VOL. XI. No. 14. JULY 2, 1914

THE JOURNAL OF PHILOSOPHY
PSYCHOLOGY AND SCIENTIFIC METHODS

A DEFINITION OF CAUSATION. IV

WE have in the preceding paper passed in review all the kinds of causal process found under the heading "Properties of Matter"; the second principal group of cases marked out at the beginning of our investigation. Before going to the third group, that of electrical phenomena, it may be well very briefly to summarize the results so far.

From the point of view of one who looks for new light on the nature of causation the results of the second group were decidedly disappointing. Under the first group, Mechanics, we found two types clearly emerging: a certain serial type called a self-repeater, and one of composition. Under the "Properties of Matter" we obtained some cases which were believed to reduce to the mechanical types, but more which gave no result at all. Perhaps the science of the near future will either reduce these to mechanical or electrical cases, or prove them *sui generis*. But at present the scientific analysis gives no further data. This is, of course, one of the imperfections inherent in empirical method; yet there seems no dodging it, if we wish to remain in touch with what is assuredly known of the existent world.

Meanwhile an objection arises, accusing us of a certain arbitrariness. It may be asked, why do we not accept elasticity, density, etc., as ultimate causes, connected with their effects, motion, acceleration, etc., and needing no further analysis? We were willing, under Dynamics and Statics, to accept motion, mass, position, etc., as ultimate causes, leading to motion, etc., as effects. We did not then accuse science of having not analyzed sufficiently. Does it not, therefore, seem that we are biased in favor of a mechanical type, and so long as causes have not been reduced to such a type, arbitrarily consider them not yet analyzed? Certainly such a bias has no just place in an empirical investigation. This objection, however, misunderstands the situation. There is a definite ignorance with regard to elasticity, density, friction, such as is not the case in motion causing motion. With the dynamical phenomena, both cause and effect are distinguishable and verifiable by sense-observation. With elasticity, only the effect, the recovery of normal size and shape, is identifiable. There is no verifiable preceding or accompanying phenomenon in the body which can be distinguished from this recovery and assigned as its cause. It is believed that there is such a phenomenon, and that it is of the nature of a system of intramolecular stresses; but these stresses are not definitely described. Elasticity, then, is at present a name for an effect to which science has not yet correlated a cause. The only requirement for our investigation is that we are provided with a definite, identifiable cause, and a definite, identifiable effect distinguished from that cause. *Then* we can undertake, by analysis, to bring to light the relations that hold between them; and upon the application of this method to all known types of causation, to base a general definition of the causal relation. But this we can not do until *both* the cause and the effect are furnished. It is not that the mechanical has any inherent advantage over any other causal explanation; in the field of electricity, indeed, we shall find a different type, which is quite as clear and good. And what has been said of elasticity holds as well, *mutatis mutandis,* of resistance, density, crystallization, etc.

We pass now to the third group of causal events. The recent advances in this field are summed up in what is known as the electron-theory. Our task is then to expound the fundamental laws of that theory, which are believed to account for electrical phenomena so far as they are at present explained.

III. Causation in the Field of Electricity

The electron-theory explains a great many of the phenomena of light, heat, electricity, magnetism, chemistry, which, so far, mechanical theories have been unable to explain. It does not supersede or account for the laws of mechanics, but rather invokes them in its applications. The causal sequences that obtain among electrons are not then to be regarded as more fundamental than those of mechanics. "The corpuscular [electron] theory of matter with its assumptions of electrical charges and the forces between them is not nearly so fundamental as the vortex-atom theory of matter, in which all that is postulated is an incompressible frictionless fluid possessing inertia and capable of transmitting pressure."[1] We have here simply a type of events as yet unreduced to any other type. It is not however based simply upon hypothesis. "We have direct experimental proof of the existence of these corpuscles."[2] Yet, while certain of the laws governing electrons are established, there is, it appears, by

[1] J. J. Thomson, "Corpuscular Theory of Matter," page 2.

[2] *Op. cit.*, page 2.

no means a fixed body of doctrine throughout this region. The best we can do is to state some of the typical views, indicating where they are hypothetical rather than generally accepted. The type of causation in each view will serve to illustrate scientific procedure in this field, so far as that can be done at present. We begin with the more generally accepted properties of electrons.

1. Each electron at rest repels every other electron at rest. "The fundamental property of the electron which distinguishes it from ordinary matter is that it repels another electron, instead of attracting it, as two pieces of matter would do."[3] "Every electron placed at a distance of 1 cm. from another electron repels it with a force of 1.16×10^{-19} dynes" (p. 29). Each electron has thus a definite and constant negative charge, whose action consists in just this force of repulsion; and the electric current consists in the motion of such electrons. When, however, they move at a very high velocity, the mass increases, "just as if the ether in that space were set in motion by the passage through it of the lines of force proceeding from the charged body, and . . . the increase in the mass of the charged body arose from the mass of the ether set in motion by the lines of electric force."[4] This apparent increase of mass does not hold of the electron itself; that remains constant. Furthermore, this repulsion decreases with the square of the distance,[5] which means (analogously to what we saw in the case of gravitation) that the repulsion from a given electron runs outward uniformly in a straight line.

The cause of this repulsion is not assigned. "We shall not attempt to go behind these forces and discuss the mechanism by which they might be produced."[6] The most that can at present be done is to see the way in which this repulsion acts. The above account makes it comparable with gravitation, or any pressure or tension acting uniformly in a straight line. Whether considered as holding between electrons at rest or as between them when moving, it clearly has the same logical structure as was found above in Statics and Dynamics, in the serial type.

2. The same laws hold between the positive atoms at rest, *i. e.*, those which have been deprived, each of the same number of electrons.

3. "Every electron attracts every neutral atom from which one electron is removed, when placed at a distance of 1 cm. from it, with the same force—*viz.*, 1.16×10^{-19} dynes, or if two, three, etc., electrons have been removed, with a force two, three, etc., times that amount."[7] This attraction varies also inversely as the square of the

[3] E. Fournier d'Albe, "The Electron Theory," page 23.

[4] "Corpuscular Theory," page 29.

[5] *Ibid.*, page 1.

[6] "Corpuscular Theory," page 1.

[7] *El. Th.*, page 29.

distance. When this attraction is combined with the original motion of the electron, the resultant is an orbital motion of the electron. "These electrons are usually associated with atoms of ordinary matter, round which they describe circular or elliptical orbits, with periods approaching those of visible light-waves."[8] (This statement, however, seems to embody no more than a hypothesis). Such orbits are conceived analogous to those of our solar system; they are resultants of attraction and some original motion, and as such are believed to obey the laws of Kepler and Newton.[9] The cause of this attraction is not assigned, but the way in which it behaves is, once more, that of a force acting uniformly in a straight line, as in the case of the repulsions between like charges.

4. ". . . electrons moving side by side through the ether attract each other with a force proportional to their speed, and inversely proportional to the square of their distance apart."[10]

We now come to a distinctly hypothetical part of the theory. The attraction between electrons in motion, it is supposed, "balances their electrostatic repulsion as soon as they travel with the velocity of light" (p. 147). Also, "when an electron and a positive atom travel side by side through the ether, their original attraction is balanced by mutual repulsion, so that, again, when they travel with the velocity of light, they exert no mutual force" (*ibid.*). As the forces described in 1, 2, and 3, above were electrostatic forces holding between electrons at rest, so these are electrodynamic, between electrons in motion. They are claimed to be the fundamental events, in terms of which magnetism, radiant energy, etc., are described. According to the view here set forth, it follows that the way in which one electron influences another moving beside it is to retard its motion; if that other is stationary, to impart to it a motion in the opposite direction.[11] "A change of momentum of an electron produces a change of momentum in every other electron in the opposite direction" (p. 281). This mode of behavior is alleged to be exemplified in the electromagnetic wave, which is constituted by the swinging of one electron giving rise to that of another, and so on; each one swinging in a plane perpendicular to the direction of the propagation, and the propagation taking place with the velocity of light. As these waves produce radiant heat or light according to their length, it seems fair to say that the typical causal events (on this view) in the whole field of radiant energy and electricity are

[8] *Op. cit.*, page 280.

[9] *Cf.* the calculations, *op. cit.*, page 32, made on that assumption.

[10] *Op. cit.*, pages 280–281.

[11] *Op. cit.*, pages 176–79.

believed to be the motion of an electron and the imparting of motion from one electron to another.

The analogy between these types of causation and the mechanical ones is striking. "It is as if the electric momentum—the momentum of electric charges—destroyed in one body reappeared in another, just as in a collision there is a transfer of momentum" (pp. 237–238). Each electron imparts a momentum equal to its own, in a reversed and parallel direction, to the next; this imparts in the same way to the next, and so on. It is not, apparently, done by contact. Nevertheless the electron's influence is conceived as permeating all the adjacent continuous space, just as gravitation does; for it acts upon other electrons at any given distance from itself. It is supposed to be continuously active as we go out from it in space. We are reminded of the case of the bar balanced on a fulcrum, where the downward pressure is communicated from one end to the other. The causal process in the present case seems to lend itself to just such a serial description as in that one.

Much of this is tentative (if I understand the matter correctly). The law of inverse squares is, however, I believe, generally admitted, and the uniform action which that implies may be reduced, as in the case of gravitation, to the same serial type as uniform motion. We are concerned only to show that the kind of explanation that is offered in the hypothetical portion of the doctrine is of the same general logical structure.

A quite different mode of conceiving the fundamental electron-event is that of F. Bohr.[12] This dispenses with the notion of an ether-wave entirely, replacing it by that of the projection of a uniformly rotating electron; the rotation accounting for the periodicity of light. This is clearly a dynamical type,—though its source is, I believe, not explained as yet. On the whole, then, it seems safe to say that there are at present no causal explanations offered in this field which can not be reduced to a type analogous to that found in mechanics.

Before proceeding to analyze the results obtained for the cause-effect-situation, in general, we must notice an already mentioned definition, apparently grounded on an exhaustive empirical survey, by Professor Ostwald. According to him, whenever we have causation, we have an uncompensated difference between two adjacent intensive or potential quantities of the same kind. Now this is indeed empirically grounded, but—in accord with the standpoint of energetics—it neglects analysis. The "*Stärken*" or potential factors are not analyzed, nor is the way in which the effect follows, or even the effect itself, subjected to scrutiny. The definition, though not as

[12] *Philos. Mag.*, July, 1913.

abstract as those of Russell or Natorp, remains less than concrete. Had Professor Ostwald attempted to perform analysis, he would, I venture to think, have found that this formula is largely tautological. For intensive quantities, as he uses them, are those factors of energy which contain a time-factor in their definition. The capacity-factors—mass, specific heat, etc.,—do not change; the "*Stärken*"—velocity, temperature, etc.,—do change. If there is change then, *i. e.,* an event, it must occur in the intensive factors; they are defined as those which alone can change, in a given system. Why it is that there must be a difference between the two is equally obvious. For otherwise there would be no opportunity for change. If all bodies moved with the same velocity relatively to one another, there would be no relative motion, and no mechanical events. If all bodies, and empty space as well, were at the same temperature, there would be no heat-events. In fact, we have already found that an uncompensated potential factor *meant* loss of equilibrium—whence is obvious the tautological character.

Nevertheless there is a sense in which Ostwald's formula is not tautological. It is a generalization from the second law of thermodynamics, which is by no means a tautology. That law has a positive meaning which Ostwald's formula has in its generalized form lost. It tells us that the event which happens is in *one certain direction* rather than its opposite. Heat in available form decreases rather than increases. The tendency is always toward lower rather than higher levels. Professor Bergson regards this as so fundamental a trait of all material processes as to warrant us in defining the very nature of matter thereby. Whether that may turn out true or not, it is a universal empirical character for which, so far as I know, no explanation has been given by science. As to the mode in which it acts, it is believed to be that of a uniform process; whether it takes the form of pressure, motion, radiation, or electric potential. The reason why this empirical property holds, then, not being yet assigned, we have here no data for analysis; but the mode in which the change from higher to lower levels occurs has already been examined in the fields of mechanics and electricity.

The final task is to define more precisely the invariant structure which we have found in all the cases which yielded results. Two types were discovered: that of composition, and that of a series which was called a self-repeater. Are these distinct, or reducible to one type? Let us first briefly recall them, and then proceed to analysis.

The invariant *composition* was found in Statics and Dynamics, and afterward seen to be universally present, since every phenomenon is a complex one. Its nature was extremely simple: two forces, motions, accelerations, pressures, combine to produce a third in which they are preserved intact. That is all we need now recall of that type. The invariant *series* was found, in Statics, in the principle of transmissibility and in the moment of a force about an axis; in Dynamics, in uniform motion, rest, collision of bodies; in the law of inverse squares holding for so many phenomena outside Mechanics, as gravitation, attraction, and repulsion of electrons, and in fact throughout the field of electron-theory. Its structure was temporarily defined as determined by a first term x, which was followed or accompanied by a second term x', essentially similar to x in all respects save one or more definitely named differences (position, magnitude). It was suggested, but not proved, that this would suffice to generate an endless number of following cases, x'', x''', etc. We have now to study these two invariants more closely.

Each starts from a duality. By this is meant that in each the cause is *two terms* with a certain relation between them. This is self-evident in the case of composition, where the two factors plus the relation of *combination* determine unambiguously the resultant. In the case of the series, as was already suggested, two terms with a relation of sameness—and also of difference—between them determine unambiguously the remainder. Let us now see by some further analysis that this is really the case. The terms are conceived different in ordinal position—if that phrase may properly be used of less than three. If not, we must find some other phrase such as temporal position or spatial direction. The first is more fundamental than the second *only* in the sense that the second is defined by reference to the first, but not conversely. It is not meant that the first gives rise to, or necessitates, or in any way accounts for, the second. *E. g.*, the first term x of the series is a body in a certain spatial position at a certain time; the second term is a body in another spatial position at another time. The latter body is defined as *the same as* the first (*whatever* that may mean), and its spatial and temporal positions are defined as *later than or beyond* those of the first. Of course this definition is theoretically reversible, and we might define the first by means of the second; this would indeed be *a priori* possible, but would not be an adequate account of the particular types we have been studying. Here is where our procedure once more definitely diverges from the method of *a priori* logic, which finds asymmetry reducible to symmetry.[13] The reason why we must diverge from this interpretation is that we are dealing with existences. In the existential world, things do not conform to this

[13] J. Royce, following C. L. Franklin and A. B. Kempe, in *Trans. Am. Math. Soc.*, Vol. 6, pages 353–415, and in *this* problem, B. Russell, *Proc. Arist. Soc.*, 1912–13, pages 10–11, 15, 21.

ideal symmetrical arrangement; the relation of earlier to later is existentially irreversible. We accept this irreversibility as fact, recognizing that the relations between these two members are not symmetrically describable except by abstraction from this irreversibility. Science regards the first as existing without the other, but not conversely; x exists before x', when x' does not as yet exist; but it is not true that x' exists when x has no existence; for the past has a certain title to existence which the future has not. This is a fundamental attribute of time, which *a priori* methods generally neglect. Stated in terms of time, then, our point is that the past has a certain existential rank higher than that of the future. Accordingly the second member of the series must be defined by reference to the first but not conversely; otherwise our definition of the series would be abstract and would neglect this empirical character of precedence. But the first member does not, of course, suffice to define any series; it does not determine that there *is* a second term which is defined in terms of it. It *needs* the second term in order to constitute the series, as truly as that term needs it. And in all this we claim to be stating the structure of the empirical process. It is that character of the inquiry alone which precludes the symmetry and consequently renders ultimate the difference between the two terms.

The first two members give rise to the series in the following manner: x is followed by x' where x' is defined by its sameness with x, and also has a certain additional difference whereby it is made a distinct case whose existence is independent of x. We now concentrate our attention on the sameness, for it is this relation that will bear the burden of what follows. It is a given fact that the relation "followed by x'" is associated with x; or as science and ordinary thought put it, x has the particular property of being followed by x'. Now this simply means that this property is predicated of x, or is in the last analysis in part to be identified with x, as a thing is identified with its properties.[14] That is, it has the relation of sameness with x. So we have, in addition to the above relation, x' is the same as x, this further relation, x is the same as this particular property of being followed by x'. We have, then, two relations, which may be symbolized thus: $x'Rx$, and xSy, when y is this property in question. The inference is inevitable and uniquely determined, $x'RSy$; x' has the relation to y which is the "relative product" of the relations R and S. Now in this case R and S are both the relation "the same

[14] The externality of relations would seem to forbid this, but the matter may be stated consistently with that view. x has the relation "followed by x'" and x' is the same with, or defined by x, hence x' has the relation "followed by itself."

as,'' which happens to be a transitive relation. Hence the result must read, $x'Ry$; the ''relative product'' is equivalent to the original relation. Translating this, we have: x' has the property of being followed by itself, *i. e.,* it implies another case of itself.

If x' implies *another case* of itself, this latter will be defined just as x' was defined, consistently of course with its distinctness from x'. Its relation to x' will be seen to be exactly analogous to the relation of x' to x. It is, in fact, the third member of the series, which we called x''. Being exactly analogous to x', it in turn will be subject to the same reasoning, and will imply one more case of itself, which is the fourth member of the series, x'''. It is clear that the series must continue indefinitely.

Thus, it is claimed, the first two terms, related as we found them to be, determine the rest of the series. Inasmuch as more seems to come out of the premises than was put into them, we may profitably make some comments before going further, and meet some obvious objections.

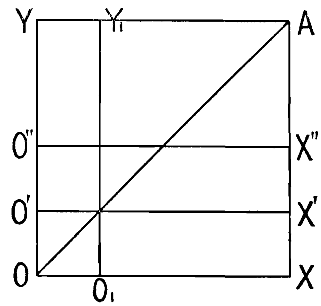
Notice, first, that every term and relation here found is strictly particular—or individual if that is a less universalized term. x and x' clearly are so. The ''property of being followed by'' sounds general, but is not here used as general; we dealt only with this particular case of it. We *found* by inspection of the situation that it must *become* so; but that is not determined either by the statement ''x' is the same as x,'' or by the statement, ''x is the same as the property.'' Nor is it tacitly presupposing a universal when we call x'' ''another case'' of x'; for two cases only are meant, and it is not yet known that there could be more. Two cases of course do not constitute a universal—nor do any finite number of cases.

Notice, second, that the result is in no way tacitly taken for granted at the outset. When we say x' is the same as x, it is a very natural criticism to reply that we really mean ''x' is the same as x in respect to the property of being followed by x'.'' If this were true, we should have already begged the result; we should not have defined the series by the two terms and their relation alone, but by two terms so defined as directly to include an endless series of terms. But it is not true. To define x' to be the same as x is not, *so far,* to define it as the same as this property of x. For x can and does exist without this property: in the case of a body at rest it is not followed by the x' which follows it in the case of motion. The being followed by x' is not a necessary consequence of x; we saw this above, when we showed that x' is not dependent for its existence upon x. But even if the property of being followed *were* a necessary consequence of the existence of x, we should not have tacitly assumed it; for in general it does not follow, if $x'Rx$ and xSy, that $x'Ry$—as we saw just now. It *does* follow in the case of a transitive relation such as sameness; and the fact is so familiar that we know instantly it *will* follow and so we seem to have assumed it at the beginning. Indeed, if this criticism were not very probable, we could have phrased the whole matter much more simply thus: x is followed by x' which resembles x and therefore must be followed by a particular case resembling itself, x''.

A second natural criticism is an accusation of self-contradiction. If this, and the answer to it, are pushed very far, we shall land in a dialectic. This I believe to be no sound objection to the criticism; but it has been expressly barred from this investigation. We desire only to show that the above description of the series uses nothing which is contradictory to the ordinary scientific usage. That two terms can be the same while different does not seem contradictory to common reflection; for we speak, and the scientist speaks, of the *same* (and not merely an equal) mass in different positions, the *same* (and not merely an equal) momentum in different bodies, etc. That a term can be "followed by itself" does sound paradoxical; but when the phrase: "by another case of itself" is substituted, it does not sound paradoxical at all. For that is a property we ascribe to any universal. One case of it is (in part) just like another and may follow or precede it. We claim, then, merely that our definition is no more contradictory than the practise and speech of science. If that is philosophically condemned, let it be so; but if not, neither should our view be condemned. It may be remembered that *one* reason for considering science as not ultimate has already seemed to disappear, *viz.*, a generalization we made from the principle of composition. But of course others remain.

Two related factors, then, seem necessary and sufficient to every cause of the serial type, and these two determine what follows, *i. e.*, the effect. One alone would, it appears, never produce anything. It would be a potentiality only. A mere potentiality would never produce an actual event: the series would not start but for the second member being added to the first. Thus all causes of the serial type may be considered as *Auslösungen:* a potentiality *plus* a motion or change. The case of spark and gunpowder is not at all unique. But this is only one side of the matter. The two main types, the series and the composition, are alike in more than the duality of the causal member.

The principle of composition may be stated in serial form. Every force acts in a certain line, and is defined by the acceleration it imparts along that line, represented by a certain length. The resultant line is uniquely determined by the component lines, in length and direction. How happens this determination? Let us take a simple case. Imagine the components to be *OX* and *OY,* the resultant *OA.* How then is the certainty secured, that *OA* and



nothing else is the resultant? This situation may be put in serial form in many ways, one of which is as follows. The length *OX* from *O* takes us just as far in a fixed direction from *OY* as the length *O'X'*. The length *O'X'* takes us just as far in the same direction from *OY* as the length *O''X''*. And so on; the length *YA* being one of the members of this series. The series is quite analogous to the self-repeater above. We do the same thing for the lengths *OY,* O_1Y_1, etc., and thus arrive at *XA.* Now it is being a member of such a series that determines *A* to be the required terminus of the resultant. Not *OA,* but some other line *OB* would be the resultant, if the series were not a valid one, truly describing the nature of space. The point *A* is determined to be where it is by virtue of being in both of these series at once. Thus viewed, the composition of forces is not a mysterious union of two entities to produce by some magic a third, but the meeting-point of two repeating series.

The necessary and sufficient condition of the series was found to be two terms in a certain relation. That is, given that much, the rest *of necessity and by pure deduction* follows. Hence the series contains *necessary connection.* Had Hume examined specific cases of causation, or had he even told us what necessary connection meant, the present almost universal philosophical skepticism in regard to its existence might not have come into being. One great obstacle, too, to his search, was that he treated a cause as *one* instead of *two.* At any rate, the denial of real necessity has usually been made without a fair examination of the evidence. If the evidence here offered is sound, and the analysis of it correct, we seem to have obtained an answer to Hume of a very different character from the usual one: an answer which finds necessity in the empirical contents of experience, rather than in the form imposed on it by mind. We have examined those contents by themselves, as science alleges that they occur in the material world. No hypothesis of the presence of mind as law-maker was needed to account for necessary connection between cause and effect, nor any assertion of the independent "subsistence" of universals. In fact we have, if no mistake has been made, been able to derive the concept of the universal, *i. e.,* that which may have *any* number of instances, from a system of two particular terms and a particular relation. Of course this result has not yet been subjected to criticism, such as the different schools of philosophy might make. One suggestion may perhaps without impropriety be offered. If the method here pursued gives us a result (the objective existence of necessary connection) which philosophers have been in the past unable to reach, and in default of which they were driven to various modern ''schools'' to account for the persistent human error of believing in it—if it gives us this, does there not seem to be less reason to fear criticism? For if we are correct, we have something, the lack of which has occasioned a number of modern philosophic systems. And if such a view *could* survive their criticism the practical belief of every one that events are necessitated by their causes is justified. Nature, in fact, on our view, deduces itself from its past—so far as there is causation.

The meaning of causation which applies to the external world appears then to be: two facts or events such that one precedes the other, temporally or logically, and the second is defined by the first, *i. e.,* the same as the first; a second case of it with added differences. This constitutes the cause. The definition would repay further analysis, I think, but let this now suffice. The effect is the logical deduction from this, the necessary consequence; a never finished series. In practise we generally single out the member, or members, of this series, that for the purpose in hand interests us, and consider it or them the cause with reference to what follows, or the effect with reference to what precedes. We may now see why the momentum of a body *must* be conserved, why the law of inertia *must* hold; in short, why anything that is caused *must* be what it is. That does not, of course, enable us to say that causation is everywhere present. There may very well be indeterminate beginnings; that there are, I have elsewhere tried to show. If there are, then the universe would seem to be a growing one: for since no causal series is ever finished, new beginnings would simply add to the content of the universe, and the series started by them—if any—would never be destroyed. But aside from this interesting speculation, the main results of this investigation are, that there is necessity in the existent world, and that it is not an absolute *a priori* necessity, but one *derived from* the existence of a dyadic relation. No necessity, probably, could be derived from *one term alone,* where Hume and his successors always looked for it, but only from two. Whether or not this result is valid for ultimate reality, is another question.

DARTMOUTH COLLEGE. W. H. SHELDON.