The Quantum Theory and Reality

The doctrine that the world is made up of objects whose existence is independent of human consciousness turns out to be in conflict with quantum mechanics and with facts established by experiment

by Bernard d'Espagnat

Any successful theory in the physical sciences is expected to make accurate predictions. Given some well-defined experiment, the theory should correctly specify the outcome or should at least assign the correct probabilities to all the possible outcomes. From this point of view quantum mechanics must be judged highly successful. As the fundamental modern theory of atoms, of molecules, of elementary particles, of electromagnetic radiation and of the solid state it supplies methods for calculating the results of experiments in all these realms.

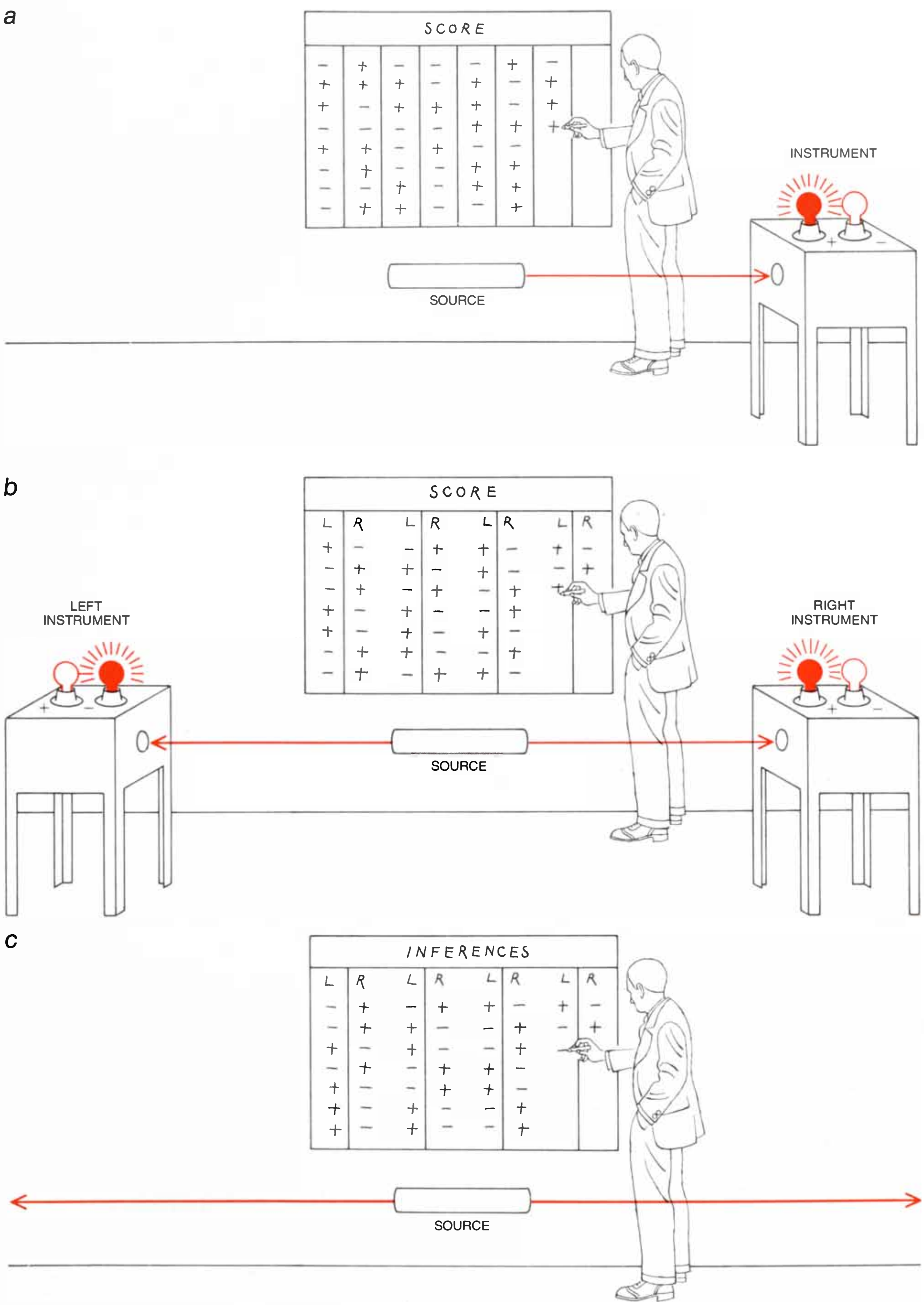
Apart from experimental confirmation, however, something more is generally demanded of a theory. It is expected not only to determine the results of an experiment but also to provide some understanding of the physical events that are presumed to underlie the observed results. In other words, the theory should not only give the position of a pointer on a dial but also explain why the pointer takes up that position. When one seeks information of this kind in the quantum theory, certain conceptual difficulties arise. For example, in quantum mechanics an elementary particle such as an electron is represented by the mathematical expression called a wave function, which often describes the electron as if it were smeared out over a large region of space.

This representation is not in conflict with experiment; on the contrary, the wave function yields an accurate estimate of the probability that the electron will be found in any given place. When the electron is actually detected, however, it is never smeared out but always has a definite position. Hence it is not entirely clear what physical interpretation should be given to the wave function or what picture of the electron one should keep in mind. Because of ambiguities such as this many physicists find it most sensible to regard quantum mechanics as merely a set of rules that prescribe the outcome of experiments. According to this view the quantum theory is concerned only with observable phenomena (the observed position of the pointer) and not with any underlying physical state (the real position of the electron).

It now turns out that even this renunciation is not entirely satisfactory. Even if quantum mechanics is considered to be no more than a set of rules, it is still in conflict with a view of the world many people would consider obvious or natural. This world view is based on three assumptions, or premises that must be accepted without proof. One is realism, the doctrine that regularities in observed phenomena are caused by some physical reality whose existence is independent of human observers. The second premise holds that inductive inference is a valid mode of reasoning and can be applied freely, so that legitimate conclusions can be drawn from consistent observations. The third premise is called Einstein separability or Einstein locality, and it states that no influence of any kind can propagate faster than the speed of light. The three premises, which are often assumed to have the status of well-established truths, or even self-evident truths, form the basis of what I shall call local realistic theories of nature. An argument derived from these premises leads to an explicit prediction for the results of a certain class of experiments in the physics of elementary particles. The rules of quantum mechanics can also be employed to calculate the results of these experiments. Significantly, the two predictions differ, and so either the local realistic theories or quantum mechanics must be wrong.

The experiments in question were first proposed as "thought experiments," intended for the imagination only. In the past few years, however, several versions of them have been carried out with real apparatus. Although not all the findings are consistent with one another, most of them support the predictions of quantum mechanics, and it now seems that unless some extraordinary coincidence has distorted the results the quantum-mechanical predictions will be confirmed. It follows that the local realistic theories are almost certainly in error. The three premises on which those theories are founded are essential to a common-sense interpretation of the world, and most people would give them up only with reluctance; nevertheless, it appears that at least one of them will have

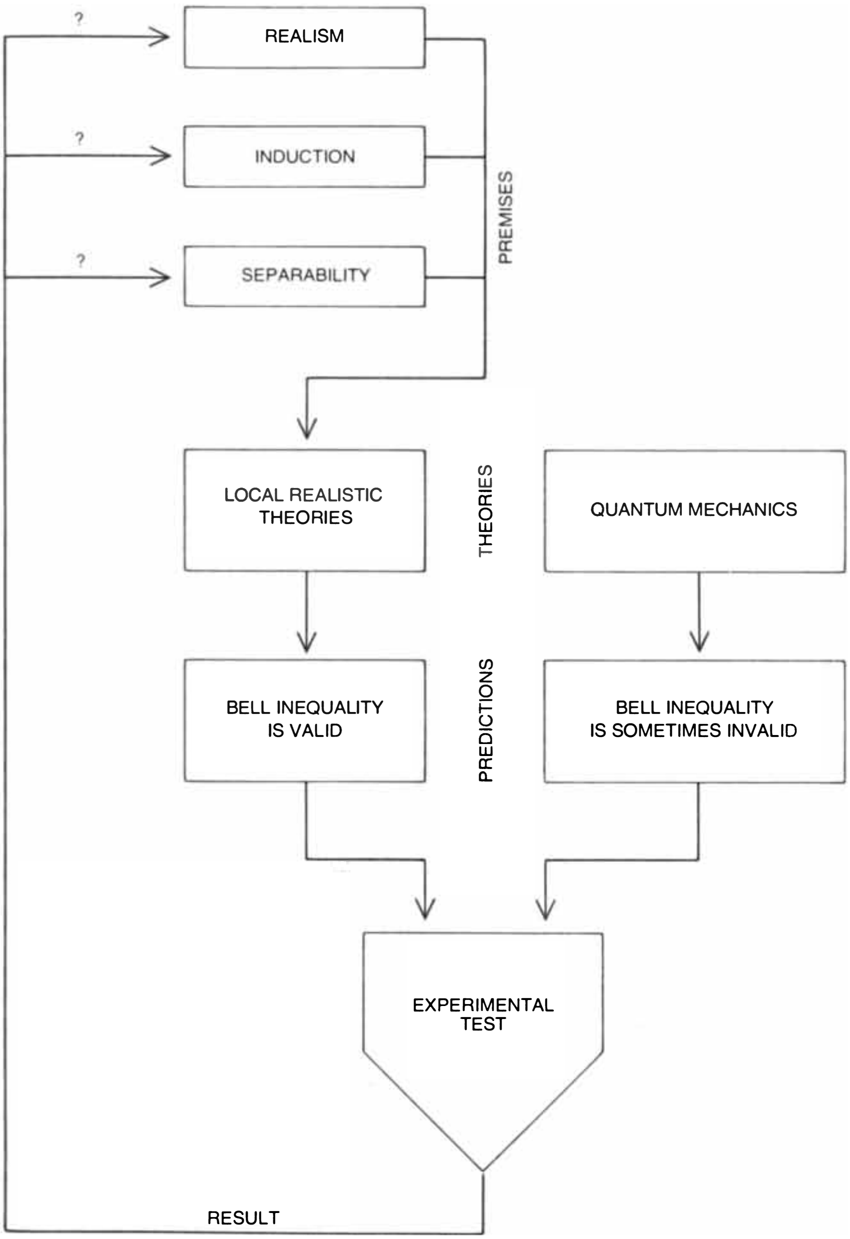
CORRELATIONS BETWEEN DISTANT EVENTS can form the basis of conclusions about the structure of the world. Suppose a physicist sets up an experiment in which subatomic particles such as protons are fired one at a time into an instrument that can give only two possible readings, plus and minus (*a*). He finds that for some protons the reading is plus and for others it is minus, but he cannot tell whether the instrument measures some real property of the protons or merely records random fluctuations. The physicist then arranges two identical instruments with a source that emits two protons simultaneously (*b*). He observes a strict negative correlation: whenever one instrument reads plus, the other reads minus. On the basis of this correlation the physicist concludes that a real property of protons is responsible for the readings and that its value is determined before the protons leave the source. If the sample of particles measured meets certain statistical tests, he can go on to infer that every pair of protons emitted by the source consists of one proton with the property plus and one with the property minus, even if neither proton is submitted to a measurement (*c*). The conclusions are reasonable if three premises are accepted as valid: that at least some properties of the world have an existence independent of human observers, that inductive inference can be applied freely and that a measurement made with one instrument cannot influence the result of a measurement made with the other instrument. A more restrictive form of the last premise forbids such influences only if the two measurements are so nearly simultaneous that the influence would have to propagate faster than light. The premises can be identified as realism, the free use of induction and separability; the more restrictive version of the separability premise is called Einstain separability or Einstein locality. Any theory that incorporates them is called a local realistic theory.


a
SCORE
INSTRUMENT
SOURCE
b
SCORE
L R L R L R L R
LEFT INSTRUMENT
RIGHT INSTRUMENT
SOURCE
c
INFERENCES
L R L R L R L R
SOURCE

to be abandoned or modified or in some way constrained.

The experiments are concerned with correlations between distant events and with the causes of those correlations. For example, suppose two particles a few meters apart are found to have identical values of some property, such as electric charge. If this result is obtained once or a few times, it might be dismissed as coincidence, but if the correlation is detected consistently in many measurements, a more systematic explanation is called for. It would make no difference if the measured values were always opposite instead of the same; the correlation would then be a negative one, but its magnitude would be just as great, and it would be just as unlikely to arise by chance.

Whenever a consistent correlation between such events is said to be understood, or to have nothing mysterious about it, the explanation offered always cites some link of causality. Either one event causes the other or both events have a common cause. Until such a link has been discovered the mind cannot rest satisfied. Moreover, it cannot do so even if empirical rules for predicting future correlations are already known. A correlation between the tides and the motion of the moon was observed in antiquity, and rules were formulated for predicting future tides on the basis of past experience. The tides could not be said to be understood, however, until Newton introduced his theory of universal gravitation.



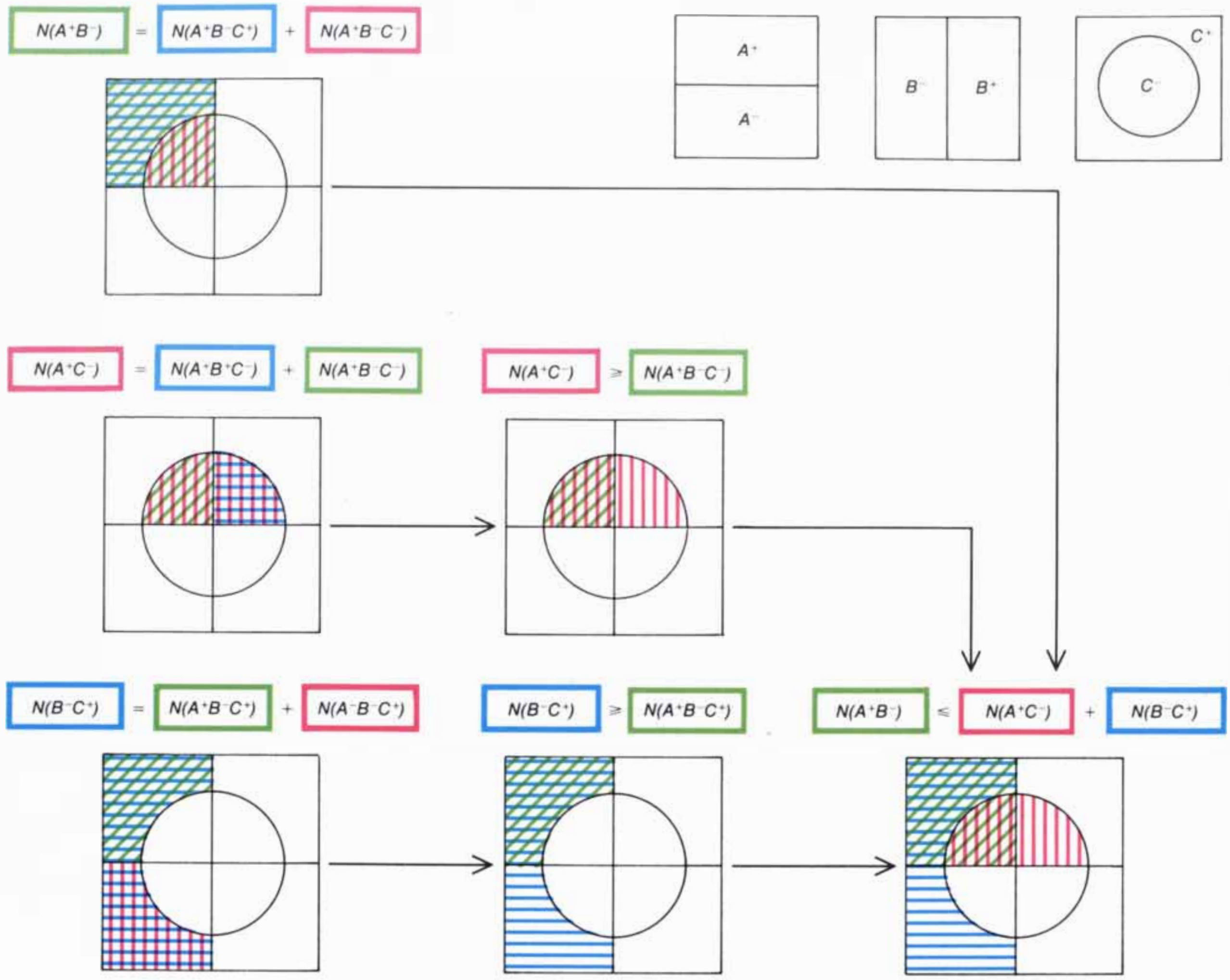
LOCAL REALISTIC THEORIES and quantum mechanics make conflicting predictions for certain experiments in which distant events are correlated. In particular, local realistic theories predict that a relation called the Bell inequality will be obeyed, whereas quantum mechanics predicts a violation of the inequality. There is strong experimental evidence that the inequality is violated in the way predicted by quantum mechanics. Local realistic theories therefore seem to be untenable, and at least one of the premises underlying those theories must be in error.

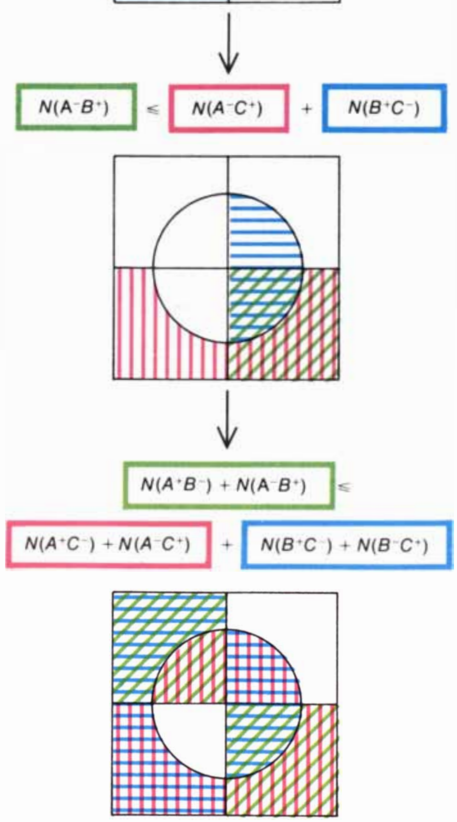
The need to explain observed correlations is so strong that a common cause is sometimes postulated even when there is no evidence for it beyond the correlation itself. Whether or not this procedure can always be justified is a central issue in the conflict between quantum mechanics and local realistic theories. The correlations in question are between observations of subatomic particles, where a quantum-mechanical description, with its attendant epistemological hazards, is indispensable. The predictions of local realistic theories, however, can be illustrated by considering how correlations between distant events are explained in a more familiar context, where quantum mechanics need not be introduced.

Imagine that a psychologist has devised a simple test, which a subject must either pass or fail, so that there can be no ambiguity in the results. The psychologist finds that some people pass and some fail, but he does not know what distinguishes the two groups other than their performance on the test itself. In other words, he cannot tell whether the test measures some real aptitude or attribute of the subjects or whether the results are haphazard.

It seems there is no general solution to this problem, but in a special case it might be solved. Suppose the test is administered not to a series of individuals but to a series of married couples and that a strong correlation is detected in their answers. The procedure might consist in separating the husbands from the wives before the test and then giving the test to each of them in isolation. When the results are analyzed, it is found again that part of the population has passed and part has failed, but in the case of each couple where the husband passed so did the wife; similarly, whenever the husband failed so did the wife.

If this correlation persists after many couples are tested, the psychologist is almost sure to conclude that the response of each subject is not determined randomly at the time of testing. On the contrary, the test must reveal some real property or attribute of the subjects. The property must already be present in the subjects before they are tested, and indeed before they are separated. Chance may have had some influence on the development of the property,





BELL INEQUALITY, formulated by John S. Bell of the European Organization for Nuclear Research (CERN), can be proved in two stages. The inequality applies to experiments with particles that have three stable properties, *A*, *B* and *C*, each of which can have the values plus and minus. Thus there are 2^3, or 8, possible classes of particles, corresponding to the eight regions of the diagrams shown here. If a particle has been found to have the properties A^+ and B^-, then it must be a member either of the class $A^+B^-C^+$ or of the class $A^+B^-C^-$. Hence if $N(A^+B^-)$ represents the number of such particles, it must be equal to the sum $N(A^+B^-C^+) + N(A^+B^-C^-)$. In a similar way it can be shown that $N(A^+C^-)$ is equal to $N(A^+B^+C^-) + N(A^+B^-C^-)$, from which it follows that $N(A^+C^-)$ is greater than or at least equal to $N(A^+B^-C^-)$. The same reasoning leads to the conclusion that $N(B^-C^+)$ must be greater than or equal to $N(A^+B^-C^+)$. These three relations can now be combined to yield a further inequality, which asserts that the number of A^+B^- particles cannot exceed the sum of the A^+C^- particles and the B^-C^+ particles. The same relation holds if all signs are reversed to give the inequality $N(A^-B^+) \leq N(A^-C^+) + N(B^+C^-)$. The last two inequalities can be added to yield a relation among all individual particles for which two properties have opposite values.

since not all the couples possess it, but that influence must have been exerted at some time before the husbands and the wives were separated. It was only then, while the husbands and the wives were still united, that they could have acquired any traits that would induce them to respond consistently the same way. Thus the correlation is explained by attributing it to a common cause antecedent to the test.

One other explanation that must be excluded in deriving this conclusion is the possibility that husbands and wives could communicate with each other while they were taking the test. If some means of communication were available, there would be no need for any tested attribute to exist beforehand. Whichever spouse was given the test first could choose a response at random and send instructions to the other, thereby creating the observed correlation. In giving a psychological test it would not be hard to guard against subterfuge of this kind. In the extreme case the tests could be made so nearly simultaneous, or husbands and wives could be tested at sites so far apart, that a signal moving no faster than light could not arrive in time to be of any value.

Once having decided that the test measures some real property of individuals, the psychologist can take a further step and make an inductive inference. If the couples already tested constitute an unbiased sample of some population of couples, and if the sample meets certain statistical standards, the psychologist can infer that any couple taken from the same population will be made up of a husband and a wife who either both possess or both do not possess the property measured by the test. By the same principle he can conclude that in any large, unbiased sample of couples who have not yet been tested some of the couples will have the property and some will not. The confidence of these assertions approaches certainty as the size of the sample increases. Hence both the correlation within couples and the existence of differences between couples are inferred to exist even in the segment of the population that has not been submitted to any test.

These conclusions rest on the same three premises that form the basis of local realistic theories. Realism is a necessary assumption if one is to believe at least some tests measure stable properties that exist independently of the experimenter. It was necessary to assume the validity of inductive inference in order to extrapolate from the observed data to the segment of the population that had not yet been tested. Separability was incorporated in the assumption that husbands and wives being tested cannot communicate with each other. If the tests are given simultaneously, so that any signal passing between husbands and wives would have to propagate faster than the speed of light, the assumption is equivalent to Einstein separability.

At first the conclusions drawn from this hypothetical experiment in psychology seem to follow quite obviously from the data. An epistemologist might nonetheless maintain that the conclusions are uncertain. In particular an epistemologist trained in the foundations of quantum mechanics might argue that there is no logical necessity for accepting the three premises of the psychologist's argument; hence neither would it be necessary to conclude that a correlation existed between the husbands and wives before they were tested, or that differences existed between the couples before any tests were given. The psychologist is likely to find these objections laughable, an expression of misplaced doubt or of a very unscientific adherence to paradox. In the literature of quantum mechanics, however, there are numerous arguments similar or equivalent in form to this one, all purporting to show that correlations and differences need not exist until they are measured.

A singular feature of quantum mechanics is that its predictions generally give only the probability of an event, not a deterministic statement that the event will happen or that it will not. The wave function employed to describe the motion of an elementary particle is often interpreted probabilistically: the probability of finding the particle at any given point is proportional to the square of the wave function at that point. As I mentioned above, a wave function can sometimes be spread out over a large region, which implies that the probability can also be broadly distributed. Of course, when a measurement is actually made at some chosen point, the particle must either be detected or not be detected; the wave function is then said to collapse. Suppose the particle is detected. The question of epistemological interest is then: Did the particle have that definite position all along, even before the measurement was made?

The conclusions of the psychologist, if they could be transferred to this context, would imply that the position of the particle was well defined from the start, just as the attribute discovered in some members of the population was deduced to have existed before any tests were given. According to this argument the position of the particle was never indeterminate but was merely unknown to the experimenter.

Most authorities on the quantum theory would disagree. One exception among physicists was Einstein, who throughout his life remained dissatisfied with the probabilistic nature of the interpretations generally given to quantum mechanics. He based his most incisive criticism of those interpretations on an argument that was somewhat similar to the one I have attributed to the psychologist. In 1935 Einstein published a paper with two young colleagues, Boris Podolsky and Nathan Rosen, in which he stated his objections explicitly. He did not maintain that the quantum theory is wrong; on the contrary, he assumed that at least some of its predictions must be correct. What he proposed was that the quantum-mechanical description of nature is incomplete or approximate. The motion of a particle must be described in terms of probabilities, he argued, only because some of the parameters that determine the motion have not yet been specified. If the values of these hypothetical "hidden parameters" were known, a fully deterministic trajectory could be defined.

A number of counterarguments to Einstein's proposal have been formulated. For now I shall mention only one of them, which is based on the criterion of utility. It is immaterial, the argument states, whether or not hidden parameters exist, or whether differences between married couples exist in the absence of a test. Even if they do exist, they should not be incorporated into any theory devised to explain the observations, and so they can be said to have no scientific existence. The exclusion of the hidden parameters is justified by the conjunction of three facts. First, the mathematical formalism of the theory is simpler if any hidden parameters are ignored. Second, this simple formalism predicts results that are confirmed by experiment. Third, adding the hidden parameters to the theory would give rise to no supplementary predictions that could be verified. Thus the assertion that hidden parameters exist is beyond the reach of experiment and is a proposition not of physics but of metaphysics.

This defense of the conventional interpretation of quantum mechanics dismisses any hidden parameters as being superfluous and ultimately, perhaps, meaningless. Recent theoretical developments have shown that their actual status is quite different. The hypothesis that hidden parameters exist does in fact lead to experimental predictions differing from those of quantum mechanics. Hidden-parameter theories, and local realistic theories in general, place a limit on the extent to which certain distant events can be correlated; quantum mechanics, in contradistinction, predicts that under some circumstances the limit will be exceeded. Hence it should be possible, at least in principle, to devise an experimental test that will discriminate between the two theories.

Suppose a physicist has devised a test that can be carried out on subatomic particles such as protons. After many trials he finds that some protons pass the test and others fail, but he does not know whether he is measuring some real property of the protons or merely observing random fluctuations in his apparatus. He therefore tries applying the test not to individual protons but to pairs of them. The protons that make up each pair are initially in close proximity, having been brought together by some well-defined procedure that is the same for all the pairs. The protons are then allowed to separate, and when they have moved some macroscopic distance apart, they are tested, simultaneously for some pairs and with an interval between the tests for the remaining pairs. The physicist discovers a strict negative correlation: whenever one proton in a pair passes the test, the other proton invariably fails.

The situation of the physicist has obvious similarities to that of the psychologist giving a test to married couples, and the same reasoning might be applied to the results of the physical experiment. If realism, the free use of induction and Einstein separability are all accepted as premises, then the physicist is justified in concluding that his test does measure some real property of protons. For the correlation to be explained the property must exist before the protons in each pair are separated, and it must have some definite value from then until the measurement is made. Furthermore, if additional pairs of protons are prepared by the same method, the physicist knows that in each case one proton will have the property and one will not, even if neither proton is actually tested.

Is there any real test that can be carried out on subatomic particles with results like these? There is. It is a measurement of any one component, defined along some arbitrary axis, of the spin of a particle. The spin attributed to a subatomic particle is analogous only in some respects to the spin angular momentum of a macroscopic body such as the earth. For the purposes of this discussion, however, there is no need to introduce the details of how spin is treated in quantum mechanics. It will suffice to note that the spin of a particle is represented by a vector, or arrow, that can be imagined as being attached to the particle. A projection of this vector onto any axis in three-dimensional space is the component of the spin along that axis. A well-established but nonetheless surprising property of protons (and many other particles) is that no matter what axis is chosen for a measurement of a spin component the result can take on only one of two values, which I shall designate plus and minus. (A measurement of a component of the earth's spin would give very different results; depending on the direction of the component, it could have any value from zero up to the total angular momentum of the earth.)

A strict negative correlation between spin components is observed when any two protons are brought together in the quantum-mechanical configuration called the singlet state. In other words, if two protons in the singlet state are allowed to separate and the same component of spin is subsequently measured on both particles, it will always be plus for one proton and minus for the other. There is no known means of predicting which particle will have the plus component and which the minus component, but the negative correlation is well established. It makes no difference what component of the spin the experimenter chooses to measure, provided the same component is measured for both particles. It also makes no difference how far the protons travel before the measurement is made, as long as there are no perturbing influences, such as other particles or radiation, along their paths.

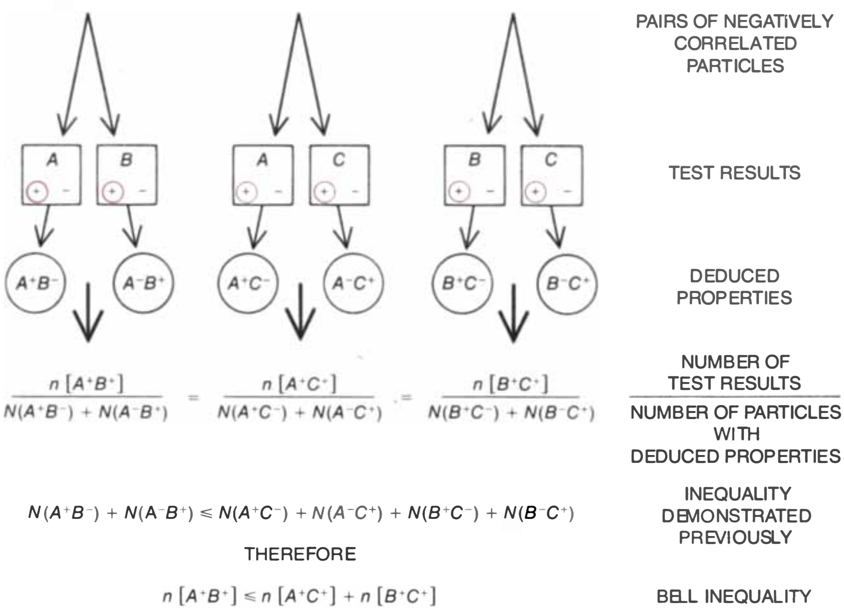
In this simple measurement there is no conflict between the predictions of quantum mechanics and those of local realistic theories. A conflict can arise, however, when the experiment is made somewhat more complicated.

The vector that represents the spin of a particle is defined by components along three axes in space, which need not necessarily be at right angles to one another. For a vector associated with a macroscopic object in everyday life, one would assume as a matter of course, and with good reason, that all three components have definite values at all times; the value of a component might be unknown, but it cannot be undefined. When this assumption is applied to the spin vector of a particle, however, it becomes highly suspect, and indeed in the conventional interpretation of quantum mechanics it is dismissed as an instance of a hidden-parameter theory. The problem is that no experiment can be devised, even in principle, that would provide information about the simultaneous values of all three components. A single instrument can measure only one spin component, and in doing so it generally alters the values of the components. Hence in order to learn the values of three components three measurements would have to be made in succession. By the time the particle emerged from the third instrument it would no longer have the same spin components it had when it entered the first instrument.

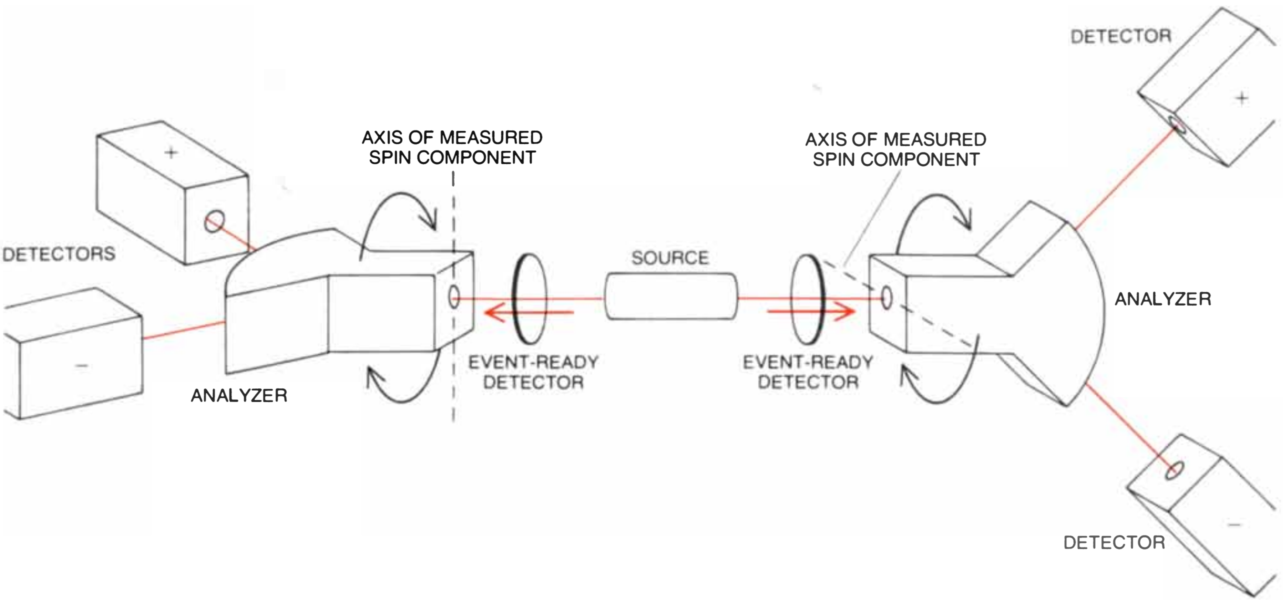
Although no instrument can measure more than one spin component at a time, a device can be built that is capable of being adjusted to measure the spin component along any one of three arbitrarily chosen axes. I shall designate these axes *A*, *B* and *C* and note the results of experiments as follows. If the spin component along axis *A* is found to be plus, it is labeled A^+; if the component along axis *B* is minus, it is given as B^-, and so on. The physicist can now prepare a large batch of protons in the singlet state. He finds that if he measures component *A* for both protons in each pair, some protons are A^+ and others are A^-, but whenever one member of a pair is A^+, the other member is always A^-. If he decides instead to measure component *B*, he observes the same negative correlation: whenever one proton is B^+, its singlet partner is B^-. Similarly, a C^+ proton is invariably accompanied by a C^- one. These results hold no matter how the axes *A*, *B* and *C* are oriented.

It is important to emphasize that in these experiments no proton is submitted to a measurement of more than one spin component. Nevertheless, if the physicist accepts the three premises of local realistic theories, he can draw conclusions from these findings about the values of all three components, following an argument much like that of the hypothetical psychologist. Considering a fresh batch of proton pairs in the singlet state on which no spin measurement has yet been made (and perhaps on which no such measurement will ever be made), he can infer that in every pair one proton has the property A^+ and the other has the property A^-. Similarly, he can conclude that in every pair one proton has the property B^+ and one B^- and one has the property C^+ and one C^-.

These conclusions require a subtle but important extension of the meaning assigned to a notation such as A^+. Whereas previously A^+ was merely one possible outcome of a measurement made on a particle, it is converted by this argument into an attribute of the particle itself. To be explicit, if some unmeasured proton has the property that a measurement along the axis *A* would give the definite result A^+, then that proton is said to have the property A^+. In other words, the physicist has been led to the conclusion that both protons in each pair have definite spin components at all times. The components may be unknown, since the physicist cannot say which proton in a pair has the property A^+ and which has the property A^- until a measurement along axis *A* has been made, but he can argue from the premises of local realistic theories that the values are quite definite even in the absence of any measurements. This view is contrary to the conventional in-



SECOND STAGE OF THE PROOF extrapolates from the case of single particles for which two properties are known to that of pairs of particles, each particle of which is tested for one property. The pairs are created in such a way that there is always a strict negative correlation for any property considered separately, that is, if one particle in a pair has the property A^+, the other must have the property A^-. Because of this correlation, if one particle in a pair is found to be A^+ and the other is found to be B^+, it is possible to deduce both properties of both particles. The doubly positive test result can arise only if one particle has the two properties A^+B^- and the other has the properties A^-B^+. Hence the number of such doubly positive test results, which can be designated $n[A^+B^+]$, must be proportional to the total number of particles with the properties A^+B^- and A^-B^+. Similar proportionalities can be derived for the number of doubly positive results observed when pairs of particles are tested for properties *A* and *C* and for properties *B* and *C*; these are the quantities $n[A^+C^+]$ and $n[B^+C^+]$. The constant of proportionality depends only on the number of pairs submitted to each set of tests and on the total number of pairs, and so the constant is the same in all three cases. It follows that the three ratios of the number of doubly positive test results to the number of individual particles that can give rise to those results must also be equal. A relation has already been demonstrated between the numbers of individual particles with the indicated properties; it is the inequality proved in the illustration on page 162. If that inequality is to hold, there must be a similar inequality between the numbers of doubly positive test results. This is the Bell inequality. The proof is valid only if the three premises of local realistic theories are assumed to be valid.



THOUGHT EXPERIMENT would test the Bell inequality by measuring the components of the spin of protons or other elementary particles. A spin component is a projection along some axis of the proton's intrinsic angular momentum; each component can have only two possible values, which can be designated plus and minus. The experiment, which assumes the availability of perfect instruments, would have a source where pairs of protons are brought together in a quantum-mechanical configuration called the singlet state. The pairs would then be broken up, and the protons would fly apart in opposite directions. "Event-ready" detectors would issue a signal whenever a suitable pair of protons had been emitted. Each proton would then enter an analyzer, where it would be deflected to one of two detectors depending on the value of its spin component along the axis defined by the analyzer. If the analyzers were set to measure the spin components along the same axis, a strict negative correlation would be observed. If one analyzer were rotated, so that they measured different components, local realistic theories predict that the correlation observed would be no greater than that allowed by the Bell inequality regardless of what the angle between the analyzers was. Quantum mechanics predicts a violation of the Bell inequality for some angles.

terpretation of quantum mechanics, but it is not contradicted by any fact that has yet been introduced.

The strict negative correlation for protons in the singlet state is expected only when the same spin component is measured on both protons. What happens when the instruments are set to measure different components? To be precise, consider the following experiment. Pairs of protons are brought together in the singlet state by the same method employed in the earlier experiments and are allowed to separate under exactly the same conditions. Each proton is then tested for just one spin component, *A*, *B* or *C*, but which one of the components is measured in each case is determined entirely at random. Sometimes by coincidence the same component will be measured on both protons in a pair; those results are discarded, since they provide no new information. The remaining pairs must then be made up of either one proton tested along axis *A* and one tested along axis *B*, or one tested along axis *A* and one along axis *C*, or one along axis *B* and one along axis *C*. For the sake of brevity I shall refer to the pairs in each of these three populations as *AB*, *AC* and *BC*. A pair that on testing yields the results A^+ for one proton and B^+ for the other can be labeled an A^+B^+ pair. The number of such pairs observed can be represented by the notation $n[A^+B^+]$. Can any relation among these quantities be expected?

In 1964 John S. Bell of the European Organization for Nuclear Research (CERN) discovered such a relation. For any large sample of singlet proton pairs Bell showed that the tenets of local realistic theories impose a limit on the extent of correlation that can be expected when different spin components are measured. The limit is expressed in the form of an inequality, which is now called the Bell inequality. Given the experimental conditions described above, it states that the number of A^+B^+ pairs cannot exceed the sum of the number of A^+C^+ pairs and the number of B^+C^+ pairs. The inequality can be expressed in symbols as

$$n[A^+B^+] \leq n[A^+C^+] + n[B^+C^+].$$

Many similar inequalities could be constructed with the various symbols transposed or with the signs reversed. Because the directions along which the spin components are defined were chosen arbitrarily, all such formulations are interchangeable, and I shall discuss only this one.

The Bell inequality can be proved, within the context of local realistic theories, through a straightforward argument in the mathematical theory of sets. It is convenient to begin with an assumption contrary to fact: that some means exist for independently measuring two components of the spin of a single particle. Suppose this impossible instrument has revealed that a particular proton has the spin components A^+ and B^-. The third component, *C*, has not been measured, but it can have only one of two values, plus or minus; hence the measured proton must be a member of one of two sets of protons, either the set with spin components $A^+B^-C^+$ or the set with components $A^+B^-C^-$. There are no other possibilities.

If many protons with the spin components A^+B^- are detected, one can write an equation about their number:

$$N(A^+B^-) = N(A^+B^-C^+) + N(A^+B^-C^-).$$

In order to avoid confusion the symbol $N(A^+B^-)$ has been employed to represent the number of individual protons with the two spin components A^+ and B^-; the symbol $n[A^+B^-]$ gives the number of proton pairs in which one particle has the component A^+ and the other has the component B^-. The equation states the obvious fact that when a set of particles is divided into two subsets, the total number of particles in the original set must be equal to the sum of the numbers in the subsets.

The protons found to have the spin components A^+C^- can be analyzed exactly the same way. Every such proton must be a member either of the set $A^+B^+C^-$ or of the set $A^+B^-C^-$, and the total number $N(A^+C^-)$ must be equal to the sum $N(A^+B^+C^-) + N(A^+B^-C^-)$.

bulletin highlights the new developments.

New material replaces old material in the living text, so that the information is there — up-to-date, at your fingertips.

A CME program of eight patient management problems offered over a 12-month period is available at no extra cost. Authorities from first-line institutions create the problems under the direction of Stanford's Office of Postgraduate Medical Education and the Educational Testing Service.

As an organization accredited for continuing medical education, the Stanford University School of Medicine designates this continuing medical education activity as meeting the criteria for 32 credit hours in Category 1 for Educational Materials for the Physician's Recognition Award of the American Medical Association provided it has been completed according to instructions.

This program is approved by the American Academy of Family Physicians for 32 Elective Hours of CME credit.

Trial Offer

Please send us the card and you will receive over a period of 60 days at no cost, the two-volume text and two monthly updates. You may also take a CME test for credit. At the end of 60 days, if you decide to continue the subscription, we will bill you for $150 for the full 12 months; otherwise just return the volumes to us.

We urge you to subscribe now. We anticipate a price increase in January 1980 to $218 for the first year and $165 for the annual renewal as of January 1981. If you subscribe in 1979 for $150, you will be able to renew next year at $150, a saving of $83 over two years.

So please mail the card today and let us take the hassle out of keeping up.

SCIENTIFIC AMERICAN MEDICINE

415 Madison Avenue, New York, N.Y. 10017

Please enroll me as a subscriber to SCIENTIFIC AMERICAN *Medicine*. On receipt of this coupon you will send me the advanced two-volume text described in your announcement and update it regularly by sending me new monthly subsections. I understand that the annual subscription of $150 for SCIENTIFIC AMERICAN *Medicine* is tax-deductible. If I am not entirely satisfied, I may cancel at any time during the first 60 days, returning all materials for a complete refund.

☐ Please enter my subscription for SCIENTIFIC AMERICAN *Medicine*

☐ I shall also enroll in the CME Program

☐ I enclose a check made out to SCIENTIFIC AMERICAN *Medicine* for $150

☐ Please bill me

Please add sales tax for California, Illinois, Massachusetts, Michigan, Ohio and New York

Name ____________________

MD Specialty ____________________

Address ____________________

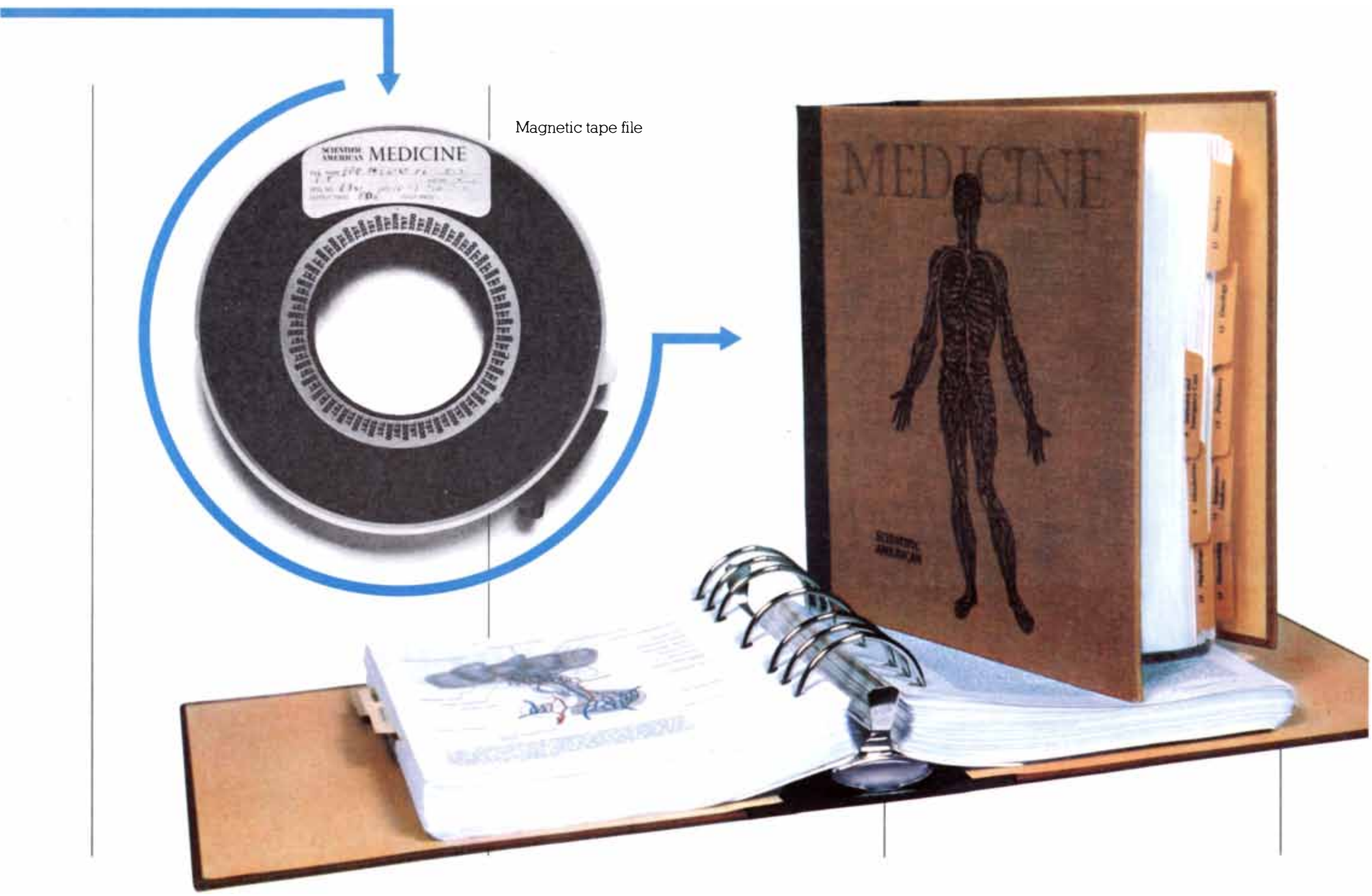
City ____________________

State ____________________ Zip ________

Signature ____________________

Subscribers outside of the U.S. and the possessions will be charged extra for shipping and handling; updates will be sent by surface routes unless air mail delivery is requested.

1E



Magnetic tape file

A further step can now be taken. If the number of protons $N(A^+C^-)$ is equal to $N(A^+B^+C^-) + N(A^+B^-C^-)$, then it must be greater than or at least equal to $N(A^+B^-C^-)$. (The two sets will be equal if the B components of all the particles' spins happen to be minus, so that the subset $(A^+B^+C^-)$ is empty; otherwise $N(A^+C^-)$ will be larger. In other words, a part of the whole cannot be greater than the whole.) The same reasoning can be applied once again to prove that the number of protons with spin components B^-C^+ must be equal to the sum $N(A^+B^-C^+) + N(A^-B^-C^+)$ and hence that $N(B^-C^+)$ must be greater than or equal to $N(A^+B^-C^+)$.

Consider again the first equation derived above:

$$N(A^+B^-) = N(A^+B^-C^+) + N(A^+B^-C^-).$$

It has just been demonstrated that $N(B^-C^+)$ is greater than or at least equal to $N(A^+B^-C^+)$, which is the first term on the right side of the equation. It has also been shown that $N(A^+C^-)$ is greater than or equal to $N(A^+B^-C^-)$, which is the second term on the right side of the equation. It is therefore permissible to make the appropriate substitutions in the equation, changing the equals sign to one signifying "less than or equal to." The result is the inequality

$$N(A^+B^-) \leq N(A^+C^-) + N(B^-C^+).$$

Although this inequality is hereby formally derived, it cannot be tested directly by experiment because no instrument can independently measure two spin components of a single proton. The experiments under consideration, however, are carried out not on individual protons but on correlated pairs of them, and there is no need to make such impossible measurements. Suppose one proton in a pair is submitted to a measurement of its spin component along the A axis and is found to have the value A^+. No other measurements are carried out on this particle, but its singlet partner is tested for the component along the B axis and the result is found to be B^+. The latter measurement, which might be made at a distant site after the protons have been moving apart for some time, conveys additional information about the state of the first proton. To be explicit, the existence of a strict negative correlation implies that the first proton, which is already known by direct measurement to have the spin component A^+, must also have the component B^-.

By this means the observation of a pair of protons one of which has the spin component A^+ and the other the component B^+ can be employed as a signal indicating the existence of a single proton with the components A^+B^-. Furthermore, it can be demonstrated by a statistical argument that $n[A^+B^+]$, the number of such doubly positive pairs, must be proportional to $N(A^+B^-)$, the number of individual protons with the spin components A^+B^-. In the same way $n[A^+C^+]$ must be proportional to $N(A^+C^-)$ and $n[B^+C^+]$ must be proportional to $N(B^-C^+)$. The constant of proportionality in all three cases is the same. For single protons each of which is subjected to an imaginary double measurement an inequality has already been proved, showing that $N(A^+B^-)$ can be no greater than the sum of two terms: $N(A^+C^-) + N(B^-C^+)$. It is now possible to replace each of these unmeasurable quantities by the corresponding numbers of doubly positive proton pairs. The resulting expression is

$$n[A^+B^+] \leq n[A^+C^+] + n[B^+C^+].$$

This is the Bell inequality.

Of course the inequality is proved by this argument only if the three premises of local realistic theories are considered valid. Indeed, it is here that the premises have their most important application and ultimately their most questionable one. If the premises are granted, at least for the sake of argument, it should be clear that the Bell inequality must be satisfied. Moreover, the orientation of the axes A, B and C has nowhere been specified, so that the inequality should be valid regardless of what axes are chosen. The only possible violation of the inequality would result from a statistical fluke, where many particles with the spin components A^+ and B^+ happened to appear through random coincidence. The probability of such a coincidence approaches zero as the number of particles tested increases.

The Bell inequality constitutes an explicit prediction of the outcome of an experiment. The rules of quantum mechanics can be employed to predict the results of the same experiment. I shall not give the details of how the prediction is derived from the mathematical formalism of the quantum theory; it can be stated, however, that the procedure is completely explicit and is objective in the sense that anyone applying the rules correctly will get the same result. Surprisingly, the predictions of quantum mechanics differ from those of the local realistic theories. In particular, quantum mechanics predicts that for some choices of the axes A, B and C the Bell inequality is violated, so that there are more A^+B^+ pairs of protons than there are A^+C^+ and B^+C^+ pairs combined. Thus local realistic theories and quantum mechanics are in direct conflict.

The conflict raises two questions. First, what are the experimental facts of the situation? Is the Bell inequality satisfied or is it violated? Whatever the outcome of an experimental test there must be a flaw of some kind either in the rules of quantum mechanics or in local realistic theories. The second question therefore is: What premise underlying the refuted theory is at fault?

The thought experiment proposed in 1935 by Einstein, Podolsky and Rosen called for measurements of the position and momentum of particles. The experiment on spin components of protons was first discussed in 1952 by David Bohm of Birkbeck College in London, but still in the context of a thought experiment. It was not until 1969, after Bell had introduced his inequality, that real experiments exploring these questions were contemplated. The feasibility of such experiments was discussed by John F. Clauser of the University of California at Berkeley, R. A. Holt of the University of Western Ontario and Michael A. Horne and Abner Shimony of Boston University. They found that for a practical experiment the Bell inequality would have to be generalized somewhat, but a meaningful test of the alternative theories would still be possible.

The technical difficulty of the experiments should not pass unmentioned. In a thought experiment both protons of every pair always reach the instruments and the instruments themselves always yield an unambiguous measurement of the spin component along the chosen axis. Real apparatus cannot reproduce these results. The detectors are never perfectly efficient: many protons are simply not registered at all. Because of the imperfections of the instruments the number of protons counted in each category cannot be interpreted directly; instead an allowance must be made for the inefficiency of the detectors, which adds to the uncertainty of the results.

Of seven experiments reported since 1971, six have not concerned measurements of the spin components of protons but have instead measured the polarization of photons: the quanta of electromagnetic radiation. Polarization is the property of a photon that corresponds to the spin of a material particle. In one series of experiments atoms of a particular element and isotope were raised to an excited state by the absorption of laser light and then allowed to return to their original energy level in two steps. At each step a photon with a characteristic energy or wavelength was emitted. The photons moved off in opposite directions, and they had opposite polarizations. In other words, if the polarization of both photons was measured along any single direction, a strict negative correlation was observed.

The differences between ideal instruments and real ones are quite plain in these experiments. There is no single device that can intercept a photon and report directly on its polarization. Instead two devices are necessary, a filter and a detector. The filter is designed to allow the passage of those photons that have the selected polarization and to stop or deflect all others; the detector counts the number of photons that pass through the filter. Neither of these components is perfect, so that the failure to register a photon does not necessarily mean that it had the wrong polarization.

Experiments have also been done on the polarization of gamma rays, which are high-energy photons. The gamma rays were created by the mutual annihilation of electrons and their antiparticles, positrons. Such an annihilation gives rise to two gamma rays, which are emitted in opposite directions and have opposite polarization. The experiments are therefore formally equivalent to the atomic ones, but the apparatus required is quite different. In general detectors are more efficient for high-energy photons, but polarization filters are more efficient for low-energy ones.

One experiment has measured the correlations of spin components of protons and therefore closely resembles the original thought experiment. The pairs of protons are created by injecting protons of comparatively low energy into a target made up partly of hydrogen atoms. The nucleus of a hydrogen atom consists of a single proton. When an incident proton strikes a hydrogen nucle-

EXPERIMENT	DATE	PARTICLES STUDIED	RESULTS
Stuart J. Freedman and John F. Clauser, University of California at Berkeley	1972	Low-energy photons emitted during transitions in calcium atoms.	In agreement with quantum mechanics.
R. A. Holt and F. M. Pipkin, Harvard University	1973	Low-energy photons emitted during transitions in atoms of mercury 198.	In agreement with Bell inequality.
John F. Clauser, University of California at Berkeley	1976	Low-energy photons emitted during transitions in atoms of mercury 202.	In agreement with quantum mechanics.
Edward S. Fry and Randall C. Thompson, Texas A. & M. University	1976	Low-energy photons emitted during transitions in atoms of mercury 200.	In agreement with quantum mechanics.
G. Faraci, S. Gutkowski, S. Notarrigo and A. R. Pennisi, University of Catania	1974	High-energy photons (gamma rays) from annihilation of electrons and positrons.	In agreement with Bell inequality.
L. Kasday, J. Ullman and C. S. Wu, Columbia University	1975	High-energy photons (gamma rays) from annihilation of electrons and positrons.	In agreement with quantum mechanics.
M. Lamehi-Rachti and W. Mittig, Saclay Nuclear Research Center	1976	Pairs of protons in the singlet state.	In agreement with quantum mechanics.

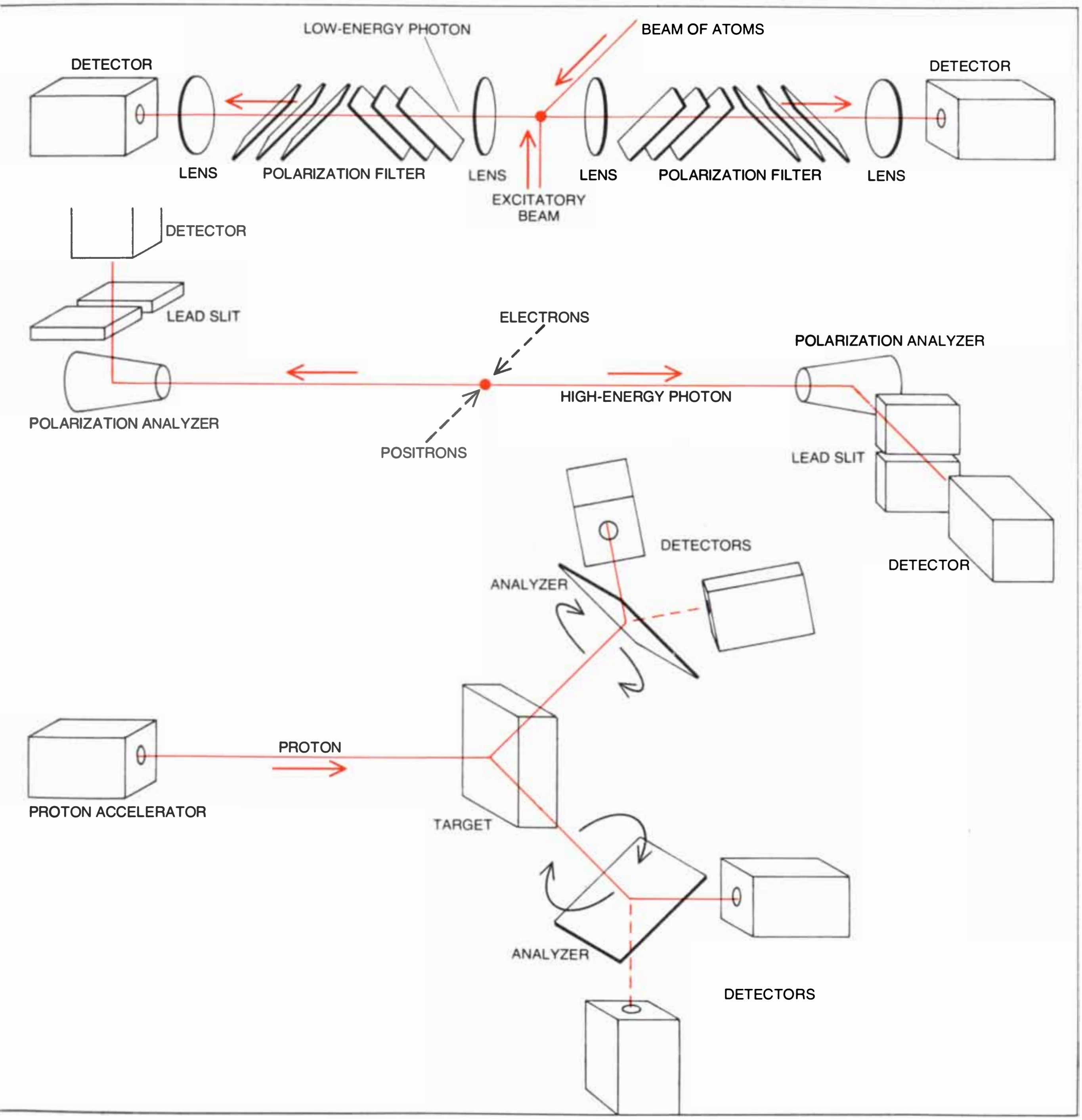
REAL TESTS OF THE BELL INEQUALITY have been carried out by seven groups of investigators. Only one of the experiments measured the spin components of protons; the others studied the polarization of photons, or quanta of electromagnetic radiation. In four experiments pairs of low-energy photons with opposite polarization were emitted by atoms that had been raised to an excited state. Pairs of oppositely polarized gamma rays, or high-energy photons, were created in two other experiments by the mutual annihilation of electrons and their antiparticles, positrons. In the remaining experiment protons from a particle accelerator struck a target made up partly of hydrogen; the accelerated protons and the hydrogen nuclei formed pairs in the singlet state. Five of the experiments gave results in violation of the Bell inequality and in agreement with quantum mechanics. That the Bell inequality is violated is now generally accepted. The cause of the discrepancy in the results of the other two experiments is uncertain.

us, the two protons interact briefly and enter the singlet state. Both then leave the target, sharing the momentum of the incident proton, but if they are undisturbed, they remain in the singlet state. Preliminary measurements of the same spin component on both protons give opposite results.

The instruments for an experiment with proton pairs again consist of filters and detectors. In the one experiment that has been completed the filter was a carbon foil, which scattered each proton into one of two detectors depending on the value of the measured component.

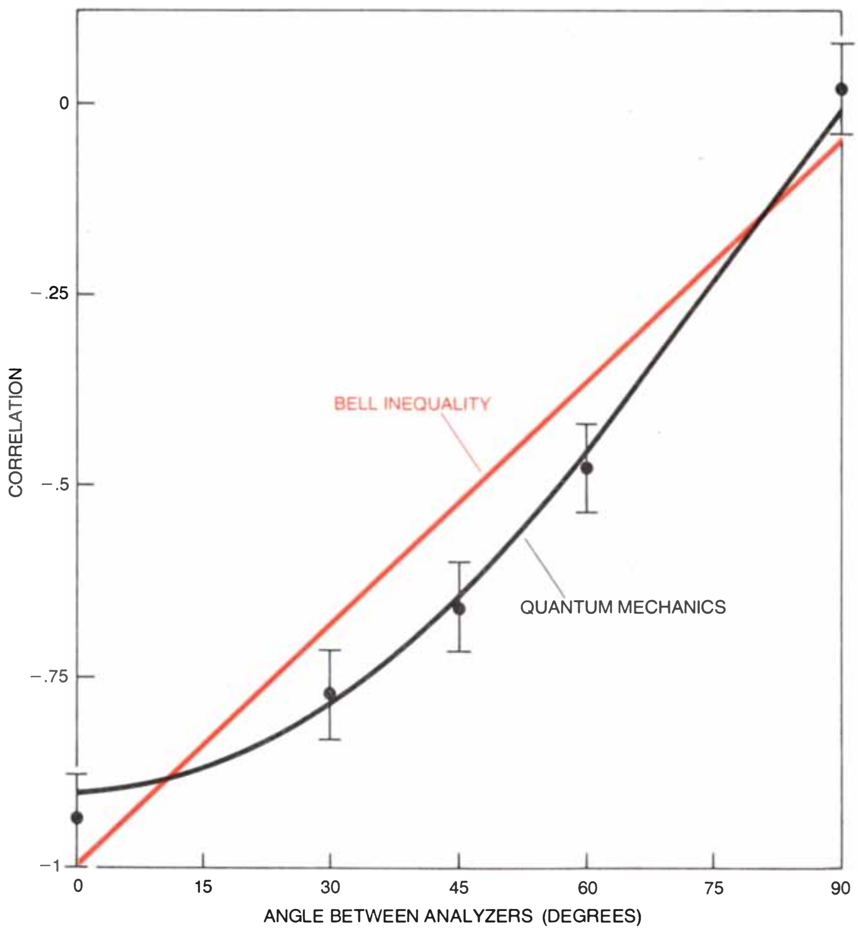
Regardless of what particles are being studied, the experiment consists of three series of double measurements. Three axes, *A*, *B* and *C*, are selected; in general the angles between them are set to the values where the maximum discrepancy between quantum mechanics and local realistic theories is expected. One filter is then set to admit particles with the polarization or spin component A^+ and the other is set to pass particles with the component B^+. After a large enough sample of particles has been recorded in this configuration the filters are rotated to measure the components along axes *A* and *C* and further data are recorded. Finally the filters are reoriented again to axes *B* and *C*. The coincidences recorded in each configuration are counted and corrections are made for the inefficiency of the apparatus. It is then a matter of simple addition to compare the results with the Bell inequality.

Of the seven completed experiments five endorse the predictions of quantum mechanics, that is, they indicate a violation of the Bell inequality for some choices of the axes *A*, *B* and *C*. The other two give correlations no greater than those allowed by the Bell inequality and



therefore support local realistic theories. The score is thus five to two in favor of quantum mechanics. Actually the support for quantum mechanics is much stronger than this ratio would seem to imply. One reason for attributing greater credibility to the five experiments that violate the Bell inequality is that they represent a larger sample of data and are therefore statistically more significant. Some of those experiments were done after the two anomalous results were reported and included refinements in the instrumentation designed explicitly to avoid any biases that might account for the two discrepant results. Clauser and Shimony have pointed out that there is also an epistemological justification for disregarding the two experiments that are in disagreement with the majority. Quantum mechanics predicts a larger correlation between events and local realistic theories predict a smaller one. A great variety of systematic flaws in the design of an experiment could destroy the evidence of a real correlation, yielding results within the limit set by the Bell inequality. On the other hand, it is hard to imagine an experimental error that could create a false correlation in five independent experiments. What is more, the results of those experiments not only violate the Bell inequality but also violate it precisely as quantum mechanics predicts. For the results of the five experiments to be produced by random coincidence would require an extraordinary statistical fluke that is not credible given the number of particles that have now been detected.

Further tests of the Bell inequality are under consideration, and at least one additional experiment is already in preparation. Most physicists concerned with these problems, however, have substantial confidence, based on the five consistent results, that the issue has already been decided. For some choices of the axes *A, B* and *C* the Bell inequality is violated in nature, and local realistic theories are therefore false.



RESULTS OF AN EXPERIMENTAL TEST of the Bell inequality show that it is clearly violated. The experiment is the one that employed pairs of protons in the singlet state, which was carried out by M. Lamehi-Rachti and W. Mittig of the Saclay Nuclear Research Center in France. The negative correlation between the values of different spin components is given as a function of the angle between the settings of the two analyzers. A correlation of −1 would indicate that the components invariably had opposite values. The Bell inequality states that the correlation at any angle must be on or above the colored line. The observed correlations at 30, 45 and 60 degrees are below the line. The results not only violate the Bell inequality but also are in good agreement with the predictions of quantum mechanics, which fact adds to their credibility. The violation of the Bell inequality implies that at least one of the three premises of local realistic theories must be false; Einstein separability is considered the most plausible candidate.

If it can be considered as having been demonstrated that local realistic theories are in error, which of the three premises underlying those theories is to blame? A first step in answering this question should be to make sure no additional assumptions were made in formulating the experimental test.

As it happens, at least one subsidiary assumption was needed. Because of the limitations of practical instruments, it was necessary to generalize the Bell inequality slightly, and that generalization must be assumed to be valid; it cannot be proved. It seems most unlikely, however, that this circumstance could alter the phenomena in such a way that the results of the experiments not only would violate the Bell inequality but also would be consistent with the predictions of quantum mechanics. In any case it is possible more refined experiments will test the inequality without the generalization. Because the subsidiary assumption is susceptible to an experimental test it seems less fundamental than the other three, and so it will not be considered further here.

Another area that might be scrutinized for unacknowledged assumptions is the proof of the Bell inequality. Indeed, it seems the proof does depend on the assumed validity of ordinary, two-valued logic, where a proposition must be either true or false and a spin component must be either plus or minus. Some interpretations of quantum mechanics have introduced the idea of a many-valued logic, but those proposals have nothing to do with the reasoning applied in this proof. Indeed, in the context of the proof it is difficult even to conceive of an alternative to two-valued logic. Unless such a system is formulated it seems best to pass over the problem.

The entire series of experiments founded on the ideas of Einstein, Podolsky and Rosen is sometimes regarded as merely a test of hidden-parameter theories. The experiments do indeed test those theories, but it should be emphasized that the existence of hidden parameters is not an additional premise of local realistic theories. On the contrary, the existence of parameters specifying the deterministic properties of a particle was derived from the three original assumptions. Remember that the psychologist did not assume that his invented test measured any real attribute of the tested subjects; instead he deduced the existence of such an attribute after observing a strict correlation. In the same way the existence of hidden parameters was derived from the negative correlation detected when a single spin component was measured on pairs of protons in the singlet state.

It is probably not possible to prove rigorously that no other supplementary assumptions enter into the argument supporting the local realistic theories. The chain of reasoning is simple enough, however, that if other assumptions are implicit in it, they should be easily recognized. None has yet been pointed out. It therefore seems that attention must be focused on the three premises of realism, the free use of induction and Einstein separability.

Of the three premises realism is the most fundamental. Realism can be stated formally as the belief that a mere description of data is not all that should be required of a theory. Even an empirical rule for predicting the patterns of future measurements is not enough. The mind demands something more: not necessarily determinism—there is nothing intrinsically irrational about randomness—but at least objective explanations of observed regularities, or in other words causes. Underlying this demand is the intuitive notion that the world outside the self is real and has at least some properties that exist independently of human consciousness.

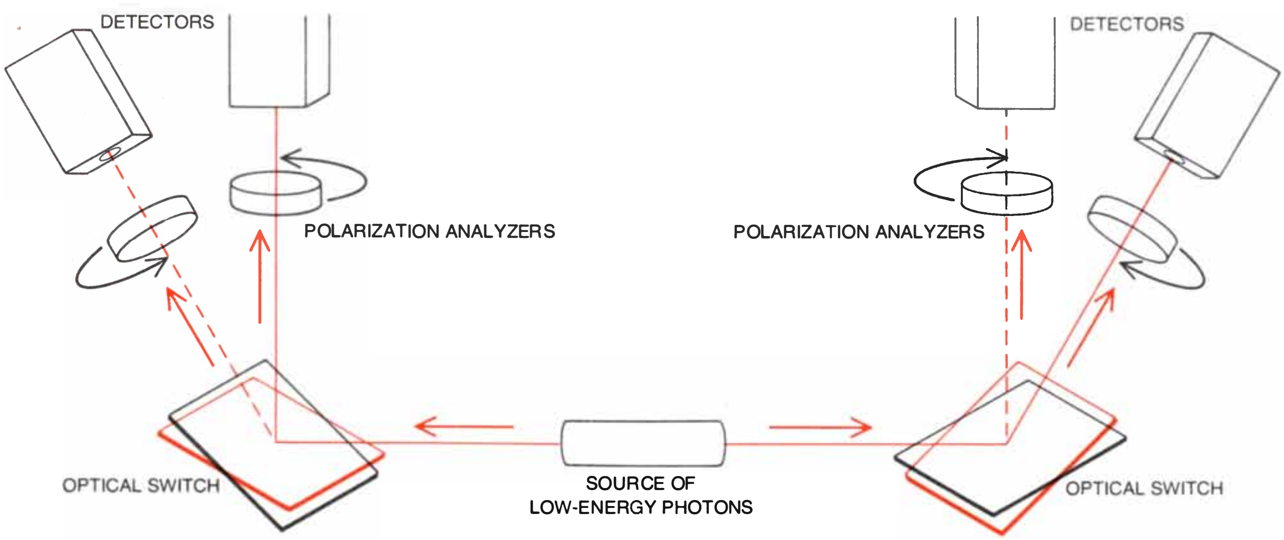
A number of philosophers, who can collectively be called positivists, have rejected the realistic viewpoint. The positivists do not assert that the world external to the mind does not exist; they merely dismiss as meaningless any statement about an external reality that does not refer directly to sensory impressions. In the 20th century some radical positivists have had an appreciable, if indirect, influence on the thinking of theoretical physicists.

The sense of paradox induced by the finding that the Bell inequality is violated can certainly be alleviated by adopting a positivist attitude, and such a course of action was first proposed long ago. When all the consequences of abandoning realism are considered, however, it is too great a renunciation to have much appeal. In the context of this experiment positivism asserts that it would be meaningless to attribute anything resembling a definite spin component to a particle before the component is measured; that the only quantity with any verifiable reality is the observation itself, the sensory impression; and that the psychologist's demand for an objective explanation of the remarkable correlation he observes should ultimately be rejected. If this refusal to seek underlying causes of observed regularities is applied consistently, it trivializes the entire scientific enterprise. Science is reduced to a set of recipes for predicting future observations from a knowledge of past ones. Any notion of science as "the study of nature" is impossible; nature is a phantom. One can imagine a physics grounded on positivist principles that would predict all possible correlations of events and still leave the world totally incomprehensible. Given the extreme consequences of abolishing realism, one is inclined to cling to this first premise.

Realism enters the argument supporting local realistic theories at another point: it is the justification for postulating the free use of induction. It is induction that enabled the physicist to extrapolate from a series of observed negative correlations to the conclusion that any two protons in the singlet state have opposite values of any single spin component, even if none of the components is measured. The extrapolation was an essential step in the proof of the Bell inequality, but it is clearly insupportable if the concept of unmeasured properties has no meaning.

This use of induction might be regarded by some as a weak link in the chain of argument. Shortly after the paper by Einstein, Podolsky and Rosen appeared, Niels Bohr published a reply in which he defended the completeness of the quantum-mechanical description of nature; the basis of his criticism was that Einstein's use of induction was unwarranted. Bohr's reply is a central document in what has come to be known as the Copenhagen interpretation of quantum mechanics. His reasoning amounts to an argument that a particle and an instrument adjusted to make a specific measurement on it constitute in some respects a single system, which would be altered in an essential way if the setting of the instrument were changed. For this reason it is not allowable to make any inferences about the state of a particle without specifying at the same time the settings of the instruments that will interact with the particle.

Bohr's views have been widely influential, and in a sense rightly so; after all, the recent work under discussion here has shown that in these matters he was



EINSTEIN SEPARABILITY will be tested rigorously in an experiment now being prepared by Alain Aspect of the Optics Institute of the University of Paris. Earlier experiments tested only the less restrictive separability principle: the settings of the analyzers were determined well in advance, so that some influence of one measurement could be communicated (by an unknown mechanism) to the other measurement at a speed well below the velocity of light. This possible explanation of the observed correlation is extremely unlikely, but it would be excluded entirely if the settings of the analyzers were changed so quickly that a signal moving no faster than light could not pass from one detector to the other in time to influence the result of the second measurement. In Aspect's experiment, which will measure the polarization of low-energy photons, this condition will be met. Two sets of analyzers and detectors will be provided for each photon, and the analyzers will measure different components. A fast optical switch will determine which analyzer the photon enters only when it is too late for the decision to influence the other measurement (assuming that the hypothetical influence propagates no faster than light). The switch is shown as a moving mirror; actually the switching will be accomplished by ultrasonic waves on the surface of a crystal.

closer to the truth than Einstein was. Nevertheless, when Bohr's ideas are considered in their essence, they are subject to objections much like those that were raised against a retreat to positivism. Because realism provides the ultimate rationale for the free use of induction, it can be argued that Bohr was not a realist, or at least not a consistent one. Any explanation of the distant-correlation experiments that relies on Bohr's reply to Einstein, Podolsky and Rosen may turn out to be inconsistent with even a moderate version of realism.

If realism and the free use of induction are to be retained, the violation of the Bell inequality can be explained only by giving up the assumption of Einstein separability. In the psychological experiment separability was understood to imply that the husbands and the wives, once they were separated, could not communicate with each other. In the physics experiment the separability assumption expressed the intuitively reasonable idea that the spin components of one proton have no influence over those of the other proton, provided the two particles are far apart. The more restrictive assumption of Einstein separability forbids such an influence only if it would have to propagate with a speed greater than the speed of light. As I have shown, this assumption must now be regarded as highly questionable.

Before considering the consequences of this conclusion it should be pointed out that none of the experiments completed so far has rigorously tested the assumption of Einstein separability. In those experiments the settings of the instruments were determined well in advance (on the time scale of particle physics). Therefore the setting of one instrument could conceivably affect events observed at the other instrument, or it could modify hidden parameters at the source of the proton pairs; in either case there would be no need for the influence to travel faster than light. An experiment with instruments whose setting can be changed rapidly could exclude this possibility. The decision to measure a certain spin component with one detector would not be made until it was too late for any influence of that decision to reach the other instrument or the source, even at the speed of light, in time to alter the outcome of the second measurement. Such an experiment is now being done by Alain Aspect of the Optics Institute of the University of Paris.

Quite apart from the question of how fast a hypothetical influence could travel from one instrument to another, the influence itself seems extremely implausible. It would be required to alter the distant observations in precisely the manner needed to produce the observed violation of the Bell inequality. Hence it seems best to search for some other explanation and to assume, pending the results of Aspect's experiment, that if ordinary separability is violated, Einstein separability will be violated too.

I have discussed a pair of protons as if they were independent entities that come together in the target and then move apart again. They can also be regarded as the elements of a single physical system that is created during the first interaction and becomes progressively more extended in space until it is disrupted by the first measurement. With respect to separability these descriptions are equivalent. In each case a violation of Einstein separability requires instantaneous action at a distance, either between independent systems or within a single extended system.

Must the principle of the finite propagation of signals therefore be abandoned? To that question no rash answer should be given. The principle was introduced as a premise of the theory of relativity, which cannot be made consistent without it. Moreover, signals that outrace light give rise to bizarre paradoxes of causality in which observers in some frames of reference find that one event is "caused" by another that has not yet happened. It turns out, however, that the instantaneous influences that seem to be at work in the distant-correlation experiments do not require such a drastic revision of accepted ideas. It seems quite certain these influences could not be employed to transmit any "useful" information, such as orders or instructions. No event that causes another event can be linked to it through this mechanism; the instantaneous influences can pass only between events that are related by a common cause. Hence the concept of a signal could be redefined in such a way that only those means of communication that transmit useful information would be called signals. The principle of the finite velocity of signals would then be preserved.

Even this solution impairs scientific realism to some extent. The basic law that signals cannot travel faster than light is demoted from a property of external reality to a feature of mere communicable human experience. Although this represents a step toward philosophical positivism, the concept of an independent or external reality can still be retained as a possible explanation of observed regularities in experiments. It is necessary, however, that the violation of Einstein separability be included as a property, albeit a well-hidden and counterintuitive property, of that independent reality. It should be noted in passing that Bohr's refutation of Einstein's argument for hidden parameters introduces an implicit violation of separability. It is founded on a strange indivisibility of the system of particles and the instruments of observation.

The argument that proceeds from an

observed correlation to the Bell inequality to the violation of Einstein separability is not particularly complicated, but it is indirect. Could the same result have been obtained in some more straightforward way? As it happens, it could not have been demonstrated without the Bell inequality, but it could have been suspected, and in fact it was. The suspicion arose from the fact that the wave function for a system of two or more particles is generally a nonlocal entity, which is considered to collapse suddenly or even instantaneously when a measurement is made. If the wave function is regarded as a kind of bizarre real jelly, the instantaneous collapse obviously violates Einstein separability. This naive argument was never taken very seriously, however, because the conventional interpretation of quantum mechanics does not identify the wave function of a system with whatever is meant by the reality of the system. Bohr, for example, considered the wave function a mere tool for doing calculations. Besides, the wave function for a system of several particles describes them only in an approximation that ignores the theory of relativity, and so its structure hardly seems a reliable argument against Einstein separability. For these reasons it was possible until a few years ago to believe in an independent, external reality and simultaneously to regard Einstein separability as a completely general law bearing on that reality.

One conceivable response to the distant-correlation experiments is that their outcome is inconsequential. The experiments themselves might represent a rare and therefore interesting test of quantum-mechanical phenomena observed at long range, but the results are merely what was expected. They show that the theory is in agreement with experiment and so provide no new information. Such a reaction would be highly superficial. It is indeed true that the experiments, now that they have been completed, have turned out to have little to do with quantum mechanics. That does not make them trivial; rather, it indicates that their real bearing is elsewhere. A discovery that discredits a basic assumption about the structure of the world, an assumption long held and seldom questioned, is anything but trivial. It is a welcome illumination.

Most particles or aggregates of particles that are ordinarily regarded as separate objects have interacted at some time in the past with other objects. The violation of separability seems to imply that in some sense all these objects constitute an indivisible whole. Perhaps in such a world the concept of an independently existing reality can retain some meaning, but it will be an altered meaning and one remote from everyday experience.