The great tradition in philosophy, from Aristotle to Kant, was that philosophy legislated the methodology and foundations of science. It can be claimed that, in spite of the many centuries separating Aristotle and Kant, it is still true that the three most important foundational works on science were Aristotle's *Posterior Analytics*, with many points amplified in the *Physics* and the *Metaphysics*, Descartes' *Principles of Philosophy*, and at the other end of the period the very specific working out of the foundations of physics in Kant's *Metaphysical Foundations of Natural Science*, with the more general lines of argument being given in the *Critique of Pure Reason*. It is not difficult to trace the enormous impact of Kant on physics in the nineteenth century, especially German physics, and also psychology, even though Kant was skeptical of providing the kind of foundations for psychology he gave for physics.

A different kind of foundational effort was made by logical positivism. In this case the effort was more to say what was not science but bad metaphysics, rather than to lay down a detailed foundation for science itself. Certainly in the tradition of logical positivism there was nothing so close to the actual spirit of classical physics as is to be found in Kant's *Metaphysical Foundations of Natural Science*, or, earlier, in Descartes' *Principles*.

But those days are gone and done for. I am skeptical that we shall ever find a revival of the view that philosophy can seriously legislate the foundations of any science. Indeed, I shall even question as we examine the matter in more detail that there is a serious sense in which there should be the foundations of any of the major sciences. The enterprise of foundations, I want to claim, has become inevitably and irreducibly pluralistic in character. The analysis of certain problems or their solutions, because of their wide conceptual interest, has a foundational character. But there is not some epistomological or metaphysical view that can be used to organize in a definitive way classification of problems as foundational in nature. There is not some selected and small list of problems that are regarded as the central problems of the foundations of any one discipline. Of course, some physicists still talk this way, but the record speaks for itself: whenever one range of problems is solved that were regarded at one point as foundational and fundamental in an absolute sense, a new range of problems replaces them. I see no reason to be other than skeptical about the ultimate nature of the physical
universe being settled, whether we are concerned with the final version of the big bang or the final statement of the fundamental forces. In fact, to make a skeptical prediction, I think it likely that the inappropriateness of the detailed analysis of forces in Kant's *Metaphysical Foundations of Natural Science* will be matched by a corresponding datedness for the current views of the fundamental physical forces a hundred years hence.

The old theological drive for certainty and salvation is hard to control and I am sure that there will be continual attempts to put this or that scientific discipline on an "ultimate" foundational basis, but all that will result in practice is a partial solution of some interesting problems, which is a good outcome, or what is a bad outcome—the development of a new form of scholasticism irrelevant to current scientific work.

Let me give some examples to illustrate this general remark. Mathematicians have currently lost interest in foundations as classically conceived. The development of classical foundations has become a technically sophisticated and important subdiscipline but its philosophical role has nearly faded away. A different kind of example where foundational scrutiny is still actively involved in the main scientific developments is the intense Bayesian controversy in statistics. An example of still another sort is provided by quantum mechanics. Partly because the literature on physics is now so diverse and so large, but also because of the focus of much of the foundational literature, there are large parts of the foundational literature on quantum mechanics that are really only known to specialists. A good example would be the now quite extensive literature on quantum-mechanical logic. Another area of greater interest to physicists in general, but still a subject that has become too specialized to follow in detail, is the continuing controversy about the existence of hidden variables. The controversy about hidden variables continues to be an active area of interest, even to some experimental physicists, but it has to be regarded as a foundational subject, not as one of the most important areas of current research in physics.

I mention these various examples just to give a descriptive sense of the way in which foundational interests interact with a particular discipline. What I have to say is not meant to be evaluative but I also want to emphasize that what I have to say is not meant to be a permanent or static descriptive analysis. The proper attitude, it seems to me, is very much not only pluralistic but dynamic. The periods of great interest in foundations in a discipline as a whole are periods that wax and wane with particular features of the development of the discipline.

The kind of sweeping viewpoint that Aristotle or Kant tried to put forth aggressively in defense of the central role of philosophy is out of the question now. Current research in physics, for example, is too complicated, technical,
and diverse even for physicists to understand all the various subdisciplines. It is a hopeless task for philosophers to think of offering some kind of underpinnings for this vast intellectual enterprise. I simply pick physics as an example. This is certainly true for other disciplines as well. The disciplines are held together by a traditional conglomeration of ideas, which often become separated over time. There is for most scientific disciplines no serious unified sense of foundations even possible.

This may sound pessimistic and skeptical about any role philosophers may have. This is not my view. There is a role for philosophy in relation to the sciences. We are no longer Sunday's preachers for Monday's scientific workers, but we can participate in the scientific enterprise in a variety of constructive ways. Certain foundational problems will be solved better by philosophers than by anyone else. Other problems of great conceptual interest will really depend for their solution upon scientists deeply immersed in the discipline itself, but illumination of the conceptual significance of the solutions can be a proper philosophical role.

In the rest of this lecture I will try to illustrate these general ideas by considering three examples of scientific problems and results in a given area that have philosophical interest—indeed philosophical interest in relation to long-standing problems in the philosophy of science. But as should be evident from what I have already said, I do not mean to suggest that the three examples I have chosen lead to anything like a philosophical claim about science of the sort we associate with Aristotle or Kant.

The first example deals with randomness and determinism in classical physics, the second with hidden variables in quantum mechanics, and the third with the nature of visual space.

So, of my three examples, two are taken from physics and one from psychology. Examples from other disciplines could as easily have been selected but the selection of three problem areas had to be made not in terms of some metaphysical criterion of interest but in terms of problems I happen to know something about.

**Determinism and Randomness**

One of the great issues in the philosophy of science in the twentieth century has been the conflict between the deterministic features of classical physics and the development of probabilistic models of all kinds of natural phenomena, with randomness as a central feature of such models. Quantum mechanics, of course, in the view of many persons, has shown once for all that there exist significant natural phenomena that are in principle indeterministic. I have something more to say about quantum mechanics in my second example. What I want to challenge now in a decisive way is the conventional...
picture that classical mechanics is deterministic and therefore in no sense random. There are several ways of getting at the demonstration that this is a mistaken dichotomy, but I think the most striking example and, indeed, one of the most striking theorems in the entire history of classical mechanics arises from detailed consideration of a special case of the three-body problem, which is without doubt the most extensively studied problem in classical mechanics. The special case is this. There are two particles of equal mass $m_1$ and $m_2$ moving according to Newton's inverse-square law of gravitation in an elliptic orbit relative to their common center of mass, which is at rest. The third particle has a nearly negligible mass, so it does not affect the motion of the other two particles, but they affect its motion. This third particle is moving along a line perpendicular to the plane of motion of the first two particles and intersecting the plane at the center of their mass—let this be the $z$ axis. From symmetry considerations, we can see that the third particle will not move off the line. The restricted problem is to describe the motion of the third particle.

To obtain a differential equation in simple form, we normalize the unit of time so that the temporal period of rotation of the two masses in the $x, y$-plane is $2\pi$, we take the unit of length to be such that the gravitational constant is one, and finally $m_1 = m_2 = \frac{1}{2}$, so that $m_1 + m_2 = 1$. The force on particle $m_3$, the particle of interest, from the mass of particle 1 is:

$$F_1 = \frac{m_1}{z^2 + r^2} \cdot \frac{(-z, \overline{r})}{\sqrt{z^2 + r^2}},$$

where $r$ is the distance in the $x, y$-plane of particle 1 from the center of mass of the two-particle system $m_1$ and $m_2$, and this center is, of course, just the point $x = y = z = 0$. Note that $\frac{(z, r)}{\sqrt{z^2 + r^2}}$ is the unit vector of direction of the force $F_1$. Similarly,

$$F_2 = \frac{m_2}{z^2 + r^2} \cdot \frac{(-z, \overline{r})}{\sqrt{z^2 + r^2}}.$$

So, simplifying, we obtain as the ordinary differential equation of the third particle

$$\frac{d^2z}{dt^2} = -\frac{z}{(z^2 + r^2)^{3/2}}.$$

The analysis of this easily described situation is quite complicated and technical, but some of the results are simple to state in informal terms. Near the
escape velocity for the third particle—the velocity at which it leaves and does not periodically return—the periodic motion is very irregular. In particular, the following remarkable theorem can be proved. Let \( t_1, t_2, \ldots \) be the times at which the particle intersects the plane of motion of the other two particles. Let \( s_k \) be the largest integer equal to or less than the difference between \( t_{k+1} \) and \( t_k \) times a constant.\(^1\) Variation in the \( s_k \)'s obviously measures the irregularity in the periodic notion. The theorem, due to the Russian mathematicians Sitnikov (1960) and Alekseev (1969a,b), as formulated in Moser (1973), is this.

**Theorem 1.** Given that the eccentricity of the elliptic orbits is positive but not too large, there exists an integer, say \( \alpha \), such that any infinite sequence of terms \( s_k \) with \( s_k \geq \alpha \), corresponds to a solution of the deterministic differential equation governing the motion of the third particle.\(^2\)

A corollary about random sequences immediately follows. Let \( s \) be any random sequence of heads and tails—for this purpose we can use any of the several variant definitions—Church, Kolmogorov, Martin-Löf, etc. We pick two integers greater than \( \alpha \) to represent the random sequence—the lesser of the two representing heads, say, and the other tails. We then have:

**Corollary.** Any random sequence of heads and tails corresponds to a solution of the deterministic differential equation governing the motion of the third particle.

In other words, for each random sequence there exists a set of initial conditions that determines the corresponding solution. Notice that in essential ways the motion of the particle is completely unpredictable even though deterministic. This is a consequence at once of the associated sequence being random.

From a general philosophical standpoint, what this example suggests above all is that the classical dichotomy between deterministic and indeterministic phenomena is not really the one that has been the major worry. What we are in many contexts mainly concerned with is not determinism but prediction. What the theorem shows is that the real dichotomy is between determinism and prediction, not between determinism and randomness. In other words, we can have systems that are both deterministic and random, and we can also have systems that are deterministic but completely unpredictable in their behavior. In the present context it is not appropriate to attempt a detailed disentangling of the relationships between the four concepts of determinism, indeterminism, randomness, and predictability, but I hope that I have been able to suggest in these rather brief remarks that the
relationship is not that which is often claimed philosophically. There is another point to make in this connection that bears on my general thesis about the relation between philosophy and the sciences. In discussions of determinism, a well-known paper of Montague "Deterministic Theories" (1974) is often cited. Montague proves some useful general theorems about determinism in a setting that he formulates precisely for classical mechanics, but from a mathematical standpoint the proofs of the theorems are all quite simple, and from a physical standpoint no really interesting phenomena are treated. In contrast, I would say, by looking more deeply at results in a particular science, in this case mechanics, we are led to genuinely surprising results, as reflected in Theorem 1, whose proof demands the full resources of modern work in mechanics.

Bell’s Inequalities in Quantum Mechanics

Bell’s inequalities are formulated for measurements of quantum-mechanical spin of pairs of particles originally in the singlet state. A variety of specific experimental realizations has been given in the literature. Let $A$ and $A'$ be two possible orientations of apparatus I, and let $B$ and $B'$ be two possible orientations of apparatus II. Let the measurement of spin by either apparatus be 1 or -1, corresponding to spin $1/2$ or $-1/2$, respectively. By $E(AB)$, for example, we mean the expectation of the product of the two measurements of spin, with apparatus I having orientation $A$ and II having orientation $B$. By axial symmetry, we have $E(A) = E(A') = E(B) = E(B') = 0$, i.e., the expected spin for either apparatus is 0. Note that we now use the notation $A$, $A'$, $B$, and $B'$ for the random variables whose values are the results of spin measurements in the four positions of orientation. It is, on the other hand, a well-known result of quantum mechanics that the covariance (or correlation) term $E(AB)$ is $-\cos \theta(A, B)$, where $\theta(A, B)$ is the difference in angles of orientation $A$ and $B$. Again, by axial symmetry only the difference in the two orientations matters, not the actual values $A$ and $B$.

On the assumption that there is a hidden variable that renders the spin results conditionally independent, i.e., that there is a hidden variable $\lambda$ such that $E(AB|\lambda) = E(A|\lambda) E(B|\lambda)$, Bell (1964) derives the following inequalities:

\[-2 \leq E(AB) + E(AB') + E(A'B) - E(A'B') \leq 2,\]
\[-2 \leq E(AB) + E(AB') - E(A'B) + E(A'B') \leq 2,\]
\[-2 \leq E(AB) - E(AB') + E(A'B) + E(A'B') \leq 2,\]
\[-2 \leq -E(AB) + E(AB') + E(A'B) + E(A'B') \leq 2.\]

(This form of the inequalities is due to Clauser, Horne, Shimony, and Holt, 1969.)
The work described thus far falls in a rather standard way within physics, but the problem is of such general interest and connects to so many other issues in philosophy, that it is important to see how Bell’s inequalities can be pursued further in a way that does not really depend upon additional physical assumptions but on general matters of probability and logic.

The first step to mention is Fine’s (1982) proof that Bell’s inequalities hold for the four random variables $A, A', B,$ and $B'$, if and only if there exists a joint probability distribution of the four random variables compatible with the four given covariances. Note that it will be part of the joint distribution to fix the two covariances that are not determined by the experimental data, namely, the covariance of $A$ and $A'$, and the covariance of $B$ and $B'$.

Bell obtained the inequalities by reasoning from the existence of a hidden variable. It is also straightforward to show that a joint probability distribution compatible with the given covariances implies Bell’s inequalities. What is surprising and interesting about Fine’s result is that the inequalities are sufficient for a joint distribution. On the other hand, the result is mathematically special. For $N > 4$, satisfaction of Bell’s inequalities for every quadruple of the $N$ random variables is not a sufficient condition for existence of a joint distribution.

A second result, related in a more general way to this discussion, is an earlier theorem of Suppes and Zanotti (1981) that relates the existence of a hidden variable to the existence of a joint probability distribution. The theorem as originally stated by Zanotti and me assumed the random variables had only only two values but, as Paul Holland pointed out (Holland & Rosenbaum, 1986), the generalization to a finite number of values is immediate. So the theorem on the existence of a hidden variable or, as it is more generally called in the philosophical literature, a common cause is as follows:

**THEOREM 2.** Let $X_1, \ldots, X_n$ be finite-valued random variables. Then a necessary and sufficient condition that there is a random variable $\lambda$ such that $X_1, \ldots, X_n$ are conditionally independent given $\lambda$ is that there exists a joint probability distribution of $X_1, \ldots, X_n$.

In the statement of the theorem, $\lambda$ is of course what the physicists would call a hidden variable. What is philosophically interesting about this theorem is that if no restrictions, for example, physical assumptions about the nature of the hidden variable, are made, then always trivially we can find one for any phenomenon for which there exists a joint probability distribution. Moreover, we can find a hidden variable that is deterministically related to the phenomenological variables.

Of course, when a negative result is anticipated, as in the case of quantum mechanics, it is reasonable to put no conditions whatsoever on the nature
of the hidden variable, for then a negative result is as strong as possible. But what happens in the case of quantum mechanics is also clear. This reduces the problem of a hidden variable just to the question of a joint probability distribution's existing for given random variables. This is a question that arises, one might say, in a ubiquitous way in quantum mechanics; for example, in general the position and momentum of a particle do not have a joint distribution.

What Theorem 2 shows is that we have in the general case a complete reduction of the existence of a hidden variable to the existence of a joint probability distribution of the phenomenologically given random variables. Note that although the theorem is stated for finite-valued random variables, continuous distributions may be approximated arbitrarily well by such discrete distributions.

The next step is to look more carefully in a methodological way at what is involved in the existence or nonexistence of a joint distribution as data are collected in any particular empirical situation. When data are recorded for several random variables in what we might term the standard way, then there is no problem of the existence of a joint distribution. Without trying to define this standard approach in a general way, let me illustrate by a couple of vivid examples. Suppose we are concerned with the distribution of height and weight in the population of entering students in American universities in the fall of 1986. We record a large sample chosen with appropriate methodology of sampling, and as we observe each student we measure height and we measure weight. For each individual observed we put in our data records the height and the weight of the individual. It is an implicit assumption of such procedures that it does not really matter within a few moments which variable we measure first. So the measurement of one variable does not have any impact whatsoever on the measurement of the second. If we measure height first then our procedure for measuring height does not affect the outcome of the following weight measurement. This assumption about sequence in time is dependent upon the interval of time between the two measurements being quite short. If we measured height when the students were entering the university, and measured weight four years later, we would have a joint distribution if we identified appropriately each individual, but it would not in any sense be the joint distribution we had originally planned to study, namely, the "simultaneous" distribution of height and weight.

What is suggested by these remarks is that the nonstandard cases can be classified into several different natural categories. For example, we can obtain a joint distribution of height for individuals where we are measuring height separated by a fixed number of years. Such temporal distributions are of great interest, and it is in fact disappointing how poor the data are on
such a longitudinal variable as height in terms of good information about the sample paths of children's increase in height. In the case of such temporal separation there is no reason to suppose that the first measurement in any way interferes with the second.

A second kind of case occurs when the measurement of the first variable definitely interferes with the measurement of the second—in the sense that the first measurement distorts the nature of the object being measured in such a way that it affects in a significant fashion the result of the second measurement. Here the classical scientific cases are to be found in quantum mechanics. If we measure position of a particle, then in general we affect the particle's state in making that measurement and therefore when we measure momentum we get a different measurement than we would have anticipated getting if we had reversed the procedure and measured momentum first. In other words, we cannot get a "simultaneous" distribution of position and measurement for particles of atomic or subatomic size. We obtain a joint distribution but not the one in which we are interested.

There is a way of describing this situation that has not been used too often but that I think is important from a philosophical standpoint. We can easily claim that identity conditions have been violated in the following sense. When we measure the position of a particle, we change in an essential way the state of the particle and therefore the particle we are now observing is not, in one clear sense, the same particle.

We need to be somewhat careful in the characterization of identity conditions in these situations. We might want to hold on to a bare identity of the particle, but claim that what is important is that the properties of the particle do not have a continuing identity in time. So we cannot get a joint distribution of position and momentum because when we measure position, for example, we change the state of the particle in such a way that, if we now want to measure momentum, the momentum of the particle is significantly different from the momentum of the particle before the measurement of position. The identity, in other words, of the property of momentum has been destroyed. So when we talk about identity conditions here the appropriate thing, in general, is to talk about properties, although in some cases we can also be faced with the destruction of the particle itself, as in the case of the observation of photons.

This violation of identity conditions is not peculiar to quantum mechanics. In all kinds of situations, where interaction is expected between properties and where a measurement or treatment affects one property, we can anticipate the identity of another property of an object being destroyed or, if a less extreme term is preferred, changed. A simple but clear example is the following. Suppose the producer of a certain achievement test wants to
determine if the two forms of the test are parallel. One simple way to do this would be to give test A to students and then immediately give test B. If test A had no impact on the state of the student’s skill or competence being measured, then immediate retest with test B would be a good way to determine that test A and test B were parallel forms measuring the same competence in the student. Yet almost all psychological ideas about testing would hold that immediately giving test B after test A would lead to a poor measurement of parallelness of the tests, for the impact of having just taken test A would measurably change the student’s response to test B. Invasive measurements of physiological properties can have similar interference effects, even though we like to think that ordinary physiological measurements used for purposes of assessing the state of health of an individual do not significantly interfere with each other.

The third category represents the extreme case of modification, which has already been hinted at. In this case the first measurement destroys the object, and consequently the second measurement is not even possible with bare identity conditions of the object holding. The classic case in quantum mechanics is the measurement of properties of photons. For many kinds of measurements of photons one measurement is all we can make, but this is not special again to quantum mechanics and is not true in general for photons. A familiar example is that of sampling procedures for testing quality of objects. Many quality-assurance programs require destruction of the objects that are sampled and in many such measurements of quality only one significant measurement is made because that significant measurement, the one of importance, leads to destruction of the object. When complicated objects are tested for quality assurance, as, for example, by the Underwriters Laboratory, we are faced with progressive destruction of the object rather than destruction by a single measurement. In this case, ordinarily strong assumptions are made that gradual destruction of the object will not distort successive measurement on parts that have not been destroyed. Ordinarily we feel quite comfortable with the decomposition assumptions that are made in these cases.

A fourth kind of case of great philosophical and theoretical interest is when the measurements cannot be made in principle but are assumed to exist or perhaps even have values that can be inferred from other measurements that are made. We can return to Bell’s inequalities to find good examples of this last category. Note that when we ask for the joint distribution of random variables $A, A', B,$ and $B'$, we are not given in the Bell inequalities the two missing covariances, $E(AA')$ and $E(BB')$. In other words, we do not observe the correlations between measurements taken on the same side of the measuring apparatus with different settings and of course at different times. There is no natural way to do this. We send a particle through the
apparatus, a single particle in principle if not in practice, and we measure
the correlation—which is the same as the covariance for these random vari-
ables whose expectations are zero—and we observe, for example, correlation
$E(AB')$. But we have no natural way of identifying what we would be talk-
ning about in talking about correlations for separate measurements at separate
times of $A$ and $A'$ or correspondingly of $B$ and $B'$. Consequently, in asking
about the existence of a joint distribution we are simply asking if there can
exist numerical assignments to the two missing covariances, that is, $E(AA')$
and $E(BB')$, such that a joint distribution consistent with all the six covari-
ances can be given. This kind of question is an unusual kind of question. It
is not at all natural, from an experimental standpoint, to ask for the values
of these two missing expectations.

Let me focus very sharply on this question. It seems to me it is not at
all clear what are the identity conditions we are focusing on, either at the
level of properties or at the level of “bare” particles, when we ask for the
two missing covariances. These are covariances that we would not naturally
inquire about. Let us consider a similar situation of a very simple sort from
a setting that is surely noncontroversial. Suppose we have two treatments
for a certain kind of cancer. We give one treatment to some patients and the
other treatment to other patients. We cannot ask for the correlation between
the two treatments because no individual is being given both treatments. To
ask for the correlation of the two treatments does not, from an experimental
standpoint, make sense. Introduction of the correlation of the two treatments
rests upon some further theoretical assumptions not obvious at all on the
surface.

Now in quantum mechanics the whole point of the Bell inequalities is
that they are violated by appropriate choice of angles of measurement for
the four random variables so that no joint distribution exists. We might ask,
well, even though the joint distribution does not exist, can we theoretically
compute the missing covariances $E(AA')$ and $E(BB')$? As far as I can see the
answer is strictly negative: We cannot. I conclude that, looked at from a con-
ceptual standpoint and keeping in mind the identity conditions we naturally
impose for properties and for “bare” particles, the tests of hidden variable
theories generated by the Bell inequalities and the Bell-type experiments are
not as straightforward as it would be natural to expect. At the very least, we
cannot write out the data tables to generate a joint distribution in the way
that would be, in any ordinary experimental situation, straightforward. The
inference about the nonexistence of hidden variables must be at best quite
indirect.
**Visual Space**

One of the classic problems in the philosophy of science has been the analysis of the nature of physical space. As everybody knows, the discovery of non-Euclidean geometries in the nineteenth century and the development of the theory of relativity in the twentieth century have changed forever the long-held idea that physical space is necessarily Euclidean in character. Much has been made by philosophers of all sorts of the conceptual importance of the changes in our theories of physical space.

Much less attention has been devoted to the nature of visual space, that is, the psychological space in which we see objects. This visual space has a lot of special characteristics. First, we must think of it in binocular terms. Second, the space is certainly not homogeneous in the way in which Euclidean space is. The observer looking out in front of himself with a different viewpoint on what lies straight ahead, as opposed to what lies to the right or left, immediately imposes natural distinctive directions in visual space and thus upsets our ideas of homogeneity so familiar in the discussion of physical space. Surprisingly, however, the analysis of visual space has not gone into this problem from a foundational standpoint in very great depth. I will not have more to say about it here, although I recognize its importance, and it is easy enough to see it generates an axiomatic problem that as far as I know has not yet been solved at all, that is, to formulate visual space with appropriate and direct account taken of the facts just mentioned.

Returning now to the main question, in a previous article (Suppes, 1977) I looked at the history of discussions of this problem, beginning with Euclid. Here I want to concentrate on the various methodologies that have been considered for studying the nature of visual space and also some of the results that have been obtained experimentally. The subject is complicated. The number of experiments is large, and often the nature of these experiments is involved, especially in terms of the actual parameters estimated from data. I shall therefore not cover in anything like serious depth all aspects even of the restricted questions I want to consider, but I hope to be able to say enough to show that the problem of the nature of visual space is in itself an interesting philosophical one, even if we should not attach to it the same primary importance that has been historically attached to the nature of physical space. Perhaps the central point to emphasize in the context of the present lecture is that philosophical speculations about visual space conducted independent of consideration of the very large modern psychological literature on the question seem naive and wholly inappropriate. On the other hand, the traffic can be two-way: I think philosophers have something to contribute in their own way to the conceptual discussion of psychologists on the nature of visual
space. I hope that some of the comments I make will give a sense of the kind of help each group may give the other.

**Methodology.** What would seem to be, in many ways, the most natural mathematical approach to the question of the nature of visual space has also been the method most used experimentally. It consists of considering a finite set of points. Experimentally, the points are approximated by small point sources of light of low illumination intensity, displayed in a darkened room. The intuitive idea of the setting is to make only a finite number of point-light sources visible and to make these light sources of sufficiently low intensity to exclude illumination of the surroundings. The second step is to ask the person making visual judgments to state whether certain geometrical relations hold among the points. For example, do points \(a\) and \(b\) appear to be the same distance from each other as points \(c\) and \(d\)? (Hereafter in this discussion I shall refer to points, but it should be understood that I have in mind the physical realization in terms of point-light sources.) Another kind of question might be, Does the angle formed by points \(abc\) appear to be congruent or equal in measure to the angle formed by points \(def\)?

Another approach to such judgments is not to ask whether given points have a certain relation but rather to permit the individual making the judgments to manipulate some of the points. For example, first fix points \(a\), \(b\), and \(c\) and then adjust \(d\) so that the distance between \(c\) and \(d\) appears the same as the distance between \(a\) and \(b\). Although the formulation may sound metric in character, the judgments are often of a qualitative nature—for example, that of congruence of segments, which I also formulate here as equidistance of points. However, in other experiments, magnitude estimates of the ratio of distances are required, in order to apply metric methods of multidimensional scaling.

Once such judgments are obtained, whether on the basis of fixed relations or ratios, or by adjusting the position of points, the formal or mathematical question to ask is whether the finite relational structure representing the experimental data can be embedded in a two- or three-dimensional space of a given type—Euclidean, hyperbolic, etc. The dimensionality depends upon the character of the experiment. In many cases the points will be restricted to a plane and therefore embedding in two dimensions is required; in other cases, embedding in three dimensions is appropriate.

By a **finite relational structure**, I mean as usual a relational structure whose domain is finite. To give a simple example, suppose that \(A\) is the finite set of points and the judgments we have asked for are judgments of equidistance of points. Let \(\approx\) be the quaternary relation of congruence. Then to say that the finite relational structure \(A = (A, \approx)\) can be embedded in three-dimensional Euclidean space is to say that there exists a function \(\varphi\)
defined on $A$ such that $\varphi$ maps $A$ into the set of three-dimensional Cartesian vectors of real numbers and such that for every $a$, $b$, $c$, and $d$ in $A$ the following relation holds:

$$ab \approx cd \iff \sum_{i=1}^{3} (\varphi_i(a) - \varphi_i(b))^2 = \sum_{i=1}^{3} (\varphi_i(c) - \varphi_i(d))^2,$$

where $\varphi_i(a)$ is the $i^{th}$ coordinate of $\varphi(a)$. Note that the mapping into vectors of real numbers is just mapping visual points into the Cartesian representation of three-dimensional Euclidean space. In principle, it is straightforward to answer the question raised by this embedding procedure: Given a set of data from an individual's visual judgments of equidistance between pairs of points, we can determine in a definite and constructive mathematical manner whether such a Euclidean embedding is possible.

Immediately, however, a problem arises. This problem can be grasped by considering the analogous physical situation. Suppose we are making observations of the stars and want to test a similar proposition, or some more complex proposition of celestial mechanics. We are faced with the problem recognized early in the history of astronomy, and also in the history of geodetic surveys, that the data are bound not to fit the theoretical model exactly. The classical way of putting this is that errors of measurement arise, and our problem is to determine if the model fits the data within the limits of the error of measurement. In examining data on the advancement of the perihelion of Mercury, which is one of the important tests of Einstein's general theory of relativity, the most tedious and difficult aspect of the data analysis is to determine whether the theory and the observations are in agreement within the estimated error of measurement.

Laplace, for example, used such methods with unparalleled success. He would examine data from some particular aspect of the solar system, for example, irregularities in the motion of Jupiter and Saturn, and would then raise the question of whether these observed irregularities were due to errors of measurement or to the existence of "constant" causes. When the irregularities were too great to be accounted for by errors of measurement, he then searched for a constant cause to explain the deviations from the simpler model of the phenomena. In the case mentioned, the irregularities in the motion of Jupiter and Saturn, he was able to explain them as being due to the mutual gravitational attraction of the two planets, which had been ignored in the simple theory of their motion. But Laplace's situation was different from the present one in the following important respect. The data he was examining were already rendered in quantitative form and there was no question of having an analytic representation. Our problem is that we are
faced simultaneously with the problem of both assigning a measurement and determining the error of that measurement. Because of the complexity and subtlety of the statistical questions concerning errors of measurement in the present setting, for purposes of simplification, we shall ignore them, but it is absolutely essential to recognize that they must be dealt with in any detailed analysis of experimental data.

Returning to the formal problem of embedding relations among a finite set of points into a given space, it is surprising to find that the results of the kind that we need for this perceptual problem are apparently not to be found in the enormous mathematical literature on geometry. There is a large literature on finite geometries; for example, Dembowski (1968) contains over 1200 references. Moreover, the tradition of considering finite geometries goes back at least to the beginning of this century. Construction of such geometries by Veblen and others was a fruitful source of models for proving independence of axioms, etc. On the other hand, the literature that culminates in Dembowski's magisterial survey consists almost entirely of projective and affine geometries that have a relatively weak structure. From a mathematical standpoint, such structures have been of considerable interest in connection with a variety of problems in abstract algebra. Some general theorems on embedding of finite structures in projective and affine planes are given in Szczerba and Tarski (1979) and Szczerba (1984).

The corresponding theory of finite geometries of a stronger type, for example, finite Euclidean, finite elliptic, or finite hyperbolic geometries, is scarcely developed at all. As a result, the experimental literature does not deal directly with such finite geometries, although they are a natural extension of the weaker finite geometries on the one hand and finite measurement structures on the other.

A second basic methodological approach to the geometrical character of visual space is to assume that a standard metric representation already exists and then to examine which kind of space best fits the data. I shall consider this approach in some detail. Of especial relevance here is multidimensional scaling, some results of which are reported.

**Luneberg theory of binocular vision.** The theory of binocular vision developed by R.K. Luneburg and his collaborators beginning in the 1940s is still the most detailed and sophisticated viewpoint to receive both mathematical and experimental attention. Much of the experimental work I report later takes as its objective testing directly the Luneburg theory or some modification of it; this is certainly true of the extensive experimental work of Tarow Indow and his collaborators.
Essentially, Luneburg wanted to postulate that the space of binocular vision must be a Riemannian space of constant curvature $K$ in order to have free mobility. It is well known that there are just three types of Riemannian spaces of constant curvature: If $K = 0$, the space is Euclidean; if $K < 0$, hyperbolic; and if $K > 0$, elliptic. Moreover, Luneburg felt the evidence is extremely strong for the conclusion that the space of binocular vision of most persons is hyperbolic. Luneburg and his collaborators adopted a metric viewpoint rather than a synthetic one toward hyperbolic space. We recapitulate some of the main lines of development here. In particular, we begin with the Luneburg (1950) axioms for determining a metric on visual space that is unique up to an isometry, that is, a similarity transformation.

Some preliminary definitions are useful. Let $A = (A, d)$ be a metric space, i.e., $A$ is a nonempty set and $d$ is a function mapping the Cartesian product of $A$ into nonnegative real numbers such that:

\[
\begin{align*}
    d(a, b) &= 0 \text{ if and only if } a = b, \quad (i) \\
    d(a, b) &= d(b, a), \quad (ii) \\
    d(a, b) + d(b, c) &\geq d(a, c), \quad (iii)
\end{align*}
\]

for any points $a$, $b$, and $c$ in $A$. In addition, $A$ is metrically convex iff for any two distinct points $a$ and $c$ in $A$ there exists a third point $b$ in $A$ such that

\[
d(a, b) + d(b, c) = d(a, c).
\]

The metric space $A$ is complete iff any Cauchy sequence of $A$ converges to a point in $A$. We define a betweenness relation $B_d$ (relative to $d$) and an equidistance relation $E_d$ in the obvious way:

\[
\begin{align*}
    B_d &= \{(a, b, c) : d(a, b) + d(b, c) = d(a, c), \text{ for } a, b, c \in A\} \\
    E_d &= \{(a, b, c, d) : d(a, b) = d(c, d), \text{ for } a, b, c, d \in A\}.
\end{align*}
\]

If we think of $B_d$ and $E_d$ as the (idealized) observed betweenness and equidistance relations in visual space, then roughly speaking any two metrics for which they are the same are related by an isometry. More explicitly and precisely, we have the following theorem.
THEOREM 3. Let $\mathcal{A} = (A, d)$ and $\mathcal{A}' = (A, d')$ be metric spaces that are complete and metrically convex, and let the betweenness and equidistance relations be the same for the two spaces, i.e., let $B_d = B_{d'}$ and $E_d = E_{d'}$. Then there is a positive real number $c$ such that for all $a$ and $b$ in $A$

\[ d'(a, b) = cd(a, b). \]

This theorem shows that it is easy to state a condition under which two metric spaces are isomorphic up to multiplication by a constant, in this case the positive number $c$. To determine that visual space must be a Riemannian space of constant curvature, still stronger assumptions are needed. In other words, just satisfaction, as such, of the numerical relations of betweenness and congruence in the sense of numerical distance is not sufficient. It is important to note this, for it might be thought that these conditions on betweenness and equidistance would be sufficient. The obvious point is that in no sense is the theorem strong enough to determine that the metric space is Euclidean, hyperbolic, or elliptic. Luneberg (1948) rightly says that the existence of such a unique psychometric distance function as expressed in the above theorem is supported by a variety of classical experiments in visual perception. In other words, there are many different experiments showing that we do have sensations of visual distance that can be represented uniquely by a metric up to selection of a unit of measurement. As Luneburg emphasizes, the assumptions of metrical complexity and completeness are needed for the uniqueness result, even though these axioms are not themselves directly tested in the relevant experiments.

Much too great a variety of spaces satisfies the hypothesis of the preceding theorem. We need to tighten the framework in order to have a limited number of spaces to investigate. Luneburg (1947, 1948, 1950) uses arguments from differential geometry to get the standard result that only in Euclidean, hyperbolic or elliptic spaces, that is, Riemannian spaces of constant curvature, is it possible to move about visual objects without deformation. The differential argument is not really satisfactory, but there is a well-known global argument not mentioned by Luneburg which also establishes this result. It is one of the most famous problems in the foundations of geometry, the Helmholtz-Lie problem on the nature of physical space.

Riemann’s famous lecture (1854), “Über die Hypothesen, welche der Geometrie zu Grunde liegen,” was responded to by Helmholtz (1868) more than a decade later in a famous paper, “Über die Thatsachen, die der Geometrie zu Grunde liegen.” Helmholtz makes it explicit that he wants to move from hypotheses to facts (Thatsachen) that underlie our conception of space. He argues that although arbitrary Riemannian spaces are conceivable, actual physical space has as an essential feature the free mobility of solid (i.e., rigid)
bodies. In metric geometry, a motion is a transformation of the space $A$ onto itself that preserves distances. Such a transformation or mapping is also called an isometry. Explicitly, if $A = (A, d)$ is a metric space, then $\varphi$ is an isometry or motion if and only if for every $a$ and $b$ in $A$

$$d(\varphi(a), \varphi(b)) = d(a, b).$$

Helmholtz based his analysis on four axioms, which we describe informally, following Freudenthal (1965). The first axiom asserts that space is an $n$-dimensional manifold with differentiability properties. The second axiom asserts there is a metric with motions as isometric transformations. The third axiom asserts the free mobility of solid bodies, which means that if $\varphi$ is an isometric mapping of a set $B$ of points onto a set $B'$ (in the same space), then $\varphi$ can be extended to a motion of the whole space. The fourth axiom requires that the motion should be periodic (and not spiraling). This is often called the monodromy axiom.

Helmholtz claimed to have proved that the only spaces satisfying his four axioms are the Euclidean, hyperbolic, and spherical spaces. Sophus Lie (1886) noticed a gap in Helmholtz's proof. Lie strengthened the axioms and solved the problem. Some years later, Weyl (1923) weakened Lie's assumptions. The details of the many subsequent contributions to the problem of weakening the axioms and retaining essentially Helmholtz's solution are to be found in Busemann (1955, Section 48) and Freudenthal (1965). The basic aim of the modern work is to eliminate differentiability assumptions, which are extraneous to the problem of characterizing the spaces that have free mobility of solid bodies. It is not appropriate here to formulate in technical detail the strongest theorems, that is, the ones with the weakest assumptions, that have been proved about the Helmholtz-Lie problem. The point is that whether we look at space either physically, as Riemann and Helmholtz certainly did, or as psychological, we want to have as a property of space—certainly to a very fine approximation—, the property of free mobility of solid bodies in the physical case and of visual images of bodies in the psychological case.

By this or other lines of argument, following Luneburg, we end up with three types of Riemannian spaces of constant curvature as the three candidates for visual space. As already remarked, I follow the usual notation to indicate the constant curvature by $K$: if $K < 0$ the space is hyperbolic, if $K = 0$ the space is Euclidean, and if $K > 1$ the space is elliptic.

Using the differential expression for a line element in Riemannian spaces we can express the line element for these three elementary spaces in the following simple canonical form:
where the sensory coordinates $\xi$, $\eta$, $\zeta$ are ordinary Cartesian coordinates in a three-dimensional Euclidean space when $K = 0$. The origin $\xi = \eta, \zeta = 0$ is selected to represent the apparent center of observation of the observer.

To present the fundamental ideas here in reasonable compass, it is necessary to skip at this point a number of technical details that are important in actual experimental applications of Luneburg's ideas. In particular, the equation for the line element is transformed once again to introduce an individual parameter $\sigma$ as well as $K$, which it is anticipated will vary from individual to individual.

As primary evidence for the hyperbolic nature of visual space, Luneburg referred to the classical experiments of Hillenbrand (1902) and Blumenfeld (1913). Let me refer here to Blumenfeld's experiments, which were improvements on those of Hillenbrand. Blumenfeld performed experiments with so-called parallel and equidistance alleys. In a darkened room the subject sits at a table, looking straight ahead, and he is asked to adjust two rows of point sources of light placed on either side of the normal plane, i.e., the vertical plane that bisects the horizontal segment joining the centers of the two eyes. The two furthest lights are fixed and are placed symmetrically and equidistant from the normal plane. The subject is then asked to arrange the other lights so that they form a parallel alley extending toward him from the fixed lights. His task is to arrange the lights so that he perceives them as being straight and parallel to each other in his visual space. This is the task for construction of a parallel alley. The second task is to construct an equidistance alley. In this case, all the lights except the two fixed lights are turned off and a pair of lights is presented, which are adjusted as being at the same physical distance apart as the fixed lights—the kind of equidistance judgments discussed earlier. That pair of lights is then turned off and another pair of lights closer to him is presented for adjustment, and so forth. The physical configurations do not coincide, but in Euclidean geometry straight lines are parallel if and only if they are equidistant from each other along any mutual perpendiculars. The discrepancies observed in Blumenfeld's experiment are taken to be evidence that visual space is not Euclidean. In both the parallel-alley and equidistance-alley judgments the lines diverge as you move away from the subject, but the angle of divergence tends to be greater in the case of parallel than in the case of equidistance alleys. Since the most distant pair is the same for both alleys, this means the equidistance alley lies outside
the parallel alley. These results have been taken by Luneburg to support his hypothesis that visual space is hyperbolic.

There is one obvious reservation to be made about Luneburg's inference that visual space is hyperbolic. There is no unique concept of lines being parallel in hyperbolic space. Indow (1979) discusses Luneburg's choice rather carefully and shows that it has some justification. Essentially he uses orthogonality to characterize being parallel. The situation is worse when visual space's being elliptic is tested by alley data, for no two lines can be parallel in such a space. A local concept must be used; for any standard choice it can be shown that in the elliptic case the parallel alley lies outside the equidistance alley.

Modern experiments. In Luneburg (1947, 1948, 1950) a number of experimental applications of the theory are sketched, for example, determination of the parameters $\kappa$ and $\sigma$ for a given observer, quantitative analysis of observational data for equidistance and parallel alleys, analysis and prediction of visually congruent configurations, and analysis of what is visually congruent to infinite horizons in physical space. Detailed analytic suggestions for experiments, quantitative analysis of the data, or determination of parameters was made later, after Luneburg's premature death, by his associate A.A. Blank (1953, 1957).

The most extensive early test of Luneburg's theory is found in the report of Hardy, Rand, Rittler, Blank, and Boeder (1953) of the experiments carried out at the Knapp Memorial Laboratories, Institute of Ophthalmology, Columbia University. Without entering into a detailed description of the experiments I summarize the experimental setup and their main conclusions. All experiments were carried out in a darkroom with configurations made up of a small number of low intensity point sources of light. The intensities were adjusted to appear equal to the observer but low enough not to permit any perceptible surrounding illumination. The observer's head was fixed in a headrest and he always viewed a static configuration—no perception of motion was investigated. All observations were made binocularly and the observer was permitted to let his point of regard vary over the entire physical configuration until a stable judgment about the visual geometry of the configuration was reached. An important condition was that all experiments were restricted to the horizontal plane.

Their main conclusions were these:

1. There is considerable experimental evidence to support Luneburg's prediction of when two configurations are visually congruent.

2. The experiments on parallel and equidistance alleys confirmed the classical results of Blumenfeld.
3. The efforts to determine the individual observer constants $K$ and $\sigma$ were not quantitatively successful. The main problem was drift of value of the constants through a sequence of experiments. The values obtained here and in related experiments supported Luneburg's hypothesis that, for most persons, visual space is hyperbolic, that is, $K < 0$.

Some closely related data and analysis are given in Blank (1958, 1961); in the main the results support the hypothesis that the curvature of visual space is negative. Other closely related experiments are those of Zajaczkowska (1956a,b).

The main group to continue in a direct way the theoretical and experimental work of Luneburg, Blank, and the Knapp Memorial Laboratories at Columbia has been the group centered around Tarow Indow, first at Keio University in Japan, and later at the University of California, Irvine campus. The list of publications extends over a period of more than two decades, and the references I give here are far from complete. Indow, Inore and Matsushima (1962a,b) reported extensive experiments conducted over a period of three years to test Luneburg's theory and, in particular, to estimate the individual parameters $K$ and $\sigma$. In the 3-point experiment, three points of light $Q_0$, $Q_1$, and $Q_2$, were presented in the horizontal plane relative to the subject, but both horizontally and vertically relative to the darkened room. $Q_0$ and $Q_1$ were fixed, and it was the task of the subject to move $Q_2$ so that the segment $Q_1Q_2$ was visually congruent to the segment $Q_0Q_1$. Conditions were similar in the 4-point experiment except that there were two points $Q_2$ and $Q_3$ to be adjusted so that $Q_2Q_3$ was visually congruent to $Q_0Q_1$. Of the 26 experimental runs with six subjects reported in (1962a), for 23 the estimated value of $K$ was in the range $-1 < K < 0$ with a satisfactory goodness of fit, which directly supports Luneburg's theory that visual space is hyperbolic. It should be mentioned that repeated runs with the same subjects showed considerable fluctuation in the value of $K$. In (1962b) the same experimental setup and subjects were used to replicate the alley experiments of Hillenbrand (1902) and Blumenfeld (1913) mentioned earlier. The equidistant and parallel alleys were in the relation observed in the earlier investigations and thus supported Luneburg's theory. But one aspect was theoretically not satisfactory. The values of $K$ and $\sigma$ estimated for individual subjects in (1962a) did not satisfactorily predict the alley data at all. Quite different estimated values were needed to fit these data.

Indow, Inoue and Matsushima (1963) repeated the experiments of (1962a,b), but with the points of light located in a spacious field. In the earlier experiments the most distant point of light was 300 cm from the subject. In this study it was 1610 cm, made possible by conducting the experiment in a large, darkened gymnasium. Qualitatively the results agreed with the
earlier experiments, but the quantitative aspects, as reflected in the estimated parameters $K$ and $\sigma$ did not.

Starting in 1967 and extending over a number of years, Indow and associates have applied multidimensional scaling methods (MDS) to the direct investigation of the geometrical character of visual space. However, there are several points about MDS to keep in mind. First, the results would be difficult to interpret if the number of scaling dimensions exceeded the number of physical dimensions. Second, MDS is most often used when there are not strong structural constraints given in advance. We know, on the other hand, that visual space is approximately Euclidean. Is the accuracy of MDS sufficient to pick up the sorts of discrepancies found in the alley experiments?

Matsushima and Noguchi (1967), using data from experiments of a Luneburg type—small light points in a dark room—and observation of stars in the night sky, obtained good fits to the Euclidean metric using MDS, with the appropriate dimensionality. On the other hand, the mapping between the physical space and the visual space determined by MDS was much more complicated than that proposed by Luneburg and in fact was too complicated to describe in any straightforward mathematical fashion. Nishikawa (1967) continued the same line of investigation by arranging the light stimuli in ways to test the standard alley results. He also suggests a theoretical approach to explain the Luneburg-type results which he replicated, on the MDS assumption that visual space is Euclidean. The essence of the approach is to assume the mapping function between visual and physical space changes substantially with a change in task and instruction. That there is such an effect seems likely but Nishikawa's theoretical analysis does not get very far. Similar theoretical arguments are advanced by Indow (1967), but he expresses appropriate skepticism about the Euclidean solution being satisfactory. Closely related empirical results and theoretical ideas are also analyzed with care by Indow (1968, and also 1974, 1975), who gives a particularly good quantitative account of the accuracy of the Euclidean model for various subjects when MDS methods are used.

Both methodologically and conceptually, it is natural to be somewhat skeptical that verbal estimates of ratios of distances—the MDS method used in the studies cited above—were sensitive and accurate enough to discriminate the Euclidean or hyperbolic nature of visual space. Indow in various places expresses similar scepticism about nonmetric MDS, whose lack of sensitivity to details is well known. A thorough discussion of these matters is to be found in Indow (1982), which also extends in a detailed way the methods of MDS to using a hyperbolic or elliptic metric as well as a Euclidean one. Although the quantitative fit is not much better than that of the Euclidean metric, the hyperbolic metric does give a better account of the standard alley data.
I restrict myself to a few other studies especially pertinent. Foley (1964a, 1964b, 1972) undertakes the important task of studying the qualitative properties of visual space, with an emphasis on whether or not it is Desarguesian. In the first two papers his answer is tentatively affirmative, and in the last one negative. Unfortunately, Foley's work represents a line of attack that has not been followed up by other investigators.

A significant and careful experimental study that reaches some different conclusions about visual space is that of Wagner (1985). The methodology of the work is notable for two reasons. First, the experiments were conducted outdoors in full daylight in a large field with judgments about the geometrical relations of 13 white stakes. Second, four different procedures were used for judging distances, angles and areas: magnitude (ratio) estimation, category estimation, mapping, and perceptual matching, where *mapping* means constructing a simple scale map of what is seen. Only the results for distance will be discussed here. In this case, perceptual matching was not feasible in the experimental setup and was not used.

The results for distance are surprising and interesting. The Luneburg model of hyperbolic space did not fit well at all. What did fit reasonably well is a Euclidean model of visual space, but the Euclidean visual space is a nontrivial affine transformation of Euclidean physical space. We may use x and y axes to discuss the results. The x-axis is the one perpendicular to the vertical plane through the eyes. It is the depth axis. The y-axis is the frontal axis passing through the two eyes. Let \((x, 0)\) and \((0, y)\) be two physical points such that \(x = y\), i.e., along their respective axes the two points are equidistant from the origin—the point midway between the two eyes, but in visual space \(x' = 0.5y'\)—approximately, i.e., visual foreshortening amounts to the perceived distance along the depth axis being half of the physical distance when perceived frontal distances are equated to the physical distances. Call this foreshortening factor \(c\), so that \(x' = cx\) and \(y' = y\), with \(c\) varying with subjects but being approximately 0.5. This very strong effect is highly surprising, for it has not been reported in Luneburg-type experiments, even with illumination. The surprise remains even when account is taken of the very different stimulus conditions of Wagner's experiments, although in an earlier study under somewhat similar conditions of full illumination Battro, Netto and Rozestraten (1976) also got results strongly at variance with the Luneburg predictions.

The notable omission, from a variety of viewpoints, is experimental study of projective geometry, for the essence of vision is the projection of objects in three-dimensional space onto the retina. Fortunately, Cutting (1986) has recently published a book on these matters. I will not attempt a resume of the many experiments he reports, but concentrate on one fundamental point.
The most important quantitative invariant of projective geometry is the cross ratio of four collinear points. Let \( a, b, c, \) and \( d \) be four such points. Then their cross ratio is \( \frac{ab}{ac} : \frac{bc}{bd} \). In perceiving lines in motion, i.e., from a continuously changing perspective, is it the cross ratio we perceive as invariant as the evidence for rigidity in the actual relative spatial positions of given lines? Cutting provides evidence that the answer is by and large affirmative. This result also solves La Gournerie's paradox (1859), described in Pirenne (1975): linear perspective is mathematically correct for just one fixed point of view, but almost any position in front of a painting will not disturb our perception. As Cutting points out, an explanation of the apparent paradox is that when the cross ratio of points projected onto a plane surface is preserved, it will be preserved from any viewer position. Further pursuit of this kind of projective analysis should throw further light on the Euclidean or non-Euclidean nature of visual space.

Some conclusions. Luneburg's fundamental hypothesis is the most striking of any that have been proposed for the nature of visual space just because of the relentless theoretical push on his part to work out so many of the implications of his fundamental ideas. As far as I can see he is the first person in the history of thought to make a really satisfactory detailed proposal that visual space is not Euclidean in character. There are of course predecessors going all the way back to Thomas Reid in the eighteenth century, but it is really Luneburg's virtue to have laid out the theory for the first time in anything like adequate detail.

Unfortunately, as we have seen from the many experiments surveyed, we cannot come to a simple conclusion of the kind that Luneburg would like to have found supported in as detailed a way as possible. We cannot conclude simpliciter that visual space is hyperbolic. Certainly we can give to Luneburg the point that there are simple experimental configurations in which the judgments of subjects certainly do support his hypothesis. On the other hand, there is a great variety of evidence supporting the view that with a change in experimental circumstances, for example, in the kind of lighting, very different results can be obtained. The study of visual space, like the study of other psychological phenomena, turns out to be quite sensitive to particular experimental configurations and particular experimental environments. From a broad methodological standpoint, in fact, it might be claimed that this is the most severe difficulty of developing in almost every area of psychology an adequate deep-running general theory.

In any case, it is important in thinking about visual space to contrast the variety of results to those obtained in the study of physical space. It may very well be that in an environment of black holes we shall find our ordinary ideas of physical space no longer at all valid and the nature of physical space
changing rapidly, since it depends upon the swirling environment of the black hole. But for measurements in human environments and on a human scale the great constancy of physical space is one of the most fundamental facts of the universe in which we live. The systematization of these physical facts many centuries ago was one of the most important achievements of Greek mathematics and science. It is a mistake to think we can achieve anything like a similar systematization of great general validity in the case of visual space, at least if we try to think about visual space in the way we think about classical geometry. The most obvious distinction is that visual space is in certain fundamental respects closer to classical physics than to classical geometry. What I have in mind by this remark is that context is rampant in classical physics but not at all in classical geometry. If we have two bodies interacting with each other gravitationally, we completely expect the motion of the two bodies to be disturbed by the introduction of a third, which changes the environment and thereby the context. We would in fact be astounded if no change occurred. Endless other physical examples easily come to mind. We might even say that the study of dynamics in all branches of classical physics is to a large extent the study of changing context.

By these remarks I do not mean to suggest that it will be an easy matter to move from a framework of classical geometry to one of classical physics and thereby achieve a deeper-running, more satisfactory general theory of visual space. I am only drawing an analogy when it comes to the treatment of context. I think we are as yet far from clear how to build theories to take account of the great variety of context effects that have been experimentally studied thus far. But I also do not want to suggest that I think the situation is scientifically hopeless, that the contexts are so complicated and devious that they cannot be reduced in a feasible way to a theoretical framework. We have a lot to build on, namely, the kinds of experiments that have supported very well Luneburg's ideas and the kinds of other experiments, for example, those of Foley and of Wagner, which go in a different direction but in a way that we can understand and begin to bring within the fold of a general theory. It is also important to recognize that physics operates only with a very selected body of experiments. We do not want to make the mistake of thinking that we can move in any direct way to the study of visual space in wholly natural environments. The need for the present is to enlarge the canonical experiments sufficiently to get a range of variation, but with contexts that we can manage.

The experimental study of visual space is a tedious business, pursued today in proper scientific fashion by only a small band of intrepid psychologists. In many ways, our study of visual space is still at the beginning because we do not yet have a general theoretical framework within which to operate. Philosophers in search of generalities about space need to be chary
of having too fixed or detailed views about the nature of visual space. One conclusion of considerable historical and philosophical interest is that a variety of experiments certainly do support the conclusion that visual space is not Euclidean.

NOTES

1 The constant is the reciprocal of the period of the motion of the two particles in the plane.

2 The correspondence between a solution of the differential equation and a sequence of integers is the source of the term symbolic dynamics. The idea of such a correspondence originated with G.D. Birkhoff in the 1930s.

3 Subjects were asked to judge ratios of interpoint distances.

REFERENCES


