a rigidly confined gas does change its pressure. We should judge that this is a causal relationship.

Woodward set forth sufficient and necessary conditions for causation:

1. Sufficient condition

> If (i) there is a possible intervention that changes the value of *X* such that (ii) carrying out this intervention (and no other interventions) will change the value of *Y*, or the probability distribution of *Y*, then *X* causes *Y*.

2. Necessary condition

> If *X* is a cause of *Y* then (i) there is a possible intervention that changes the value of *X* such that (ii) if this intervention (and no other interventions) were carried out, the value of *Y* (or the probability of some value of *Y*) would change.[21]

Woodward insisted that "possible intervention" be interpreted broadly. He did not wish to restrict the phrase to those operations that fall within the technical capacities of human beings at the time in question. Indeed, he made the extreme claim that

> causal claims of the form *X*'s cause *Y*'s can be true even if human beings lack the power to manipulate *X*'s and even if manipulation of *X*'s is nomologically impossible.[22]

A "nomologically impossible" manipulation of *X*s is one that violates a law of nature. One such law is the constancy of the velocity of light in all reference frames. Woodward's broad interpretation of "possible interventions" brings astronomical events within the scope of causal analysis and even allows for the possibility of instantaneous causal influence at a distance.

Woodward noted that his version of the manipulability theory implies that causal explanations are contrastive.[23] Causal claims are claims about the results of hypothetical manipulations—e.g., the value of *Y* would change in a specific way if the value of *X* were changed in a specific way, given circumstances of type *Z*. As such, a causal explanation rules out other possible outcomes of a specific manipulation.

He noted also that his version of the manipulability theory links causal claims and claims that certain associated generalizations are true. Type-level causal claims imply that relations of invariance hold. Woodward defended a "weak" view of invariance. He declared that

> I count a generalization as invariant or stable across certain changes if it holds up to some appropriate level of approximation across those changes.[24]

On Woodward's view, generalizations may display greater or lesser degrees of invariance. What counts as an "appropriate level of approximation" remains to be specified. Debates about whether a certain degree of invariance is sufficient to warrant a relation as "causal" are to be expected. Woodward insisted only that for an explanation to qualify as causal, there must be (implicit) reference to a generalization that possesses *some* degree of invariance.

The invariance of a generalization may be assessed in response to different types of interventions. One may test whether a generalization continues to hold as various boundary conditions are changed. Or one may test whether it continues to hold as the value of a variable mentioned in the generalization is changed.

To base judgments of causal status on the invariance of a generalization to changes in boundary conditions would be inappropriate. Boundary conditions under which a generalization is utilized are simply too numerous and varied. For example, background conditions under which Boyle's law is applied include, not only the mass and temperature of the gas, but also lighting conditions, electromagnetic and gravitational fields, the number of observers present, and diverse other factors.

Woodward maintained that what counts for the determination of causal status are the results of interventions that alter the value of variables that occur in a generalization. If values calculated from a generalization agree with what is observed (to an "appropriate" level of approximation) for a significant range of interventions that change the value of one of its variables, then the generalization is "invariant" (to some degree).

In a few cases, a comparison of "degrees of invariance" can be made by reference to a subset relation. An example is the relationship between Van der Waals law—$(P + a/V^2)(V - b) = kT$—and the ideal gas law—$(PV = kT)$. Interventions that change the value of the pressure of a gas at constant temperature change the value of its volume. The Van der Waals law holds for a greater range of pressure-changing interventions than does the ideal gas law. It therefore possesses the higher degree of invariance. Comparisons are easy where a subset relation such as this is obtained. In other cases, assessment of degrees of invariance depends on judgments about the relative importance of different interventions.

Woodward presented a quite inclusive characterization of causal relatedness. His version of the manipulability theory qualifies as "causal" relations that involve invariance to an "appropriate level of approximation" and allows "manipulations" which not only are beyond the present technical capacity of human beings but which also are "nomologically impossible." He emphasized, in addition, that his version of the manipulability theory is consistent with nonlocality and discontinuous spatiotemporal interaction.[25]

Notes

1. Georg Hendrik von Wright, "On the Logic and Epistemology of the Causal Relation," in *Causation*, ed. Ernest Sosa and Michael Tooley (Oxford: Oxford University Press, 1993), 118.
2. Von Wright, *Causality and Determinism* (New York: Columbia University Press, 1974), 1–2.
3. Von Wright, "On the Logic and Epistemology of the Causal Relation," 116–17.
4. Peter Menzies and Huw Price, "Causality as a Secondary Quality," *British Journal for the Philosophy of Science* 44 (1993): 194.
5. Ibid., 194.
6. Ibid., 195–98.
7. Donald Gillies, "An Action-Related Theory of Causality," *British Journal for the Philosophy of Science* 56 (2005): 827.
8. Ibid., 836.
9. Ibid., 828.
10. Ibid., 837–38.
11. Judea Pearl, *Causality: Models, Reasoning, and Inference* (Cambridge: Cambridge University Press, 2000), 27.
12. Ibid., 15.
13. Ibid., 23.
14. Ibid., 344.
15. Ibid., 29.
16. James Woodward, *Making Things Happen* (Oxford: Oxford University Press, 2003), 38.
17. Ibid., 20–22.
18. Ibid., 6.
19. Ibid., 6.
20. Ibid., 10.
21. Ibid., 45.
22. Ibid., 150.
23. Ibid., 146.
24. Ibid., 239.
25. Ibid., 147–49.

25

David Lewis and the Counterfactual Conditional View

Causal Relation as a Contrary-to-Fact-Conditional

Some causal theorists have expressed a preference for Hume's second definition of "cause":

> *c* is the cause of *e*, if, and only if, if *c* had not occurred, then *e* would not have occurred.[1]

In the usual symbolic notation, '*c* causes *e*' ≡ '$(\sim c\square \rightarrow \sim e)$, where '$c \rightarrow \square e$' is read, 'if *c* had occurred, then *e* would have occurred.'"

On this definition, the impact of a stone and a glass plate is the cause of its shattering if, and only if, the glass would have remained intact if this impact had not occurred. This seems reasonable, provided that no other cause of shattering is present at the time in question. By contrast, the concomitant beating of the wings of a nearby hummingbird is not the cause of shattering because it is false that the glass would have remained intact if this beating had not occurred. To support this judgment, one may note that there is a regularity that links stone-on-glass impacts and shattering, but no similar regularity that links hummingbird-wing-produced air vibrations and shattering.

On the above definition, causes may be events, conditions, or processes. It is not immediately obvious, however, how causal attributions are to be assessed. By definition, *c* did occur. How can we determine what would have been the result if *c* had not occurred? We cannot replay the situation. It has passed.

Necessarily, the assessment of counterfactual claims will be indirect. One approach is to contemplate how a world that is just like our own, but in which event *c* did not occur, would evolve. If event *e* did not occur in this alternative world, then we might take this to establish that the original contrary-to-fact conditional claim is true of our world.

One possibility is that in the alternative world the indicative relation [~Ca ⊃ ~Ea] follows from a true generalization of the form [(*x*) (E*x* ⊃ C*x*)]. If so, this may be taken to support the causal claim that if Ca had not occurred, then Ea would not have occurred. This strategy is unavailing in the case of the shattering of the glass plate. It is false that "every shattering of a glass plate is the result of an impact with a stone." But the strategy does appear to be successful in certain other cases, e.g.,

contrary-to-fact conditional claim	**generalization**
If that barium-containing substance had not been placed in that flame, then the flame would not have turned that shade of green	All cases in which a substance turns a flame that shade of green are cases in which a barium-containing substance is placed in that flame
If that child did not have blue eyes, then its parents would not both have blue eyes	All children of blue eyed parents have blue eyes
If the length of that pendulum had not been reduced to one-fourth its original length, then its period would not be one-half of what it is	For all cases, the period of a pendulum is proportional to the square root of its length

Given the deductive relationship between the indicative conditional statements and the corresponding generalizations, one may claim that (1) placing the barium-containing substance in the flame caused the flame to turn that shade of green; (2) the parents' blue eyes caused their offspring to have blue eyes, and (3) reducing the pendulum's length to one-fourth its original value caused its period to be one-half its former value.

Such assignments of causal status are effective only if the general laws cited are "genuine" and not merely "accidental." The "pendulum-length-period" correlation is a law. Contrast the generalization that each activation of a factory whistle is followed by an egress of workers from the factory. This generalization presumably is merely accidental.

The contrast between lawful and accidental generalizations traditionally is established in either of two ways. The first way is to note that lawful regularities support contrary-to-fact conditional claims and accidental correlations do not. The pendulum generalization supports the claim that if the length of that pendulum were reduced to one-fourth its present value, then its period would be reduced to one-half its present value. The factory-whistle generalization does not support the claim that if the whistle had not blown then the workers would have remained inside (I assume that some workers consult watches from time to time). This first strategy provides no support for the "cause-as-contrary-to-fact-conditional" view. It is unacceptably circular to assess the causal status of a contrary-to-fact-conditional (causal) claim by reference to a deductive relation to a generalization whose lawful status depends on accepting similar contrary-to-fact conditional claims.

A second way to establish the lawful status of a generalization is to show that the generalization is embedded in a theory. The length–period relationship for simple pendulums qualifies on this score. It can be exhibited as a theorem within a suitable axiomatization of Newtonian mechanics. As a theorem, the pendulum law holds only for low-amplitude swings in an idealized universe in which there are no energy losses due to internal friction or air resistance. In the real universe, this law, like Galileo's law of falling bodies and Kepler's laws of planetary motion, is false. Nevertheless, the theory dependence of a false generalization often is taken to confer lawful status upon it. This is the case even though scientists concede that the generalization is an approximation subject to *ceteris paribus* restrictions.

Particularly impressive are cases in which determinations of the value of a physical constant using diverse laws converge upon the same number. For instance, calculations of Avogadro's number from laws applicable to Brownian motion, electrolysis, radioactive decay, black-body radiation, the viscosity of gases, etc., converge upon the value 6.02×10^{23} molecules/gram molecular weight. This convergence may be taken to warrant the various laws that are embedded in the atomic-molecular theory of matter.[2] A similar warrant is provided for Planck's black-body radiation law, Einstein's law of the photoelectric effect, and various laws governing the absorption and emission spectra of molecules, since calculations of Planck's constant utilizing these laws converge upon the value 6.62×10^{-27} erg-sec.[3]

On the basis of these considerations, causal claims may be assessed as follows:

"*c* causes *e*" if, and only if, in a universe just like our own in which *c* did not occur

1. "not *e*" is true,
2. "if ~*c* then ~*e*" follows from a generalization, which generalization is,
3. a lawful regularity, established by showing that,
4. the generalization is embedded in a theory, which theory,
5. is used currently by scientists.

These requirements make the assessment of causal status depend on which theories currently are applied in the sciences. This is not objectionable. Our knowledge of causes has changed over time. This is to be expected if our causal knowledge is derived from a succession of increasingly more adequate theories.

Nevertheless, there are problems with the above position on causation. In the first place, the contrary-to-fact-conditional view of cause is subject to the same difficulties about cases of overdetermination that beset the "cause-as-necessary-condition" view.

In the second place, there are numerous events, normally accepted as causes, which cannot be linked to lawful regularities that are embedded in theories. Moreover, it seems obvious that collision processes were causal even before generalizations about the conservation of momentum and energy became embedded in a theory.

And in the third place, there are difficulties about the notion of "a universe just like our own." Can there be an alternative world "identical to our own," but in which event *c* did not occur? We may insist that the laws of nature and the values of fundamental physical constants be the same in both worlds. But event *c* in our world is set within a specific spatiotemporal context. In particular, certain events preceded it. Could there be an alternative world in which all these prior events occurred, but *c* did not occur? Is it possible to change one segment of an interconnected web of events without repercussions elsewhere in the web? If not, then we cannot cash the notion of "a universe just like our own, but in which *c* does not occur."

David Lewis on "Cause" and Counterfactual Claims: 1973–1986

David Lewis has developed a more sophisticated version of the contrary-to-fact conditional interpretation of causal relation. He acknowledged as starting point Hume's second definition of "cause," viz., "*c* causes *e* if, and only if, if *c* had not occurred, then *e* would not have occurred either."[4] Lewis began with definitions of "counterfactual dependence" and "causal dependence," introduced the notions of "cause" and "causal chain," and then specified a procedure for identifying causes.[5]

Counterfactual Dependence

Given events *c* and *e* and propositions $O(c)$ and $O(e)$, which assert that *c* and *e* occur, the family $[O(e), \sim O(e)]$ counterfactually depends on the family $[O(c), \sim O(c)]$ if, and only if,

1. $O(c) \Box\!\!\rightarrow O(e)$ and

2. $\sim O(c) \Box\!\!\rightarrow \sim O(e)$.

Causal Dependence

e causally depends on *c* if, and only if, the family $[O(e), \sim O(e)]$ counterfactually depends on the family $[O(c), \sim O(c)]$.

Cause

If $[O(c), O(e)]$ is true and *e* is counterfactually dependent on *c*, then *c* is the cause of *e*.

Causal Chain

If *c* is the cause of *e*, then there exists a chain of causally dependent events with *c* at one end and *e* at the other end.

Lewis thus took the causal relation between distinct events *c* and *e* to be the ancestral to the relation of counterfactual dependence. In the simplest case, *c* causes *e* if, and only if, $(\sim c \Box\!\!\rightarrow \sim e)$, given that *c* and *e* both occur. In complex cases, *c* causes *e*, provided that there is some set of intermediate events $x_1, x_2, \ldots, x_n$, such that $[(\sim c \Box\!\!\rightarrow \sim x_1)$ & $(\sim x_1 \Box\!\!\rightarrow \sim x_2)$ & … $(\sim x_{n-1} \sim x_n)$ & $(\sim x_n \Box\!\!\rightarrow \sim e)]$. Lewis excluded "backtracking conditionals" from the set of counterfactual conditionals permitted in causal relations. Backtracking conditionals are claims about how a past event would have been different if a present event were different. An example, previously cited (p. 154[c3]), is "if that child did not have blue eyes, then its parents would not both have blue eyes."

How are we to determine whether *c* is the cause of *e*? Lewis suggested that $[O(c) \Box\!\!\rightarrow O(e)]$ is true if, and only if, there exists a possible world in which $O(c)$ and $O(e)$ both are true, and this possible world is more similar to the actual world than any world in which *O*(*c*) is true and *O*(*e*) is false. Lewis declared that

> a counterfactual is non-vacuously true if it takes less of a departure from actuality to make the consequent true along with the antecedent than it does to make the antecedent true without the consequent.[6]

To make such evaluation more than an exercise of the imagination, we need to apply standards of similarity to alternative possible worlds. Lewis conceded that there is an uneliminable degree of vagueness to the characterization of "similar" possible worlds.[7] He insisted, however, that agreement of general laws usually is more important than local differences among matters of fact.[8]

Suppose I strike a cue ball so that it impacts an 8-ball and drives it the length of the table into the left corner pocket. Is the impact the cause of the 8-ball's descent into the pocket? Two possible alternative worlds are:

> W_A—a world in which Newton's laws are not obeyed and *O*(*c*) is true but *O*(*e*) is false.
>
> W_B—a world in which Newton's laws are obeyed, but in which there is an additional chalk cube on the table railing. In this world, *O*(*c*) is true and *O*(*e*) is true.

If W_B is more similar to the actual world than is W_A then we may take this to support the claim that *O*(*c*) is the cause of *O*(*e*). This falls short of a demonstration that *O*(*c*) causes *O*(*e*). To achieve this, one would have to show that there is *no* alternative world that is (1) the closest possible and (2) a world in which *O*(*c*) is true and *O*(*e*) is false.

Consider alternative world W_C

> W_C—a world in which Newton's laws are obeyed and there is no extra chalk cube present, but the billiards table is not perfectly level. The cue ball strikes the 8-ball exactly as in the actual world, but the 8-ball does not fall into the far left corner pocket.

W_B and W_C both differ only slightly from the actual world. Can we establish that W_B is more similar to the actual world than is W_C without appealing to our knowledge about which factors are causally relevant to the motion of a billiard ball? Unless this can be done, Lewis's program for assessing causal status involves an unacceptable circularity.

Jonathan Bennett formulated a *prima facie* causal relation that poses a problem for Lewis's appeal to worlds similar to our own.[9] Jones is attending an auction. He hopes to purchase the armoire currently on the block. The auctioneer announces that bidding will begin at $2,000. Jones raises his right hand and the auctioneer acknowledges his bid. Presumably there is a causal relationship between Jones raising his right hand (R) and the auctioneer's acknowledgment of his bid (A).

But is it true that if R had not occurred then A would not have occurred? In one closely related world, R is false but A is true. In this world, Jones instead raises his left hand. In a second closely related world, Jones nods his head, and his bid is acknowledged. Again R is false but A is true. Since these alternative worlds are nearly identical to the actual world, it would seem that, on Lewis's theory of causation, R is not the cause of A.

One defense of the contrary-to-fact-conditional analysis against this type of countercase is to redescribe the effect. In its extreme form, the effect is characterized by reference to the alleged cause. If the effect to be analyzed is not just "the auctioneer acknowledged Jones's bid" but rather "the auctioneer acknowledged a bid made by Jones raising his right hand" (A^*), then $\sim R \Box\!\!\rightarrow \sim A^*$ is true.

This type of defense does some violence to our intuitions. We were concerned initially about the cause of the auctioneer's acknowledgment of Jones's bid. We are told that our concern was misplaced and that we should have been concerned about the cause of the auctioneer's acknowledgment of the bid Jones made by raising his right hand (in a certain manner). The contrary-to-fact-conditional view remains unscathed, but we may not be willing to pay the price.

Overdetermination is a problem for the counterfactual dependence view as well. Suppose Adams and Butler simultaneously fire bullets into a pane of glass. The glass shatters. But since both the counterfactuals (1) "If Adams had not fired, then the glass would not have shattered" and (2) "If Butler had not fired, then the glass would not have shattered" are false, it would seem that there is no cause of the shattering.

This conclusion can be avoided by redescribing the effect as, e.g., "the shattering of the glass upon impact of a bullet from Adams's gun." Causation is restored. The counterfactual "If Adams had not fired, then the glass would not have been shattered upon impact from a bullet from Adams's gun" is true.

Cases of preemption also pose a problem for the counterfactual analysis of causation. Suppose David launches a stone at Goliath. It strikes his forehead, killing him instantly. At the time of David's throw, William Tell was prepared to shoot an arrow also directed at Goliath's forehead. Had William released the arrow, it would have struck Goliath, had he not been felled by David's stone a moment earlier. The effect that might have been the result of William's action was preempted by David's prior action.

We want to say that David's action caused Goliath's death. On the counterfactual analysis, David's action caused Goliath's death, if, and only if, had David's stone not struck Goliath, Goliath would not have died. The counterfactual presumably is false. Had David missed, William would have released an arrow fatal to Goliath. But if the counterfactual is false, so also is the claim that David's stone caused Goliath's death. This is counterintuitive.

One option in this situation is to select the redescription approach. If we distinguish "Goliath's death from the impact of the stone at time $t + \Delta t$ (D_s)" and "Goliath's death from the impact of William's arrow at time $t + \Delta t$ (D_a)," then the relevant counterfactual is ($\sim S \Box\!\!\rightarrow \sim D_s$), where S is David's launching the stone. Since "$\sim S \Box\!\!\rightarrow \sim D_s$" is true, S caused D_s, in spite of the action of William Tell.

There is a price to be paid for selecting the redescription approach, however. Michael McDermott emphasized that implementation of this approach may lead to a proliferation of false causal attributions. Consider the case of death row inmate Don, who is scheduled to be electrocuted at time *t*. The time of execution is delayed as a result of the malfunction of a clock. The actual death (*qua* delayed) would not have happened in the absence of the malfunction. On the counterfactual analysis, the clock malfunction qualifies as a cause of the redescribed "delayed death." But if we consider Don's death, *per se*, it is clear that the clock malfunction is not on a par with the jury's verdict, the denial of appeal, the governor's refusal to issue a stay of execution, and the passage of electricity through Don's body. We would be unwilling to say that the clock malfunction is a cause of Don's death. Moreover, there are various other events that might have delayed the execution, e.g., a power failure, the warden's heart attack, Don's fainting spell, etc. Counterfactual analysis would qualify any one of these events, if actual, as a cause of Don's demise.

McDermott insisted that there may be many factors that affect the time at which an event occurs without being causes of the event.[10] Of course, it is open to the redescription enthusiast to claim, for instance, that Don's execution at time t is not the same event as Don's execution at time $t + \Delta t$.

Lewis developed an interesting defense of counterfactual analysis against preemptive countercases.[11] He suggested that C may cause E in the absence of direct counterfactual support provided that there is an event X (or events $X_1, X_2, \ldots, X_n$) such that

1. $\sim C \Box\!\!\rightarrow \sim X$ and

2. $\sim X \Box\!\!\rightarrow \sim E$.

On the other hand, if there is neither direct counterfactual dependence of E upon C nor a chain of counterfactual dependencies, then C and E are not causally related.

One might call this the "chain of counterfactual dependencies" defense. Lewis's position is effective against cases of *early* preemption. Consider the David–Goliath–Tell example. William Tell was poised to release an arrow but did not do so because of David's preemptive strike. The counterfactual ($\sim S \Box\!\!\rightarrow \sim D$) is false, and this might lead one to conclude that David's hurling the stone (S) is not the cause of Goliath's death (D). However, there is a sequence of events $X_1, X_2, \ldots, X_n$, representing successive positions of the stone on its way to Goliath. The following counterfactual claims are true of this sequence: ($\sim S \Box\!\!\rightarrow \sim X_1$), ($\sim X_1 \Box\!\!\rightarrow \sim X_2$), … ($\sim X_{n-1} \Box\!\!\rightarrow \sim X_n$), ($\sim X_n \Box\!\!\rightarrow \sim D$). There are no intermediate events linking actions of William Tell and Goliath's death. The causal status of David's throw is restored.

Unfortunately, the chain of counterfactual dependencies defense is not effective in cases of *late* preemption. Suppose William Tell shoots an arrow just after David's launch of the stone. In this case there is a sequence of events representing intermediate positions of the arrow $Y_1, Y_2, \ldots, Y_n$. The counterfactual ($\sim Y_n \Box\!\!\rightarrow \sim D$) is false. There is a failed intermediate step in the backup process inaugurated by Tell. So far so good. However, in this case of late preemption, the counterfactual ($\sim X_n \Box\!\!\rightarrow \sim D$) also is false, since Tell's arrow would have killed Goliath in the absence of David's stone. Given the existence of late preemption, David's action is not a cause of Goliath's death on the chain of counterfactual dependencies theory. This, of course, is a counterintuitive result.

Cases of simultaneous preemption, or "trumping," highlight the inadequacy of the chain of counterfactual dependencies defense. A number of such examples of trumping have been proposed. Perhaps the most effective is the one developed by Jonathan Schaffer.[12] In a battle situation, major and sergeant simultaneously shout "charge." The soldiers subject to their command advance. There is a good reason to believe that there is a causal relationship between the major's order and the action of the troops. There is a problem, however. The relevant counterfactual conditional is "if the major had not shouted 'charge,' then the soldiers would not have advanced." But this is false. In the absence of a command from the major, the soldiers still would have advanced on the command from their sergeant. On the contrary-to-fact conditional interpretation, it would seem that if the major alone shouts "charge," he causes his troops to advance, but when the major and the sergeant both shout "charge" and the troops charge, there is no causal relationship.

Schaffer maintained that this case of trumping cannot be accommodated to a contrary-to-fact conditional interpretation. Lewis had responded to earlier examples of causal preemption by pointing to failed intermediary processes responsible for preemption. Schaffer noted that this strategy is not available for cases of trumping. The sergeant's order is present from the time the major shouts "charge" to the time the effect takes place. There is no intermediate process present between the two shouts of charge and the soldiers' advance, but absent in the case in which only the major issues the command. Since there is no intermediate link in the causal chain, the contrary-to-fact conditional interpretation of causation cannot be defended by calling attention to its failure.[13] Lewis acknowledged the force of trumping examples and subsequently revised his original counterfactual conditional analysis to accommodate "causal influences" not subject to the original analysis.[14]

David Lewis on Causation as Influence

In "Causation as Influence" (2000), Lewis broadened his analysis by subsuming the original concept of "causal relation as the ancestral of counterfactual dependence" under a concept of "causal influence." He noted that the original concept fits a significant range of causal interaction. It is a sufficient condition of "*c* causes *e*" that there be the appropriate relations of counterfactual dependence (subject to the no-backtracking constraint). However, the original concept of "causal relation" is not a necessary condition of "*c* causes *e*." There are other types of causation, among them trumping preemption and causation-by-omission. Lewis developed a concept of "causal influence" to account for them. He declared that, for distinct events *C* and *E*,

> *C* influences *E* if there is a substantial range C_1, C_2, . . . of different not-too-distant alterations of *C* (including the actual alteration of *C*) and there is a range E_1, E_2, . . . of alterations of *E*, at least some of which differ, such that if C_1 had occurred, E_1 would have occurred, and if C_2 had occurred, E_2 would have occurred, and so on. Thus we have a pattern of counterfactual dependence of whether, when, and how on whether, when and how. (As before, Causation is the ancestral: *C* causes *E* if there is a chain of stepwise influence from *C* to *E*).[15]

Causal influence admits of degrees. It allows us to resolve cases of trumping by highlighting hidden asymmetries. Consider once again the case presented by Schaffer. Major and sergeant simultaneously shout "charge." The troops do so. Had the major remained silent, the troops still would have advanced. On Lewis's original counterfactual analysis, the major's order does not cause the advance, since "$\sim M \Box\!\!\rightarrow \sim A$" is false, and there is no intermediate chain of counterfactual dependencies from *M* to *A*. As Schaffer emphasized, this is counterintuitive.

Lewis shifted attention to the question of causal influence. Variations in the major's command, ranging from "stay" to "retreat" would have been accompanied by very different responses by the troops. Variations in the sergeant's command, given no change in the major's command, would have been accompanied by little or no change in response. The major's command fulfills the requirements of causal influence, whereas the sergeant's command does not.

But suppose that it is two sergeants who simultaneously shout "charge." To defend the advance of the troops as the effect of a cause, the causal influence theorist would have to show that the troops respond differently to small differences in loudness, pitch, or posture in the commands of sergeant #1 and sergeant #2. Trumping countercases are not easily dismissed.

Although the "causal influence" theory supposedly subsumes and expands upon the original "counterfactual dependence" theory, the two theories provide different answers to overdetermination cases. For example, suppose that two adjacent rigidly mounted guns fire simultaneously at the center of a glass plate. The plate shatters. Is the firing of gun #1 a cause of the shattering? Small variations in the direction or timing of firing of gun #1, with firing of gun #2 held constant, would result in slightly different patterns of glass shards. However, small variations in the direction or timing of firing of gun #2, with firing of gun #1 held constant, also would result in slightly different patterns of glass shards. On the "causal influence" theory, it would seem that both firings cause the shattering; on the original "counterfactual dependence" theory, neither firing is a cause of the initial shattering (see above, p. 159). For this type of case, the "causal influence" theory provides the more plausible analysis.

Absences may enter causal relations as causes or effects. In "causation-by-omission," absences are causes. In cases of prevention, absences are effects.

Causation-by-omission is a difficult problem for counterfactual analysis. My failure to deposit my latest paycheck caused the bank to return without payment the last check I wrote. A counterfactual analysis seems appropriate. Had I deposited my paycheck, the bank would not have returned unpaid my last check. But then my failure to deposit the paycheck is a cause of the bank's return of the check without payment.

Lewis began his discussion of causal influence by restricting its scope to distinct events. Omissions are not events. In particular, my failure to deposit my paycheck is not an event. To declare that there is an omission is to state a negative existential claim. In the case under consideration, the claim is "it is not the case that there exists my last-paycheck-depositing," Lewis maintained that it is "only by way of such propositions, that absences enter into patterns of counterfactual dependence."[16] Of course, propositions are not causes or effects. Lewis rejected the position that negative existential propositions cause events to take place. Only events can be causes. But then there is nothing that is the cause of an effect in the case of causation-by-omission.

Preventive preemption also is a problem for the theory of causal influence. In a much-discussed example, a thrown ball is speeding toward a plate glass window. Consider the following variants:

1. Fielder *Y* is positioned on a line between Fielder *X* and the window. *X* catches the ball, but had he not done so, *Y* would have caught it.

2. There is a brick wall between Fielder *X* and the window. *X* catches the ball, but had he not done so, the ball would have struck the wall.[17]
3. The window is on the moon at the earth's horizon. *X* catches the ball, but had he not done so, the ball would have failed to reach the window.[18]

In which case, or cases, does *X*'s catch prevent the window from shattering, i.e., *cause* the window to remain intact? Assigning causal credit to *X* is completely far-fetched if the window is on the moon, doubtful if the window is behind a wall, and debatable if Fielder *Y* is the backup in case *X* fails to catch the ball.

Lewis accepted the reality of causation-by-omission and preventive preemption. He concluded the essay "Causation as Influence" with the recommendation that analyses of cases of preventive preemption reflect the "ambiguities" present in these cases. For instance, in McDermott's example of "pitcher–fielder–wall–window," the causal influence theory qualifies as "causal" the relation between Fielder *X*'s action (C) and the absence of the ball behind Fielder *X* (D). But we may be uncertain about the causal status of the relation between the absence of the ball behind Fielder *X* (D) and the pristine condition of the window (E).

The causal influence theory makes reference to alterations "not too distant" from the events under consideration. In his discussion of causation and absences, Lewis applied the "not too distant" qualification to absences. Certain alterations of absence (*D*) lead to a breaking of the window. For instance, the ball might have struck Fielder *X*'s mitt, described an arc over the wall, and broken the window. Is this alteration "not too distant" or far-fetched? Lewis maintained that

> if we are in a mood to think them [it] relevant, we should conclude that *D* causes *E*, and by transitivity *C* also causes *E*.[19]

On the other hand, if we are in a mood to think the alteration too "distant," then we should deny causal influence from *D* to *E*, and hence from *C* to *E*. It would appear that determination of mood is important in assessing causal influence. It would be helpful to have criteria of appropriateness of the respective moods. Lewis provided no assistance at this point, over and above the suggestion that "ambiguities" in the situation be acknowledged.

Helen Beebe applied Lewis's theory of explanation to reconcile the apparent conflict between the common sense belief in causation-by-absence and the fact that the absences are not events at all. Lewis had suggested that to explain an event is to "provide some information about its causal history."[20] This suggests that there is an important distinction between causal explanation and assertions of causal relatedness. Perhaps some causal explanations are not statements of causal relatedness. Beebe declared that

> a more ontologically perspicuous way of saying that Flora's failure to water the orchids caused their death is to say that the orchids died—or that their death occurred—because Flora failed to water them.[21]

According to Beebe, there is no such thing as causation-by-absence. But reference to absences occurs in explanations that implicitly contrast the causal history of an event with causal histories in closely related possible worlds. In one such world, Flora did water the orchids and the plants utilized the water to sustain life functions. Unfortunately, in the actual world, the history of the orchids is different. The causal explanation of the orchids' demise accounts for this difference. It does not state that Flora's failure to provide water was a cause of death. Thus counterfactual analysis does apply to cases of causation-by-absence *via* causal explanations, but not to "causal relations" that involve nonevents.

Is it a problem for a theory of causal explanation that the Queen of England also failed to water the orchids? Beebe thought not. Causal explanations of an event are not all on a par. Given the context within which a question is posed about the death of the orchids, reference to Flora's failure is more relevant than referring to the Queen's failure. The causal explanation that cites Flora's neglect clearly is the superior explanation. This is the case despite the fact that both "the orchids died because Flora neglected to water them" and "the orchids died because the Queen neglected to water them" are true.[22]

It is worth noting that physicists sometimes speak of forces as causes. They do not restrict the label "cause" to "events." David Fair took account of this usage. He insisted that momentum transfer alone is causal even if there is no accompanying energy transfer. In some cases, forces cause nonevents. Consider body x subject to opposed equal forces F_1 and F_2 ($F_2 = {\sim}F_1$). This is a special case of the composition of forces. Resultant force $F_R = 0$ and $a_x = 0$. Physicists take F_R to be the cause of the stationary position of body x.

Notes

1. David Hume, *An Enquiry Concerning Human Understanding* (Chicago: Open Court, 1927), 79.
2. See Jean Perrin, *Atoms* (1913), trans. D. L. Hammick (New York: Van Nostrand, 1923), 215–17.
3. See Max Planck, *A Survey of Physics* (London: Methuen, 1925), 162–77.
4. David Lewis, "Causation," *Journal of Philosophy* 70 (1973); Reprinted in *Causation*, ed. Ernest Sosa and Michael Tooley (Oxford: Oxford University Press, 1993), 194.
5. Ibid., 198–200.
6. Ibid., 197.
7. Ibid., 197.
8. Lewis, "Counterfactual Dependence and Time's Arrow," in *Philosophical Papers, Vol. 2* (Oxford: Oxford University Press, 1986), 173–213.
9. Jonathan Bennett, "Event Causation," in *Causation*, ed. E. Sosa and M. Tooley, 219–10.
10. Michael McDermott, "Redundant Causation," *British Journal for the Philosophy of Science* 40 (1995): 530.
11. Lewis, "Causation," in *Philosophical Papers, Vol. 2*, 159–213.
12. Jonathan Schaffer, "Trumping Preemption," *Journal of Philosophy* 97 (2000): 175–76.
13. Ibid., 176.
14. Lewis, "Causation as Influence," *Journal of Philosophy* 97 (2000): 182–97. Reprinted in *Causation and Counterfactuals*, ed. John Collins, Ned Hall, and L. A. Paul (Cambridge, MA: MIT Press, 2004), 75–106.
15. Lewis, "Causation as Influence," *Journal of Philosophy* 97 (2000): 190; *Causation and Counterfactuals*, 91.
16. Lewis, "Causation as Influence," *Journal of Philosophy* 97 (2000): 190–97; *Causation and Counterfactuals*, 100.
17. McDermott, "Redundant Causation," 525; John Collins, "Preemptive Prevention," *Journal of Philosophy* 97 (2000): 223–24.
18. Christopher Hitchcock, "Of Humean Bondage," *British Journal for the Philosophy of Science* 54 (2003): 21.
19. Lewis, "Causation as Influence," *Journal of Philosophy* 97 (2000): 197; *Causation and Counterfactuals*, 103.
20. Lewis, *Philosophical Papers, Vol. 2*, 217.
21. Helen Beebe, "Causing and Nothingness," in *Causation and Counterfactuals*, 304–05.
22. Ibid., 307.

26

Criteria of Causal Status

A number of claims about causality have been examined thus far. Among them are the following:

1. $c \rightarrow e$ is a causal relation, if, and only if, *e*-type events regularly follow *c*-type events (Hume).
2. Cause event (state) *c* is a sufficient condition of the occurrence of effect event (state) *e* (Mill, Hempel, Popper).
3. Cause event (state) *c* is a necessary condition of the occurrence of effect event (state) *e*.
4. $c \rightarrow e$ is a causal relation, if, and only if, there is a spatiotemporally continuous transfer of energy or momentum from *c* to *e*.
5. $c \rightarrow e$ is a causal relation, if, and only if, there is a transfer of energy or momentum from *c* to *e*.
6. $c \rightarrow e$ is a causal relation, if, and only if, *c* raises the probability of *e*.
7. $c \rightarrow e$ is a causal relation, if, and only if, the occurrence of *c* is statistically relevant to the occurrence of *e* (Salmon).
8. Causal processes propagate or transmit structure from one region of space–time to another and causal interactions modify structure (Salmon).
9. Causal processes involve possession of a conserved quantity, and causal interactions involve exchange of a conserved quantity (Dowe).
10. $c \rightarrow e$ is a causal relation, if, and only if, *e* is counterfactually dependent on *c*. If *c* had not occurred then *e* would not have occurred (Lewis).
11. $c \rightarrow e$ is a causal relation, if, and only if, one can infer the occurrence of *e* from knowledge of the occurrence of *c* (Frank).
12. Causes are events (occasioning conditions) that trigger the exercise of capacities (Harré).
13. *c* is a cause of *e*, if, and only if, by manipulating *c*, we could bring about changes in *e* (Von Wright).
14. *c* is a cause of *e* if a free agent's bringing about of *c* would be an effective means of bringing about the occurrence of *e* (Menzies and Price).
15. Causal relations furnish information that is potentially relevant to manipulation and control (Woodward).

The goal of much theorizing about causality has been to formulate a criterion that picks out all and only those relations that are causal. Suppose criterion C* accomplishes this. The following two statements then would be true about events (or states) *x* and *y*:

1. "If C* is satisfied, then *x* causes *y*," i.e., satisfaction of C* is a sufficient condition of *x* causing *y*.
2. "If *x* causes *y*, then C* is satisfied," i.e., satisfaction of C* is a necessary condition of *x* causing *y*.

Have any of the theories of causality examined the above achieved success by formulating a criterion that is either necessary or sufficient for causation? In the following analysis, I follow common usage by including omissions and preventions among potential causes. I recognize that some theorists both deny that such "entities" can be causes and affirm that these "entities" do play a role in "causal explanations."

Consider the energy–momentum transfer criterion. Satisfying this criterion is not a necessary condition of causal relatedness. Instances of omission and instances of prevention may be causal despite the absence of energy transfer. One may attempt to include instances of omission and instances of prevention as cases of energy transfer by appending counterfactual claims to the analysis. But reference to what might happen in an alternative world is irrelevant to energy balances in the real world.

There is, in addition, a problem about instances of disconnection. Causation-by-disconnection differs from causation-by-omission insofar as there is an inaugurating event that transfers energy, but this transfer is not to the effect. Jonathan Schaffer has emphasized that such instances are widespread. One of his examples is the death of a victim by a gunshot wound to the heart. He observed that

> heart piercings cause death only by disconnection. The brain is kept alive by an influx of oxygenated blood, and heart piercings cause death by disconnecting this influx.[1]

Causation-by-disconnection is present as well in the following instances:

1. A circuit breaker is thrown and a room becomes dark.
2. Fred enters a frozen food storeroom and commences to shiver.
3. An apple whose stem is cut falls to the earth.
4. Evacuation of a chamber results in the expansion of an air-filled lamb's bladder enclosed therein (an experiment performed by Robert Boyle in 1658).[2]

In general, the removal of one or more conditions (or forces) that sustain an equilibrium situation causes an effect to take place by disconnection.

Schaffer observed that the ubiquity of instances of causation-by-disconnection counts against several proposed criteria of causal status.[3] He noted that to require that energy be transferred from cause to effect would be to deny causal status to instances of disconnection.

With respect to the examples cited above, one could argue, on behalf of the energy-transfer criterion, that the cutting of an apple's stem and the expansion of the lamb's bladder do transfer energy. The energy transfer in each case is from potential energy to kinetic energy. Whether a transition from potential energy to kinetic energy satisfies the criterion that a causal relation involves energy transfer is subject to debate. However, the heart-piercing, circuit-breaking, and shivering cases are not subject to a "conversion-of-potential-energy interpretation." It would seem that Schaffer is correct that some causal relations do not require that the cause transfer energy to the effect.

Schaffer maintained that instances of causation-by-disconnection also are a problem for the process theories of Salmon and Dowe. Instances of causation-by-disconnection do not involve spatiotemporally continuous persistencies. There is no line of persistence between a disconnection and its effect. Rather, a prior persistence is severed.[4] Thus, causal relatedness does not require the transmission of conserved quantities.

According to Schaffer, an adequate theory of causation will need to include absences among the necessary and/or sufficient conditions that qualify as causes. In the case of heart-piercing, for instance, the absence of oxygenated blood in the brain is properly identified as the cause of death. Absences can be incorporated among the conditions that are necessary and/or sufficient to establish causal relatedness. They cannot be incorporated into the set of "entities" (photons, energy, momentum, spin . . .) transferred in numerous physical processes.

Energy transfer may not be a sufficient condition of causation either. It depends, in part, on how one interprets radioactive decay processes. Is the very existence of Pb^{210} the cause of its decay into Hg^{206} and an α-particle? There clearly is energy transfer, and given a change in energy distribution, of course there are "before and after" differences. Should we attach the labels "cause" and "effect" to those differences? A "yes" answer supports the position that energy transfer is a sufficient condition of causation. Phil Dowe defends this position.[5]

A "no" answer reflects the belief that if any process is noncausal, it is the spontaneous decay of a nucleus. There is no prior change that is the cause of the decay. Moreover, some radioactive processes are processes that occur with extremely low probability. Suppose a C^{14} nucleus emits an electron. The half-life of this decay process is 5,570 years. Electron emission at this particular time is highly improbable, and yet the decay did take place.

A similar analysis is appropriate for the "possession or exchange of a conserved quantity" criterion. Satisfaction of this criterion is not necessary for causation. Instances of omission, prevention, and disconnection may be causal without the continuous possession or exchange of a conserved quantity.

Satisfaction of the conserved quantity criterion may not be a sufficient condition of causation either. Energy, momentum, and charge are conserved in processes of radioactive decay. However, it may be argued that, since there is no interaction in radioactive decay, the "exchange" of conserved quantities is not causal.

Probability-raising clearly is neither necessary nor sufficient for causation. It is not necessary because there exist probability-lowering causes such as the thrombosis-lowering effect of contraception usage among fertile women (see above, p. 108). Probability-raising is not sufficient for causation because there exist cases of overlapping that increase the probability that an effect occurs, while being causally irrelevant to its occurrence (see above, pp. 110–11).

Satisfying the counterfactual dependence criterion is neither necessary nor sufficient for causal relatedness. That satisfying the criterion is not a necessary condition of causal relatedness is shown by instances of overdetermination and trumping. For these cases, "x is a cause of y" is true and "$\sim x \square\!\!\rightarrow \sim y$" is false.

That satisfying the criterion is not a sufficient condition of causal relatedness is shown by the existence of reliable diagnostic relations. The following is an example from chemistry. Consider a class of instances in which a flame test is performed on a portion of a substance at time t, and a second portion of that substance is added to a solution of Na_2SO_4 at time $t + \Delta t$. Suppose that in one instance a barium-containing substance produces a distinctive green flame at t and produces a white precipitate of molecular weight 233.4 ($BaSO_4$) when added to a solution of Na_2SO_4 at $t + \Delta t$. "Let g = the flame test conducted on substance S at t, produced a green flame" and p = "the addition) of S to a solution of Na_2SO_4 at $t + \Delta t$, produced a precipitate of molecular weight 233.4." It is true that "$\sim g \,\Box\!\!\rightarrow\, \sim p$." Only barium compounds turn flames that particular shade of green when heated. There is counterfactual dependence.

It is false that "*g* is a cause of *p*," however. This is a relation unaffected by the locations of the two events and the length of the time interval between them. There is no transfer of energy or exchange of a conserved quantity between *g* and *p*. Moreover, the order of these events does not matter, as it does for relations we ordinarily term "causal."

If a dry match is struck on an abrasive surface in the presence of oxygen, then it ignites. This is a causal relation on the counterfactual dependence criterion. But it is not the case that if a green flame test is produced then $BaSO_4$ results. The relation between *g* and *p* is merely a hypothetical relation. Given a green flame test result, *if* a sample of substance *S* is added to a solution of Na_2SO_4, then $BaSO_4$ is produced. The green flame test—production of $BaSO_4$ sequence is a countercase to the thesis that counterfactual dependence is a sufficient condition of causation. Other reliable diagnostic relations provide further countercases.

Satisfying the inferability criterion also is neither necessary nor sufficient for causal relatedness. It is false that if *c* causes *e* then *e* may be inferred from *c* (in conjunction with suitable laws). Whether *c* causes *e* depends on how the world is constituted. Failure of inference does not establish absence of causal relatedness. For instance, we cannot predict the exact path of a falling leaf, but this path may be causally determined by its geometry and mass distribution, gravitational attraction, and wind currents.

Moreover, it seems obvious that collisions caused subsequent motions even before laws were formulated to permit inferences about these motions. It would be an unusual interpretation of "cause" to maintain that collision processes were noncausal before, say 1650, and causal thereafter.

In addition, there is a problem about the proper interpretation of statistical correlations of low probability. On the inferability view, high probability is a requirement for causal relatedness. Only if *B*-states follow *A*-states with high probability is the relation linking the states a causal relation. If, on the contrary, the probability is 5 percent, then the appropriate inference from the occurrence of an *A*-state is "probably not *B*." But consider the correlation between exposure to radiation and the subsequent development of leukemia. Wesley Salmon is correct to speak of a "causal relation" in such cases, even if only 1 percent of persons exposed to a certain level of radiation develop leukemia.[6]

Satisfying the inferability criterion is not a sufficient condition of causal relatedness either. It is false that if *e* may be inferred from *c* (in conjunction with suitable laws), then *c* causes *e.* There are numerous mathematical models that succeed in "saving the appearances" without providing a causal analysis. Galileo's law of odd numbers is an early example. Galileo showed that the distances covered by a falling body in successive equal-time increments conform to the relation 1, 3, 5, 7, 9, ... This kinematic description of free fall permits the scientist to infer, from information about the state of a system in one region of space and time, its state in other regions of space and time. The inferability criterion is satisfied, but no causal relatedness is established. In chemistry, the successive states of reactant–product mixtures sometimes conform to simple mathematical relations. For instance, application of third-order kinetics—$[\sim d(NO)/dt = k_3(NO)^2(O_2)]$—enables successful inferences to be made about the relative concentrations of NO, O_2, and NO_2 in the reaction that produces (NO_2)—($2NO + O_2 \rightarrow 2NO_2$). Applications of this type satisfy the inferability criterion, but fail to provide causal analyses. Other examples are epicycle-deferent models of planetary motion and various applications of the Fibonacci series (see above).

Whether satisfying a manipulability criterion is a necessary condition of causation depends on the degree of extrapolation permitted from humanly achievable manipulations. Manipulability theorists allow "possible interventions" that human beings lack the power to accomplish, provided that these interventions are "extrapolations" from manipulations that human beings can perform. Woodward goes further and permits "possible interventions" that violate fundamental physical laws. Given such "nomologically impossible interventions," the manipulability criterion is a necessary condition of causation.

Woodward's version of the manipulability criterion is a sufficient condition of causation as well. If there is a "possible intervention"—including "nomologically impossible interventions"—that would create *X* such that *Y* would follow, then *X* causes *Y*. So far as I am aware, no one has developed. a countercase to this claim. However, this defense of the "satisfaction of the manipulability criterion" as necessary and sufficient for causation depends on an extreme position on what counts as an "intervention."

The table below summarizes the limitations of the various criteria of causal relatedness:

Criteria of Causal Relatedness	**“Noncausal” relations that qualify under the Concept**	**“Causal” Relations Disqualified by the Concept**
Constant sequential conjunction	Accidental correlations	Statistical correlations, omissions, preventions, disconnections
Exchange of energy or momentum	Radioactive decay?	Omissions, preventions, disconnections
Exchange of conserved quantity	Radioactive decay?	Omissions, preventions (incorporated by reference to alternative worlds), interactions in non-closed systems
Probability increase	Increasing the number of radioactive atoms present	Birth control pills—thrombosis, squirrel’s kick holes putt
Counterfactual dependence	Diagnostic relations	Late-preemption; trumping
Inferability	Generalizations that “save appearances”	Complex processes (e.g., falling leaves)
		Single processes governed by probabilistic causal laws
Manipulability	None?	Large-scale physical processes (e.g., gravitational lensing, plate-tectonic effects)—unless wild extrapolations from human practices are permitted

Notes

1. Jonathan Schaffer, “Causation by Disconnection,” *Philosophy of Science* 67 (2000): 286.
2. See Marie Boas Hall, *Robert Boyle and Natural Philosophy* (Bloomington: Indiana University Press, 1966), 326–8.

3. Schaffer, “Causation by Disconnection,” 289–91.
4. Ibid., 299.
5. Phil Dowe, *Physical Causation* (Cambridge: Cambridge University Press, 2000), 22–6.
6. Wesley Salmon, “Why Ask ‘Why’? An Enquiry Concerning Scientific Explanation,” *Proceedings of the American Philosophical Society* 6 (1978): 689. Reprinted in *Scientific Knowledge*, ed. Janet Kourany (Belmont: Wadsworth, 1987), 56.

27

Multiple Concepts of Causality

There are various ways to respond to the failure of theories of causality to develop criteria that specify conditions that are necessary or sufficient for causal relatedness. One might adopt the position of J. J. C. Smart, for instance. Smart declared that

> philosophically I try to eschew the notion of causality. I take the view that causality is a useful notion for plumbers, heart surgeons, irrigation engineers, and so on, but I do not want it in theoretical physics or cosmology. Regularities and symmetries, yes; causality, no.[1]

Smart's "anti-causality" stance is unconvincing. Granted that theoretical physicists appear to be preoccupied with regularities and symmetries, their creations are not free-floating. These creations include links, however indirect, to observables whose magnitudes can be measured. The cosmologist, for instance, utilizes data obtained from telescopes, photometers, and spectrometers.

To utilize an instrument to make a measurement is to invoke a concept of causality. In a typical case, apparatus *A* is constructed and is subjected to manipulation *M* in anticipation of result *R*. *R* is properly spoken of as the "effect" of *M* on *A*.

Any attempt to purge science of causal concepts is misguided. However, it may be appropriate to restrict causal analysis to the description of experimental arrangements and results, while excluding reference to causal concepts at a theoretical level. For instance, theorists of the quantum domain may present results of measurement in the causal language of classical physics, but elect to exclude causal concepts from calculations made from Schrödinger's wave equation or Heisenberg's matrices.

A second response to evidence about conflicting causal criteria is to invert Smart's position by restricting "genuine causal interactions" to the domain of science. One might require of a causal relation that there be an exchange of energy or momentum between cause and effect. Instances in which an omission is responsible for an effect may be dismissed as "noncausal" since there is no energy transfer from "cause" to "effect" in such cases. Instances of causation-by-prevention and causation-by-delay may be disqualified for the same reason.

A presupposition of the above position is that there is one and only one *objective* relationship that is causal. Christopher Hitchcock took this unique claim to be part of a "Thesis of Humean Bondage" (2003), viz.,

> in any concrete situation, there is an objective fact of the matter as to whether two events are in fact bound by *the* causal relation. (Italics added)[2]

Hitchcock maintained that this thesis is false. There is no single objective type of relationship that constitutes causation. He insisted, moreover, that talk about a "glue" or a "cement" that binds cause to effect has impeded progress in the philosophical analysis of causation.[3]

If there is no single objective type of causal relationship, a plausible response is to affirm the objective existence of more than one type of causal relatedness. Support for this response is provided by the fact that our decisions about causal attribution are context-dependent. The answers we give to questions about causation depend on our interests. The appropriate answer to the question "What was the cause of Yuri's death?" is different for the forensic pathologist (e.g., a bullet severed his carotid artery leading to a fatal loss of blood) and a member of the jury (e.g., a shot fired by his wife Marta).

Perhaps there is one type of causal relation that is an objective feature of the universe and a second type of causal relation objectively present in situations that involve issues of human responsibility and guilt. One might then stipulate different criteria of causal status for each type of causal relation. Type-one causes satisfy an energy-transfer criterion and/or a statistical relevance criterion. Type-two causes satisfy a counterfactual conditional criterion.

Elliott Sober has maintained that a "two-types-of-cause" position is required within science itself, independently of considerations of responsibility and guilt.[4] On one level, a causal relation is a relation between event-tokens in which an energy transfer takes place. Sober noted that an "event-token-cause" need not raise the probability that its "event-token-effect" take place. Thus squirrel #1's kick of a golf ball may qualify as a cause (*qua* event-token) despite the fact that, in itself, the kick lowers the probability that the ball winds up in the cup. Sober maintained that event-token causality is exhibited in genealogical relationships between the individuals of biological species.

On a second level, a causal relation is a relation between properties such that the realization of a "property-cause" raises the probability of the realization of a "property-effect" under "maximally specific background conditions." Sober noted that a weakened form of this criterion may be sufficient, viz., a property-cause must raise the probability of the property-effect in at least one background context and must not lower it in any other context.[5] It is in this sense of causation that heavy cigarette smoking causes heart attacks. The "property-causal relation" holds in spite of the failure of some relevant "event-token causal relations." Heavy smoker Sal may survive to age 95 without experiencing a heart attack, and heavy smoker Sid may suffer a heart attack that is not an event-token-effect of his smoking.

Christopher Hitchcock also recommended causal pluralism. In "Farewell to Binary Causation" (1996) he distinguished binary from ternary concepts of causation:

> Binary—"*c* causes *e*"
>
> Ternary—"*c* causes *e* relative to *c**," where *c** is a specific alternative to *c*[6]

He then subdivided the binary concept into singular and general versions, illustrated by

> Singular—"Seth's smoking causes his lung cancer"
>
> General—"Heavy smoking causes lung cancer"

Hitchcock noted that our need for a "contrastive" concept of causation is evident in a scenario developed by Fred Dretske.[7]

> Susan longs to possess a bicycle and a pair of skis displayed in a local sporting goods store. She breaks into the store, realizes that she cannot pilfer both, and rides off on the bicycle, leaving the skis behind.

Different causal claims about Susan's action can be made using identical words, but with different emphases:

1. "Susan's *stealing* the bicycle caused her to be arrested" (true).
2. "Susan's stealing the *bicycle* caused her to be arrested" (false).

Hitchcock suggested that the difference between (1) and (2) be made perspicuous by emphasizing the implicit contrast situations, viz.,

1. "Susan's stealing the bicycle, rather than buying it, caused her to be arrested" (true).
2. "Susan's stealing the bicycle, rather than the skis, caused her to be arrested" (false).[8]

The contrastive concept is the basic causal relation. However, in certain contexts, it reduces to the binary concept. If $\sim c$ is the only viable alternative to c, then $c^* = \sim c$ and the relation "c causes e" is appropriate. "Closing the circuit (c) caused current to flow (e)" is an appropriate claim, because the only viable contrast situation is not closing the circuit. In other contexts, there are innumerable contrast situations that could be cited, but the differences among them are not relevant to the causal situation. Suppose Smith shoots Jones, causing his death. Hitchcock pointed out that there are numerous other things Smith might have done, instead of shooting Jones. He might have gone for a walk, eaten breakfast, written a letter, etc.

> In this case, reference to a specific alternative is unnecessary because Smith's shooting Jones caused Jones to die *relative to any one of them*.[9]

In "Of Humean Bondage," Hitchcock surveyed the types of general theories about causation—regularity theories, counterfactual dependence theories, probabilistic theories, process theories, etc. He declared that, regardless of which general theory is selected, applications of that theory reveal several distinct types of causal relations:

> once one has the basic apparatus in place—causal laws, causal correlations, non-backtracking counterfactuals, causal processes—there are a number of interesting relations that can be defined. Some of these will be important in one context, others in another. When we are asked what causes what, we may pay attention to one of these relations in one scenario, to another of these relations in another scenario.[10]

Hitchcock maintained that cause and effect may be linked in a number of ways, among them[11]

1. Cause *c* is a member of a sequence of events leading up to effect *e*. For example, a squirrel's kick of a golf ball that then ricochets off a tree into the cup qualifies as a causal event *qua* member of a chain of events that leads to *e*.
2. Cause *c* is linked to a "component effect" e_c. For example, taking birth control pills (*c*) causes an increase in the occurrence of thrombosis among infertile women (e_c).
3. Cause *c* is linked to a "net effect" e_n. For example, taking birth control pills (*c*) causes a reduction in the occurrence of thrombosis among woman of child-bearing age (e_n). (The net effect takes account of both the increased thrombosis rate among infertile women and the decreased thrombosis rate among fertile women.)

By distinguishing "component effect" and "net effect" causal relations, Hitchcock is able to avoid the *prima facie* contradiction that taking birth control pills both decreases and increases the incidence of thrombosis. Of course, the contradiction also may be avoided by relativizing the causal claims to the appropriate classes.

4. Cause *c* is linked to effect *e* as a matter of physical law. For example, an ice cube placed in water (*c*) floats (*e*), and the children of blue-eyed parents (*c*) have blue eyes (*e*).
5. Cause *c* is linked to effect *e* relative to some particular range of alternatives or domain of variation. For example, a sudden drop in external temperature (*c*) caused Sam to shiver as he emerged from his house (*e*). Involuntary muscular contraction is one mechanism by which the body maintains internal temperature homeostasis in response to a change in external temperature. (Other mechanisms are adrenal gland activity, thyroid gland activity, and the dilation or contraction of peripheral blood vessels.)

Hitchcock recommended that we accept multiple distinct causal concepts. The search for a univocal concept applicable whenever there is causal relatedness has not been successful.

One way to implement Hitchcock's recommendation is to adopt a disjunctive criterion of causal relatedness. Taken individually, causal criteria C* and C# may be too restrictive. The range of relations that qualify as causal may be increased by formulating a disjunctive conjunctive criterion of the form

If either C* or C# is satisfied, then *x* is a cause of *y*.

A promising candidate is

> *x* is a cause of *y* provided that either there is energy transfer from *x* to *y*, or *y* is counterfactually dependent on *x*.

Satisfaction of just one criterion of the disjunction qualifies the relation between *x* and *y* as causal.

There are many noncontroversial examples of causal relatedness that involve energy transfer but not counterfactual dependence. Ned Hall advanced the following version of the "rock-hits-window" example:

> Billy and Suzy both throw perfectly aimed rocks at a window. Suzy's gets there first, breaking it. Intuition unhesitatingly picks out her throw, and not Billy's, as a cause of the breaking.[12]

Hall maintained that it is the intersection of a structure-altering process inaugurated by Suzy that is causal, despite the fact that her throw fails to exhibit counterfactual dependence. (Billy's throw would have broken the window in the absence of Suzy's throw.) In this example, our intuitions correctly single out Suzy's throw as cause of the breaking.

There also are noncontroversial examples of causal relatedness that involve counterfactual dependence but not energy transfer. Consider the case of Gardener Gus. Gus failed to water a rose bush entrusted to his care and the bush died. If, contrary to fact, Gus had watered the bush, it would not have died. Our intuitions are clear in this case as well. Gus's neglect is the cause of its death, despite the fact that there is no relevant energy transfer.

So far, so good. We accept both Suzy's action and Gus's neglect as causes. Hall proceeded to develop a more challenging scenario—the "Case of the Drunken Bad Guys":

> the bad guys want to stop the good guys' train from getting to its destination. Knowing that the switch is currently set to send the train down the left-hand track, the bad guy leader sends a demolition team blow up a section of that track. En route to its mission, the team stops in a pub. One pint leads to another, and hours later the team members have all passed out. The good guys, completely oblivious to what has been going on, send the train down the left-hand track. It arrives at its destination.[13]

Let *c* be the bad guy leader's command to the demolition team, *d* the team members' drinking at the pub, and *e* the safe arrival of the train at its destination.

Hall maintained that our intuitions are correct that *c* is not a cause of *e*. Is *d* a case of *e*? Presumably there is counterfactual dependence. If the demolition team had not become sidetracked in the pub, then they would have blown up the appropriate section of track, and the train would not have reached its destination. However, there is no energy transfer that links *d* and *e*. No structure-modifying processes link pub-drinking and the safe passage of the train to its destination.

If we assign causal status to the bad guys' drinking, because the relation of counterfactual dependence is satisfied, our intuitions about causation are offended. Hall maintained that

> it makes sense, for example, to ask whether, by entering the pub, the demolition team really helps cause the train's arrival. Why such a question makes sense is perfectly clear: by stressing the word "really" . . . one invokes a context where the central kind of causation [production] is salient; and entering the pub is not a cause of the arrival, in this sense.[14]

Hall acknowledged that *d* is a cause of *e* on the counterfactual dependence criterion, but maintained that this attribution of causal status should be overruled by appeal to the failure of the transition $d \rightarrow e$ to exhibit energy transfer.

Hall's analysis indicates a serious problem for our proposed disjunctive criterion. It would seem that sometimes satisfying the counterfactual dependence disjunct establishes causal relatedness, and sometimes it does not. It all depends on our beliefs about the case at hand. But if our prior beliefs about causation can override applications of our disjunctive criterion, then the criterion itself is compromised.

One might agree with Hall that production is more important than counterfactual dependence, but nevertheless stipulate as a methodological principle that satisfying either disjunct establishes causal relatedness. This would be to choose always the disjunctive criterion over our pre-analytic beliefs in controversial cases. We might then conclude that the drinking bout in the tavern did cause the safe arrival of the train.

However, affirmation of this methodological principle would have undesirable consequences. Since radioactive decay processes involve energy transfer, they then would qualify as "causal." And since the barium-flame-test→ precipitation of barium sulfate sequence exhibits counterfactual dependence, it also would qualify as "causal."

Applications of a disjunctive criterion increase the range of relations that qualify as causal only at the expense of qualifying as "causal" relations we believe to be noncausal.

Notes

1. J. J. C. Smart, "Laws of Nature as a Species of Regularities," in *Ontology, Causality and Mind: Essays in Honour of D. M. Armstrong*, ed. J. Bacon, K. Campbell, and L. Reinhardt (Cambridge: Cambridge University Press, 1993), 164.
2. Christopher Hitchcock, "Of Humean Bondage," *British Journal for the Philosophy of Science* 54 (2003): 10.
3. Ibid., 3.
4. Elliott Sober, "Two Concepts of Cause," *PSA 1984, Vol. II* (1985): 405–24.
5. Ibid., 421.
6. Hitchcock, "Farewell to Binary Causation," *Canadian Journal of Philosophy* 26 (1996): 268.
7. Fred Dretske, "Referring to Events," *Midwest Studies in Philosophy* 2 (1977): 90–99.
8. Hitchcock, "Farewell to Binary Causation," 276.
9. Ibid., 281–82.
10. Hitchcock, "Of Humean Bondage," 8.
11. Ibid., 21–22.
12. Ned Hall, "Causation and the Price of Transitivity," *Journal of Philosophy* (2000): 219.
13. Ibid., 210.
14. Ibid., 219.

28

Conflicting Criteria of Causal Status

The principal reason to adopt a disjunctive criterion of causal relatedness is to increase the range of relations that qualify as causal. If this could be accomplished without including *prima facie* noncausal relations, all would be well. Unfortunately, however, when disjoined criteria are applied separately, conflicts arise. The following examples illustrate the conflicts that arise from applications of different criteria of causal status.

1. *Overdetermination cases*. Smith fires simultaneously the gun he holds in his right hand and the gun he holds in his left hand. Both bullets hit a plate of glass, which shatters.

The question arises "Does the bullet fired from the left-hand gun cause the shattering?" It does according to the regularity criterion, the sufficient-condition criterion, and the energy-transfer criterion. But it does not according to the necessary condition criterion and the counterfactual dependence criterion (unless the effect is redescribed as "a shattering from a bullet from Smith's left-hand gun").

2. *Instances of causation-by-omission*. Gardener Gus neglects to water one of the rose bushes under his care in the greenhouse. The rose dies. The counterfactual dependence criterion is satisfied. If Gus had watered the rose then it would have lived. However, this alleged causal relation violates the energy-transfer criteria and the conserved quantity criterion.

In cases of causation-by-omission, there often is a conflict between the results of a "scientific" causal analysis and the results of causal analysis that invokes questions of responsibility or guilt. Consider the following example. Fraternity lad Leon steals a stop sign for his dorm room. Shortly thereafter, a car proceeds through the signless intersection and strikes another car. Leon's action presumably is relevant to the accident. Did his action "cause" the collision? A prosecuting attorney might seek to establish that claim in court. However, the energy-transfer analysis is that the "cause" of the accident was the intersection of the paths of the two cars, an intersection marked by abrupt changes of momentum, and kinetic energy. This may be the case, but in a legal context, this causal attribution may be considered irrelevant.

Our preference for an "attribution of guilt" analysis depends on our estimate of the degree of responsibility shirked. Leon may argue effectively that he should not be expected to have anticipated that an accident would result following his theft of the sign. But a mother who fails to feed her baby will be held to have caused its death, absent mitigating circumstances. A defense attorney who argues that the "real cause" of death was a change of the baby's metabolic variables would be dismissed as irrelevant. In criminal cases, an autopsy report does not exhaust questions about the cause of death.

It is our estimate of degree of responsibility that leads us to condemn greenhouse gardener Gus for the demise of the rose bush, but not to condemn the Queen of England.[1] Gus neglected to water the bush for several weeks, and the plant died. But the Queen of England also failed to water the plant during the time in question. We do not conclude that her inactivity is the cause of the plant's death. Gus was responsible for the care of the rose bush, the Queen was not.

3. *Instances of prevention*. Al hurls a lighted torch through the open window of a warehouse used to store recycled paper. At that moment, Ben pushes a metal cart beneath the window. The torch lands in the cart and no fire ensues. Did Ben's action cause the nonoccurrence of a fire? The action satisfies the sufficient-condition criterion and the counterfactual dependence requirement (if Ben had not passed the cart beneath the window then there would have been a fire in the warehouse) and the manipulability criterion. However, Ben's action is not a necessary condition of the absence of fire, and there seems to be no direct path of energy transfer to an effect. Indeed, our intuitions are troubled by reference to the non-fire as an "effect."

And yet, it seems to be reasonable to maintain that by dissipating impact energy, the controlled collapse of a car's crumple-zone caused its passenger compartment to remain intact. Again, the counterfactual condition criterion is satisfied, as is the manipulability criterion and the exercise of capacities criterion. However, energy transfer is from impact to the collapsing front end. It is the *absence* of energy transfer to the passenger compartment that permits it to remain intact.

Christopher Hitchcock has presented an interesting case of causation-by-prevention:

> Two assassins, Captain and Assistant, are on a mission to kill Victim. Upon spotting Victim, Captain yells "fire!," and Assistant fires. Overhearing the order, Victim ducks and survives unscathed.[2]

Hitchcock noted that we are uncertain whether to affirm that Captain's yelling "fire!"caused the survival of Victim. This is the case even if we restrict attention to applications of the counterfactual dependence criterion.

On the one hand, it presumably is true that if Captain had not yelled "fire!" then Victim would not have ducked, and that, if Victim had not ducked, then Victim would not have survived. If the causal relation is transitive, then Captain's yelling "fire!" caused Victim to survive. But if Captain had not yelled "fire!" Assistant would not have fired, and Victim's health would not have been at risk. Captain's yelling "fire!" was not the cause of Victim's survival.

Consider the case of preemptive prevention discussed above (pp. 164–65). A thrown ball is speeding toward a plate glass window.

1. Fielder *Y* is positioned on a line between Fielder *X* and the window. *X* catches the ball, but had he not done so, *Y* would have caught it.
2. There is a brick wall between Fielder *X* and the window. *X* catches the ball, but had he not done so, the ball would have struck the wall.[3]
3. The window is on the moon at the horizon. *X* catches the ball, but had he not done so, the ball would have failed to reach the window.[4]

In which case, or cases, does *X*'s catch prevent the window from shattering (i.e., cause the window to remain intact)? If we apply the energy-transfer criterion, there is nothing to choose among the cases. *X*'s catch terminates the ball's motion in each case. By contrast, if we apply the counterfactual conditional criterion, there are differences among the cases. Assigning causal credit to *X* is completely far-fetched if the window is on the moon, quite inappropriate if the window is behind a wall, and debatable if Fielder *Y* is the backup in case Fielder *X* fails to catch the ball. (Fielder *Y* may be nearsighted or have poor eye–hand coordination.)[5]

4. *Causation-by-disconnection*. If an effect is produced by removing a condition that sustains an equilibrium situation, then there is causation-by-disconnection. As noted above, Schaffer maintained that

heart piercings cause death by disconnecting the influx of oxygenated blood to the brain. Such cases of causation-by-disconnection satisfy the inferability criterion, the counterfactual dependence criterion, and the sufficient-condition criterion, but violate the energy-transfer criterion and the exchange of a conserved quantity criterion.

5. *Delaying causes.* Sometimes causes just delay, but do not prevent, the occurrence of effects. For instance, a recusal by Judge Judy caused a delay in the start of Ken's trial. Judy's departure from the case is neither a sufficient condition nor a necessary condition of the delay. Nor is there a causally relevant energy transfer. However, one may argue that her action is causal on a counterfactual dependence criterion. If she had not recused herself from the bench then there would have been no delay in the start of the trial. To defend this claim one would have to argue that no other delaying causes—e.g., a change of venue motion by defense counsel, Ken's heart attack—would have occurred.
6. *Probability-lowering causes.* Duffer strikes a putt. Squirrel #1 charges into the path of the ball and kicks it so that it moves at right angle to its former path, whereupon Squirrel #2 charges into its path and kicks the ball so that it rolls into the cup. Does Squirrel #1's kick cause the sinking of the putt? It does on the energy-transfer criteria, since the kick is part of a sequence of impacts that directed the ball into the hole. But Squirrel #1' s kick is disqualified from causal status by the regularity criterion, the necessary condition criterion, the probability-raising criterion, and the statistical relevance criterion, since Squirrel #1's kick lowers the probability that the ball reaches the hole.[6]
7. *Attempts to apply causal concepts to phenomena in the quantum domain.* Consider the two-slit experiment. According to Richard Feynman, this experiment epitomizes the difficulties that confront interpreters of the subatomic domain.[7] It has been performed on both photons and electrons (Figure 28.1[8]).

If slit *A* is open and slit *B* is closed, repeated passage of individual electrons to the screen produces distribution pattern *A**. If slit *A* is closed and slit *B* is open, the result is distribution pattern *B**. However, if both slits are open, repeated passage of electrons produce interference pattern *C**, and not a juxtaposition of patterns *A** and *B**.

Is it appropriate to say that the interaction of electrons and the two open slits of the diaphragm *causes* interference pattern *C*?* And if so, which criteria of causal status are satisfied in this experiment?

Insofar as there is an exchange of a conserved quantity (energy) from each electron to the screen, the conserved quantity criterion is satisfied. In addition, the experimental results conform to the inferability

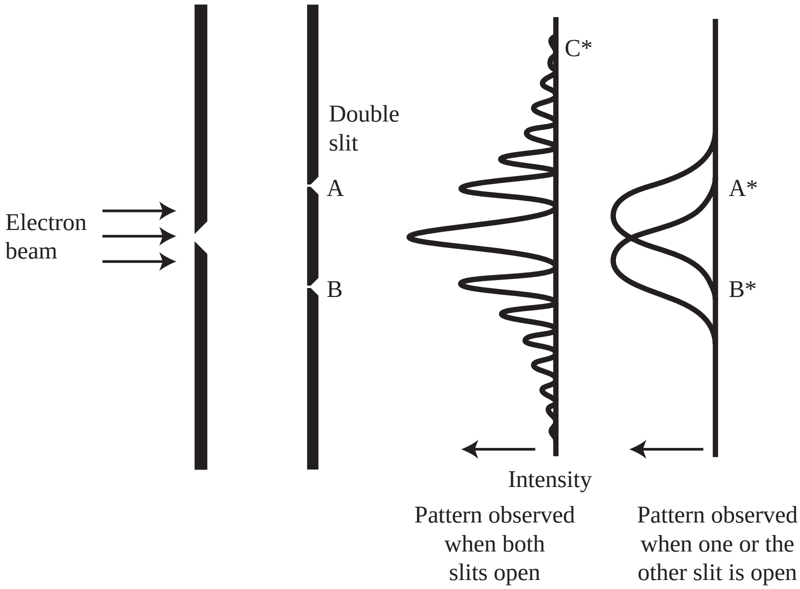


Figure 28.1 The two-slit experiment.

criterion. The interference pattern (*C**) produced when both slits are open may be calculated from the principles of quantum theory. And the probability that a single electron strikes the screen at a specific location also may be calculated from the theory.

The difficulty posed by this experiment for a theory of causation is our inability to trace a continuous spatiotemporal path from source to screen. Each individual electron passes through the diaphragm and strikes the screen such that the aggregate of hits displays pattern *C**. Pattern *C** is a typical wave-interference pattern that would be expected if two waves emanated simultaneously from slits *A* and *B*. We cannot—without destroying *C**—determine which slit an electron passed through. Individual electrons fill in appropriate positions in the interference pattern, guided by the probabilities specified by the Schrödinger wave equation. A given electron apparently "knows" where its comrades have landed and selects a "right" location to contribute to the overall interference pattern. John Gribbin analyzed the corresponding situation involving photons as follows:

> although each photon starts out as a particle, and arrives as a particle, it seems to have gone through both holes at once, interfered with

> itself, and worked out just where to place itself on the film to make its own minute contribution to the overall interference pattern.[9]

The results of the two-slit experiment are inconsistent with the requirement of spatiotemporal continuity. Gribbin pointed out that the "nonlocality" problem associated with the Einstein–Podolsky–Rosen thought experiment and the Bell inequality (experimentally confirmed by Alain Aspect et al.) also is present in the two-slit experiment.

Suppose we set up detectors at slits *A* and *B* and send electrons one by one through the apparatus. We observe what happens at the slits and the screen. Each electron is observed to pass through one or the other of the slits. But with the accumulation of electron hits on the screen, the pattern displayed is not interference pattern *C**, but rather the juxtaposition of patterns *A** and *B**. It would appear that observation of the events taking place at the slits destroys the interference pattern. The electrons no longer act as waves. This is curious. But even more curious is the effect of observing just one of the slits, say slit *A*. We detect electron particles passing through *A*. Electrons passing through *B* "seem to know" what has occurred at *A* and also behave as particles, thereby contributing to the juxtaposition pattern. Gribbin concluded that

> somehow, the electrons going through the second hole "know" that we are looking at the other hole, and also behave like particles as a result.[10]

This is a manifestation of nonlocality. Events that take place at slit *A* somehow influence what takes place at slit *B*.

The Einstein–Podolsky–Rosen experiment provides another example of nonlocality. In David Bohm's version of the experiment, a pair of electrons are emitted, for instance, by the decay of a neutral pion.[11] The spin of the pion is zero. According to the law of conservation of spin, the spin of the emitted electrons must total zero. A Stern–Gerlach magnet is located so as to measure the spin of electron #1. Suppose the result of the measurement is "spin up." One may predict from this measurement that the spin of distant electron #2 is "spin down." David Lindley's assessment of this situation is that

> it seems that a spin measurement on one electron has in effect "measured" the spin of the second, remote electron, without any actual physical intervention. The second electron has been reduced

> from a wholly indefinite unmeasured state to a definite "down" state, even though nothing has been directly done to it.[12]

Addressing the question "What is weird about this quantum connection?" Tim Maudlin stressed three features[13]:

1. The quantum connection is unattenuated. It is unaffected by distance. It does not matter how far apart the two particles are, the results of measurement are the same. This behavior is in sharp contrast to the behavior of electromagnetic and gravitational fields.
2. The quantum connection is discriminating. It is a private arrangement between the two particles that has no impact on the rest of the universe. Again, this behavior is in sharp contrast to the behavior of classical fields.
3. The quantum connection is faster than light. Maudlin noted that the speed of the quantum communication appears to be incompatible with relativistic space–time structure.[14]

Does the determination that the spin direction of electron #1 is "up" cause the measured spin of electron #2 to be "down"? On the energy-transfer theory this is not a causal relation. There is no energy transfer between the electrons.

However, the successive spin determination results on electron #1 and electron #2 do satisfy the exchange of a conserved quantity criterion. Spin is conserved, and, on the conserved quantities account, this is a causal relation. Moreover, there is counterfactual dependence. On the counterfactual dependence theory, the measurement of "up" for electron #1 caused the measured value of "down" for electron #2 if, and only if the measurement of electron #1 had been "not-up," then the measured value of electron #2 would have been "not-down." In a world that closely resembles our own, the spin direction of electron #1 is measured and found to be "down." Subsequently, the spin direction of electron #2 is measured and found to be "up." On the counterfactual dependence theory, the determination of "spin up" for electron #1 does cause the measured value of electron #2 to be "down." The pion decay sequence also qualifies as causal on the inferability theory. We can predict, from information about the system in one region of space and time, information about the system in other regions of space and time. And these predictions have been confirmed experimentally.

Ned Hall stated that production is the "central kind of causation"[15] and that production is given preference in cases for which our intuitions are uncertain. However, in the electron-spin-determination example we may not be willing to select energy transfer as the more important determinant for causal status.

The Issue of Spatiotemporal Continuity

The assessment of causal status for quantum phenomena is complicated by a seeming lack of spatiotemporal continuity between putative cause events and effect events. Does causal relatedness require spatiotemporal continuity between cause and effect?

J. R. Lucas maintained that it does. According to Lucas, there are five characteristic features of a causal relation. A causal relation is necessary, repeatable, explicable, spatiotemporally continuous, and the cause is antecedent in time to the effect.[16] To require necessity is to repudiate the regularity interpretation. Something more than constant sequential conjunction is required. To require the antecedent status of the cause and repeatability is to recognize the importance of our practical concern with causes as means of manipulating things.[17] To require explication is to recognize that the uncovering of causal relations is an essential aspect of our search for theoretical understanding. And finally we require that causal relations display spatiotemporal continuity because

> we have a view of space and time which rules out action at a distance as being essentially inexplicable.[18]

Lucas noted that we sometimes accept a relation as causal despite our inability to trace a continuous path from cause to effect. He insisted that in such cases our demand for understanding is thwarted until such time as a plausible path can be specified. One way to specify a "plausible path" is to invoke the operation of fields. However, this approach raises the well-known difficulties of wave-particle dualism. Fields are superimposed one upon another in space and time, particles are not. Invocation of fields may establish spatiotemporal continuity at the expense of creating an inability to establish material identity. Electromagnetic fields and gravitational fields from diverse sources merge in such a way that it is not possible to trace their continuous identities. According to Lucas, we require of a causal relation both spatiotemporal continuity and the ability to establish material identity.

The requirement of spatiotemporal continuity bars "causal-interaction-at-a-distance." It also stipulates that the spatial order and the temporal order are continuous.

> Between any two instants of time or point of space there is another. Any division of instants into those that are all earlier than the later ones and those that are all later than the earlier ones, itself defines an instant; there is no partition of space into two disjoint closed subsets.[19]

One may or may not accept Lucas's position that only relations that are spatiotemporally continuous in the above sense qualify as causal relations. If one does accept this continuity requirement then our observations at slit *A* in the two-slit experiment do *not* cause electrons at slit *B* to contribute to pattern *B** rather than interference pattern *C**. Although the two-slit experiment qualifies as a causal relation on the inferability criterion and the counterfactual dependence criterion, it is a noncausal relation on the spatiotemporal continuity criterion.

The requirement of spatiotemporal continuity for causal relatedness is problematic within classical physics as well. There are numerous instances of *prima facie* causal relatedness without spatiotemporal continuity. Among them are instances in which fields play a causal role. Oersted's experiment, cited above (p. 25) is a good example.

Oersted placed a magnetic needle, free to pivot in a horizontal plane, above a straight stretch of wire in an electric circuit. When the current was turned on, the needle took on a position perpendicular to the wire. He reversed the direction of current flow through the wire and observed that the needle moved to a position 180° from its initial position. When asked for the cause of the needle's behavior, the contemporary physicist will cite either the current (flow of electrons) in the wire or the magnetic field around the wire.

There is a law that correlates the magnetic field intensity around the current-carrying wire with radial distance from the wire. The variation of magnetic field intensity with distance is continuous. However, there is no chain of events—continuous in space and time—from the wire to the magnetized needle. There is only a field of *potential* effects at various intermediate distances. Consider a point at a distance x from the wire. If, contrary to fact, a unit magnetic pole had been present at x, then it would have experienced a specific force in a specific direction. There is a string of such potentials of decreasing magnitude from wire to needle. Does this mean that there is a spatiotemporally continuous process in which wire and needle participate? If not, then the requirement that cause and effect be linked by a spatiotemporally continuous process should be rejected. The alternative would be to declare scientists' usage to be mistaken when they attribute causation to the presence of electric and magnetic fields.

The effects of electromagnetic fields qualify as causal on a manipulability criterion. So also do the effects realized in the two-slit experiment. Observations of electrons at slit *A* is an intervention that alters the behavior of electrons passing unobserved through slit *B*. The pattern on the screen changes from an interference pattern to a juxtaposition pattern as a result of the intervention. On a manipulability criterion this is a causal relation despite the absence of a spatiotemporally continuous transfer of energy. Woodward rejected Lucas's continuity requirement. He maintained that a relation may be causal without the existence of a continuous connection through space and time.[20]

Unfortunately, Woodward's position does not reduce the strangeness we find in the interpretation of quantum phenomena. He seeks to extend a common sense concept of cause, rooted in the success of our everyday manipulation of objects to cases that require an instantaneous, discontinuous transfer of energy (or information). The price to be paid for thus extending causal analysis to quantum phenomena is high, however. The instantaneous transfer of energy (or information) over a distance violates a basic principle of the theory of relativity.

Notes

1. Jonathan Schaffer, "Causation by Disconnection," *Philosophy of Science* 67 (2000): 295.
2. Christopher Hitchcock, "Of Humean Bondage," *British Journal for the Philosophy of Science* 54 (2003): 10.
3. Michael McDermott, "Redundant Causation," *British Journal for the Philosophy of Science* 46 (1995): 525.
4. Hitchcock, "Of Humean Bondage," 21.
5. John Collins, "Preemptive Prevention," *Journal of Philosophy* 97 (2000): 229.
6. Cited by Elliott Sober, "Two Concepts of Cause," *PSA 1984, Vol. II* (1985): 406–07.
7. Feynman, Richard P., Robert B. Leighton, and Matthew Sands, *The Feynman Lectures on Physics, Vol. 3* (Reading, MA: Addison – Wesley, 1965), 1.

8. John Gribbin, *Schrödinger's Kittens and the Search for Reality* (Boston: Little, Brown and Co., 1995), 6.
9. Ibid., 5.
10. Ibid., 13.
11. David Lindley, *Where Does the Weirdness Go?* (New York: Basic Books, 1996), 96–100.
12. Ibid., 97.
13. Tim Maudlin, *Quantum Non-Locality and Relativity* (Oxford: Blackwell, 1994), 22–24.
14. Ibid., 24.
15. Ned Hall, "Causation and the Price of Transitivity," 219.
16. J. R. Lucas, *Space, Time and Causality* (Oxford: Clarendon Press, 1984), 33.
17. Ibid., 36–37.
18. Ibid., 42.
19. Ibid., 184.
20. James Woodward, *Making Things Happen* (Oxford: Oxford University Press, 2003), 147–49.

Conclusion

The program to provide a reductive analysis of causality has not achieved success. The leading theories identify causation with (1) constant sequential conjunction, (2) probability increase, (3) energy transfer, (4) exchange of a conserved quantity, (5) counterfactual dependence, (6) inferability, and (7) manipulability. Each of these identifications is either too inclusive or too exclusive, or both.

I believe that Ned Hall was correct to maintain that (1) it would be salutary if a single concept of causation could be formulated that fits all, and only those, instances deemed "causal" on pre-analytic grounds; (2) this has not been accomplished; and (3) prospects for a satisfactory univocal concept of causation are nil.[1] Moreover, the several candidates for a universal concept of causality conflict when applied to cases of overdetermination, omission, prevention, or disconnection.

Nevertheless, the causal theories cited above account for many instances judged causal on pre-analytic grounds. One may superimpose upon events the energy-transfer theory, the counterfactual dependence theory, the inferability theory, or the manipulability theory with expectation of a good fit in many cases. For instance, the shattering of a glass plate upon impact with a rock qualifies as causal on every one of the theories listed above. Moreover, in many cases for which applications of criteria of causal status conflict, we believe that we understand the processes under consideration despite the conflict. This is true for instances of omission, prevention, and trumping. No mystery is present when Gardener Gus fails to water a rose bush, when Carl throws a circuit breaker, or when major and sergeant simultaneously incite the troops to charge.

Brian Skyrms maintained that the collection of causal concepts discussed above may be embraced in

> an *amiable* confusion, one with real heuristic value, because in the noisy macroworld of everyday life they often go together (especially

> when their application is guided by principles of locality and temporal priority).[2]

This ideal of harmony is not achieved in the quantum realm, however. We cannot say that we understand the two-slit experiment or neutral-pion decay, even though applications of criteria of causal status conflict. The "amiable confusion" that nevertheless has "heuristic value" in the "noisy macroworld of everyday life" is not present in the interpretation of quantum phenomena. Skyrms declared that

> when we ask whether causes operate locally in the quantum domain, the old cluster concept loses its heuristic value and becomes positively misleading.[3]

It would seem that we can apply indifferently any one of a set of criteria of causal status within the "noisy macroworld" and achieve success in most cases. "Most cases" do not include instances of overdetermination, omission, prevention, or disconnection, however. In these instances, there is conflict among applications of criteria of causal status.

In the realm of quantum phenomena, there is no widespread range of agreement about causal status. Indeed, it may be argued that our beliefs about causation are systematically undermined by the results of experiments in the quantum domain.

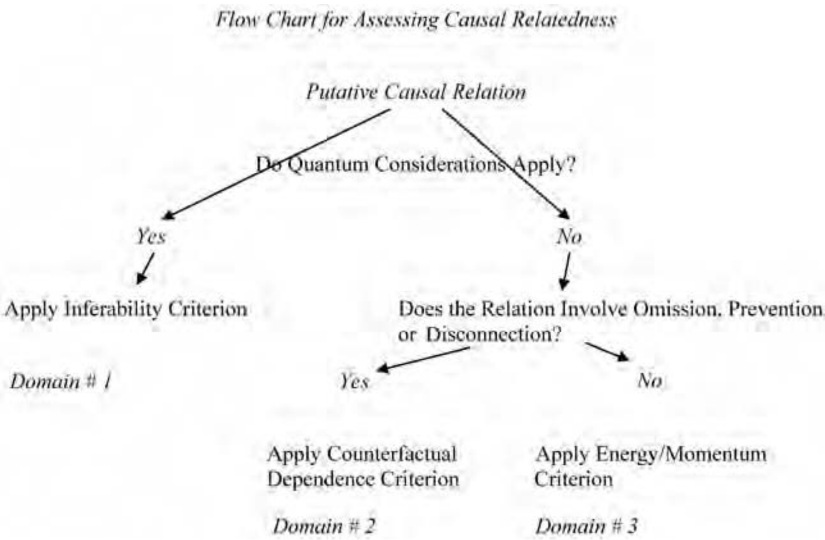
There may be a way to mitigate the skepticism engendered by the above picture, however. Suppose we were to mark out a number of nonoverlapping domains of causal discourse and accept a single theory of causal relatedness for each domain. For example, one such domain includes assertions about quantum phenomena, a second domain includes assertions made in the "noisy macroworld" that involve considerations of omission, prevention, or disconnection, and a third domain includes assertions about this "noisy macroworld" that do not involve considerations of omission, prevention, or disconnection.

It is the third domain that includes paradigmatic examples of causal relatedness. Our intuitions about causation are most secure when we contemplate collisions or other force–motion relations. Tim Maudlin pointed out that the laws of Newtonian mechanics are particularly well suited to the project of identifying causes. He noted that

> there are, on the one hand, *inertial* laws that describe how some entities behave when nothing acts on them, and then there are laws of *deviation* that specify in what conditions, and in what ways, the behavior will deviate from the inertial behavior.[4]

To the extent that we can interpret a process to be governed by "quasi-Newtonian" lawlike generalizations that specify inertial expectations and factors that produce change, we are confident about our assessments of causal relatedness.

Inferability, counterfactual dependence, and energy–momentum transfer are strong candidates for criteria for the three domains. If each domain is assigned the appropriate criterion, a flowchart for assessing causal relatedness may be constructed:



There are other ways to carve out causal domains and other choices for criteria of causal relatedness for the domains. But if the above flowchart is selected, the following sequences discussed earlier in this volume qualify as causal:

Sequence	Domain
Passage of subatomic particles through slits in a diaphragm	1
Impact of David's stone on Goliath	2
Joint command to charge by major and sergeant	2
Failure of Gardener Gus to water a rose bush	2
Piercing of Adams's heart	2
Flaming torch falls into Ben's cart, no fire occurs in paper warehouse	2
Car's front end collapses, occupants uninjured	2
Circuit breaker thrown, room becomes dark	2
Sam shivers on entering a freezer	2
α-Particle absorption	3
Squirrel's kick holes putt	3
Harry raises hand, his bid on an armoire is registered	3
Heavy smoker Sid develops lung cancer	3

By contrast, applications of the flowchart establish that the following sequences are noncausal:

> Factory whistle blows, workers leave building
>
> Bad guys get drunk, train arrives safely
>
> Sequential axial distribution of leaves emerge on the stem of a monocotyledonous plant
>
> Green flame test, $BaSO_4$ precipitation

Adopting the above flowchart has the advantage of reducing the possibilities for conflict between applications of criteria of causal relatedness. Within a given domain, decisions about causal status are made by appeal to a single criterion. Conflict is restricted to any region in which domains overlap.

There are problems with the flowchart, however. It might seem that statistical claims such as "heavy cigarette smoking causes lung cancer" and "taking birth control pills by fertile women causes a reduction in the incidence of thrombosis in this population" fail to qualify as causal.

In such cases, it is important to distinguish that which constitutes the causal relation (energy transfer) from the grounds on which the causal claim is made (predominantly statistical correlations). Heavy smoker Sid's lung cancer presumably was caused by processes of energy transfer within his body. The principal warrant for this causal relation is statistical evidence on a large number of smokers. The flowchart does accommodate judgments about causal status that are based on statistical considerations.

There are, however, other problems for the flowchart. In the first place, there is an awkward transition from causation-as-ontological-relation to causation-as-an-epistemological-claim. In the "noisy macroworld," a causal relation normally is a transfer of energy or momentum from cause to effect. In the quantum realm causation involves only an epistemological claim. If there is a causal relation in this realm, then it is possible to infer, from knowledge about a "cause situation," information about an "effect situation."

In the second place, the inferability criterion itself is suspect. The question always may be raised whether a successful inference about a quantum process is a mere saving of appearances without causal standing. After all, there are numerous successful inferences in the "noisy macroworld" that save appearances without establishing causal relatedness.

And in the third place, there remains the problematic interpretation of radioactive decay processes. These quantum processes qualify as causal on the inferability criterion, and this is counterintuitive. I leave it to the reader to decide whether the advantages of adopting a flowchart of the type illustrated above outweigh the disadvantages.

Postscript: A Descriptive Alternative

The search for a normative–prescriptive theory of causal relatedness has not achieved success. Various proposals have been advanced, among them that causal relatedness is constant sequential conjunction, counterfactual dependence, energy transfer, or exchange of a conserved quantity. None of these proposals is entirely satisfactory.

There is another approach to a theory of causal relatedness. Just as one may practice descriptive philosophy of science without issuing normative–prescriptive pronouncements, one also may elect to describe and organize causal attributions that are sanctioned within science and everyday life without making claims about conditions necessary and/or sufficient for causal status.

Ludwig Wittgenstein insisted that the meaning of a word is (usually) its *use* in the language.[5] From a Wittgensteinian perspective, to ask for *the* meaning of "cause" is like asking for *the* use of rope. The philosopher's proper task is to trace the diverse uses of "cause" within our language and to stipulate that these uses, and others that resemble them, qualify as "causes." A preliminary survey of typical causal attributions within the "noisy macroworld of everyday life" is given below:

A taxonomy of causal attributions

Type of entity	Examples
Event-tokens	"Striking the dry match on that abrasive surface caused it to ignite."
	"An electric discharge caused the thunder."
	"A meteoric impact caused that crater."
Event types	"Striking dry matches on abrasive surfaces causes them to ignite."
	"Heavy cigarette smoking causes lung cancer."
Conditions	
Necessary	"Adding hydrogen to that sample of oxygen caused the formation of water."
	"Bites from yellow-fever-fed mosquitoes cause human subjects to contact yellow fever."
Sufficient	"The fall of the guillotine blade caused the death of Marie Antoinette."
	"The addition of salt (an electrolyte) caused a lowering of the melting point of water in that bucket."
Properties of objects	"The crystalline structure of ice causes it to float in water."
	"The bonding angle between hydrogen and oxygen atoms in water causes salts to dissolve in water."

(*Continued*)

Type of entity	Examples
Changes of states of physical systems	"Heating the balloon caused its expansion."
	"Increasing the voltage within that circuit causes an increase in current flow."
States of physical system processes	"Melting ice causes a release of 80 calories per gram."
	"An increase of the temperature of the tea kettle causes the enclosed water to boil."
	"Boiling water in a covered pan causes condensation on the inner surface of its cover."
	"The buildup of charge on the plates of an electrostatic generator caused a spark discharge."
Changes of values of conserved quantities	"Rapid rotation of the submerged paddle wheels caused an increase of water temperature." (Rumford)
Fields	"Switching on the electromagnet caused a deflection of the electron beam."
	"The electric field caused a change in the path of that charged particle."
	"The magnetic field around the current-carrying wire caused a deflection of the magnetized needle." (Oersted)
Absences	
Omissions	"Susan's failure to feed the parking meter caused her to receive a ticket."
Preventions	"Safety Sam's hit caused the pass to be incomplete."
Facts	"A 23.5° inclination of the earth's axis of rotation to the plane of its revolution around the sun causes the progression of the seasons in Chicago."
	"Traders' awareness of July's employment data caused an increase of the Standard & Poor's 500 Index."

This is an extensive taxonomy of causal attributions, even without venturing into controversial linguistic usage in quantum mechanics. A fully developed descriptive theory of causal relatedness would fill in the taxa listed above, introduce new taxa as required, and extend the compilation to the quantum realm.

Notes

1. Ned Hall, "Two Concepts of Causation," in *Causation and Counterfactuals*, ed. John Collins, Ned Hall, and L. A. Paul (Cambridge: MIT Press, 2004), 255–57.
2. Brian Skyrms, "EPR: Lessons for Metaphysics," *Midwest Studies in Philosophy* IX (1984): 254.
3. Ibid., 254.
4. Tim Maudlin, "Causation, Counterfactuals, and the Third Factor," in *Causation and Counterfactuals*, 431.
5. Ludwig Wittgenstein, *Philosophical Investigations*, trans. G. E. M. Anscombe (New York: MacMillan, 1953), 20.

Index of Subjects

Index of Names