

Intellectual Autobiography, Part II, 1979–2006

I ended Part I of my Intellectual Autobiography, written in 1979, with the year 1978. So I now take up the narrative from that point onward. I begin, as earlier, with a long section on research, followed by a section on personal reflections. Although my interests have changed somewhat, I follow the same main seven divisions I used earlier: foundations of physics; theory of measurement; decision theory, foundations of probability, and causality; foundations of psychology; philosophy of language; education and computers; and philosophy and science.

IV. Research 1979–2006

Foundations of Physics

Most of my work in this period was, first, in collaboration with Mario Zanotti, as a continuation of our earlier research begun in 1974; and second with Acacio de Barros, a Brazilian physicist.

Zanotti and I wrote, between 1980 and 1991, five papers on quantum mechanics, all focused, from one angle or another, on the quantum phenomena of entanglement, first brought to prominence by the physicist John Bell. We considered the Bell inequalities in all of these articles. I won't give a technical account here of those inequalities, but they are used to show that there is no proper classical theory of the experiments to which they are applied. A sketch of the argument goes like this. In classical physics and, in fact, almost all branches of science except quantum mechanics, when we find simultaneously correlated events, such as fever and headache, we search for a common cause. Einstein believed to the end that such common causes, even if not necessarily observable in any direct way, should be found for quantum phenomena as well. The search for such causes is the search for hidden variables in the language of this enterprise, when pursued in the quantum domain. The Bell inequalities give necessary and sufficient conditions for such observable correlated phenomena as the spin of particles, often represented in entanglement experiments by the polarization of photons, to have hidden variables as common causes of the correlations. The point of the inequalities is to give a condition, just in terms of the observable correlations, for there to be such hidden causes. It turns out, not initially clear that this is the case, that the inequalities are equivalent to requiring that the random variables or probabilistic events having the pairwise correlations must have a joint probability distribution of all four together. So Bell's requirement was just equivalent to a well-known probabilistic problem for a great variety of special cases: when do random variables having certain marginal distributions have a joint distribution? The subject is messy and full of tantalizing partial

results. Some of the nice earlier results, with no connection with quantum mechanics, were given by the distinguished French mathematician Maurice Frechet, whom I met in the 1960s in Paris, but, as is usually for me and other academics, who talk to a lot of different people, I cannot remember anything about our conversations. (John Bell I did meet once and we had a very good talk about the foundations of physics at the Stanford Faculty Club; Bell, like a lot of other scientists who are well-known for conceptual innovations, had a strong natural interest in philosophy.)

I won't enter into the details of what Zanotti and I did in these five papers, except to mention one general result that has applications elsewhere. This is our proof that the existence of common causes for phenomena of any kind is exactly equivalent to their having a joint probability distribution (Suppes and Zanotti, 1981). We proved it for finite random variables, and it was easily extended to continuous random variables by Paul Holland and P. R. Rosenbaum (1986). (I am proud to say that I was Paul's Ph.D. advisor in the Department of Statistics many years before.) One immediate application is that the search for latent variables, as common causes are often called in the social sciences, always exist for any social phenomena that have a joint probability distribution, which in such research is implicitly always assumed. This means that further structural conditions must be imposed to get a result on latent variables of any scientific interest.

Years later, in fact in 2000, Acacio de Barros and I derived a set of similar inequalities for what are called Greenberger, Holt and Zeilinger (GHZ) configurations of three entangled particles. The surprising result is that we found inequalities that looked just like the Bell ones when correlations between pairs of particles are replaced by moments for three particles, i.e., expectations of the products of the spins of three particles. It was a particular pleasure for me, now 78 in 2000, to have this article published in *Physical Review Letters*, the leading physics journal in the world in the opinion of many people.

As I have emphasized in several publications, entanglement is the great mystery of nonrelativistic quantum mechanics. As far as the current evidence shows, particles can be entangled, even though separated by large distances. This seems to be a revival of action at a distance, as occult as was gravitation in the seventeenth century. Remember, Newton was quite clear that he did not understand the cause of gravity, and it violated all the precepts of the Cartesian physics that was still being taught, even at Cambridge, in the second half of the seventeenth century.

I now examine carefully some possible consequences of these ideas. The primary criterion of adequacy of a probabilistic causal analysis is that the causal variable should render the simultaneous phenomenological data conditionally independent. The intuition back

of this idea is that the common cause of the phenomena should factor out the observed correlations. So we label this principle the *common-cause criterion*. Although satisfaction of it is not the end of the search for causes or probabilistic explanations, it does represent a significant and important milestone in any particular investigation

Much of the earlier discussion of hidden variables in quantum mechanics centered around the search for deterministic underlying processes, but for some time now the literature has also been concerned with the existence of probabilistic hidden variables. It is a striking and important fact that even probabilistic hidden variables do not exist when the observable random variables do not have a joint distribution.

The next systematic concept is that of *locality*. I mean by this what John Bell (1966) meant:

It is the requirement of locality, or more precisely that the result of a measurement on one system be unaffected by operations on a distant system with which it has interacted in the past, that creates the essential difficulty. ...The vital assumption is that the result B for particle 2 does not depend on the setting \mathbf{a} , of the magnet for particle 1, nor A on \mathbf{b} .

Quantum entanglement, as is well known, violates both the principle of a common cause and the principle of locality. There are other consequences, even more directly challenging, to the edifice of much modern physics. My objective here is to describe these additional difficulties, without the necessary technical developments for a really clear and transparent exposition.

The more recent entanglement experiments demonstrate two critical features:

- (i) Entangled particles can be experimentally separated by distances of the order of 100 kilometers (Zeilinger, 2005).
- (ii) Changing the angle of the measurement apparatus for one particle can be done just before the other particle reaches its measurement apparatus. So the causal effect of this change is transmitted, in recent experiments, at least five times faster than the speed of light.

These experimental results seem to present us with the following stark alternatives:

- A1. Reject the special theory of relativity and return to the Galilean invariance of Newtonian mechanics, where even instantaneous action at a distance is possible, and any finite velocity creates no problem.

A2. Hold fast to the special theory of relativity and accept macroscopic backward causation.

Backward causation follows from (i), (ii) and the special theory of relativity. With speeds faster than that of light, particles separated by a space-like interval can causally interact. In that case there is an inertial frame F_2 of reference, related by a Lorentz transformation to the initial frame F_1 , such that in this frame F_2 the causal effect of changing the measurement angle of one particle is an event A in the future of the event B of changing the other particle. Thus, backward causation. Event A occurs, relative to frame F_2 after event B , but A is the cause of B . I find this hard to accept. But perhaps it also helps explain why relativistic quantum mechanics seems still to lack a satisfactory foundation.

In 1991 Zanotti and I published three articles on entanglement. This was our last work on foundations of physics. In that same year I met Acacio, who joined us from Brazil to teach physics in the Education Program for Gifted Youth at Stanford. (More on EPGY later.) He and I also began talking to each other about a variety of issues in foundations of physics. Our first paper, published in 1994, was on a random walk approach to optical interference. This was the beginning of several papers on a semiclassical theory of optics that postulated photons, but with definite trajectories, and the associated electromagnetic field was now “reduced to” a probability distribution of photons. We published four more papers broadly in this area and entanglement, before the 2000 article on GHZ quantum entanglement of triples of particles, I will not go into the details, but remark on two aspects of this research that has been significant for me. It represents the first time I did any sustained research with a working physicist. Back in the early 1960s, when I wrote my first article on probability in quantum mechanics, I got some help from Sidney Drell, who became one of Stanford’s distinguished theoretical physicists. But it was different with Acacio. We worked day after day in a sustained effort to understand problems we found puzzling. Earlier coworkers, like Zanotti, had a background closer to mine from philosophy, foundations of mathematics and mathematical statistics. Acacio is a true-blue physicist and we often differed sharply in our natural approach to the problems we tackled, but this meant that the union of what we did know covered much more ground.

I end this part with reference to only two other relevant research efforts. In 1998 I published a paper on pragmatism in physics, to which I return later. The second effort was begun in 2000 with Acacio to apply the theory of oscillators to the representation of brain activities. This effort is continuing, but I will say something about this, beginning in 2000, under the general heading of Foundations of Psychology.

Theory of Measurement

The big event was the completion of the treatise *Foundations of Measurement*. Volume I was published in 1971. Eighteen years later, Volume II was published in 1989, and Volume III one year later in 1990. Dave Krantz had been the lead author in finishing Volume I; I took that responsibility for Volume II; and Duncan Luce did so for Volume III. (The four authors of the volumes were David Krantz, Duncan Luce, Amos Tversky and me. Unfortunately, Amos died early in 1996, just about at the apex of his brilliant scientific career.) Writing a treatise like the one on measurement is a mixture of many intellectual activities. Perhaps the most important and the most time consuming is organizing sets of results in a coherent fashion that brings out concepts and their relations in a way that was not as evident in their original development, or in the direction of their most recent application. This sounds very general and abstract. I will give a couple of examples from Volume II. Chapter 13, which I mainly wrote, presents in the first 38 pages a variety of classical geometric structures, such as absolute, affine and projective spaces, mainly according to their axiomatic foundations. To write something detailed, but not too lengthy, took much more effort than I expected, probably because I misjudged the depth of my knowledge at the beginning.

Another problem, more serious in a way, is that I initially naively intended to do something much more original. I wanted to emphasize geometric constructions, such as the familiar affine one of bisecting a line segment, to generate only finite geometries that have exact representations. By 'exact' I mean representations unique up to the relevant classical group, for instance, the projective or affine group of transformations. Well, this is a reasonable project, but one I had not thought through, and when I began thinking about hyperbolic and elliptic spaces I needed to do more work than I had planned on. My attachment to these finite geometries marks a philosophical attitude of mine toward the foundations of mathematics. Geometry has been neglected, following the arithmetization of analysis in the latter part of the nineteenth century. Moreover, the search for a foundation has been dominated by the search for a monolithic one that would work for all of mathematics. That ideal is still in place, and, of course, I fully recognize the value of it. But as a thoroughgoing pragmatist about science and philosophy, I have with experience (for which age is a surrogate) become ever more skeptical of the need for such results. Indeed, focus on them can be intellectually mistaken. A good, but different example was the long search in philosophy until recently for an absolutely certain grounding for epistemology. We can ring the changes on the great philosophers who sought such a foundation from Kant to Carnap. In applications of mathematics, including geometry, a pluralistic view is easy to defend. I mean by this that limited use of mathematics, in fact, mainly elementary use, is made in almost all applications in physics, biology, economics or elsewhere. Such limited use ordinarily can be given a respectable finitistic foundation if one is needed.

Let me be more concrete to illustrate this last point. In order to prove the standard representations in analytic geometry of the various kinds of structures I mentioned—an enterprise which has been famous since Descartes, it is necessary to add some kind of axiom of continuity, which is usually formulated in a strong form equivalent to a completeness axiom for the real numbers, such as: any nonempty set of real numbers that has an upper bound has a least upper bound. Such an axiom, not known in this form, or its geometric equivalents, in ancient Euclidean geometry, says something that is important to be clear about in pure mathematics. But it most surely expresses a principle that will not be tested directly in any physical or other scientific experiments. Having such strong smoothness conditions was no doubt important in nineteenth-century mathematical physics, when the aim was to find symbolic methods of computation to generate closed forms, i.e., finite expressions as solutions to problems, whenever possible. Already before the end of the twentieth century it was evident that the scene had changed. Most problems that arise naturally in physics now do not have closed-form solutions. The beautiful classroom examples of classical mechanics, such as Newton's solution to the isolated two-body problem, are rare in nature or the modern physics laboratory. Numerical work more or less dominates all the way down, and what is symbolic foundation can ordinarily be given a rather straightforward constructive foundation. These comments reflect my biases about measurement as well. I personally have always, and now ever more so, stressed the need to keep the subject closer to the laboratory and the practice of actual measurement, in the lab and outside as well.

So I will only briefly consider a more radical second example. The second part of Chapter 13 is devoted to classical space-time and that of special relativity. Standard strong assumptions about the smoothness of space-time are used. But it is evident enough they are in no sense subject to a detailed positive confirmation. It needs little argument to show that no experiments as we know them will refute the assumption of a sufficiently fine discrete space-time. In fact, high-energy particle physics, when pushed, would favor the latter. Experiments that try to work with ever finer volumes of space-time do not find smoothness, but explosive reactions—this sounds too vague, but it can be technically expanded on in a standard way. For details on quantum fluctuations in a vacuum, see (Milonni, 1993).

This finitistic laboratory orientation of my recent thinking about measurement is brought out in a paper I recently wrote (Suppes, 2006) that is about indistinguishability of observations or what is being observed, the subject of Chapter 16 of Volume II, which is entitled "Representations with Thresholds", also a subject dear to my heart. But in this recent paper, rather than taking the classical psychological stance, well documented experimentally, that indistinguishability within a threshold is a nontransitive relation, I considered measurements with finite scales. So, for example, if I have weights to measure accurately to a gram, the relation of

indistinguishability of the weight of objects falling between 2 and 3 grams is transitive, but this is all that can be said. We can make up an exact number for computational purposes, but the result is indeterminate, and should, strictly be represented by a pair of upper and lower measures, the upper measure being 3 grams, and the lower one, 2 grams. Do not, when you read this, make the mistake of saying, “Well, anybody can improve such a coarse measurement as this.” The point is that, unlike the nineteenth-century fantasy that classical physical measurements could always be indefinitely improved, we now recognize that this is not the case. I can always find, at a given level of technological development of measuring apparatus, a lower bound beyond which improvements cannot go. And there are clear theoretical arguments I will not enter into here, as to why improvements will not converge in time to perfection. Making up an exact result is nice for computations, but should not be part of thinking about what the world is really like. Another point to emphasize. This indeterminate approximation precedes, rather than follows, statistical calculations, but this point also I shall not go into now in detail.

These remarks about indeterminate approximate measurements fit in well with the pragmatism I endorsed earlier. It is part of such a program to insist on plurality and partiality, as well as, indeterminacy of results. What I said, for instance, about weight, holds even more for subjective probability, or, if you wish, for propensity interpretations of probability as well.

Decision Theory, Foundations of Probability, and Causality

During the first ten years, 1979–1988, I published two books and a lot of articles under this heading. The list is too diverse and varied to summarize in a synoptic conceptual fashion. But certain themes are constant. In 1981 *Logique du Probable* was published in Paris. The text was a revision of a series of four lectures given at the College de France in November 1979. I enjoyed the extensive discussions that followed each lecture, and so, I think, did Jules Vuillemin, who, as Professor of Philosophy, at the College, sponsored my appearance. But I don’t think the book was much read in France. With its emphasis on the philosophical relevance of many concepts and results from decision theory and the foundations of probability, it strayed quite far from the main concerns of French philosophy. I ended the book with a summary of eleven principal theses I had defended. Looking back now, my favorite two are these:

Intuitive judgments that are not able to be expressed verbally are an essential and uneliminable feature of our processes of decision making.

Determinism is dead; its short existence, from Laplace to Heisenberg, Born and Dirac, has ended.

No doubt too brief to be taken seriously, but meant to summarize a longer argument in the text. Jules was a good friend for many years. He was the most learned person I have known in the history of philosophy and science. He often corrected my historical errors on many points. I continue to miss him since his death in 2001.

Three years after the publication of *Logique du Probable*, in 1984 I published *Probabilistic Metaphysics*, based on the Hågerström Lectures, which I gave in Uppsala, Sweden in 1974, on the invitation of Stig Kanger and other members of the Department of Philosophy at Uppsala. Many themes of this longer book overlapped with those of the earlier one. The headings of the five chapters after the introductory one give a quick overview of the focus: randomness in nature, causality and randomness, uncertainty, incompleteness, and the plurality of science. The seventh chapter applies many of these ideas to the analysis of language, with an appendix on probabilistic grammars. The eighth and last chapter is on rationality and the orientation is close to that of the earlier book in French.

On one major point I have since changed my mind. In *Probabilistic Metaphysics* I argued strongly that deterministic theories of most natural phenomena must be false, and that determinism as a general view of the natural world is a false one. I have been persuaded by striking invariance results in ergodic theory established by my colleague Don Ornstein and others (see especially Ornstein and Weiss, 1991). They prove the impossibility of distinguishing between the correctness of a deterministic theory and a stochastic one of the classic billiard-ball example of motion, once only a finite accuracy of measurement, bounded away from zero, is assumed. In this case, two theories that are mathematically inconsistent cannot even in principle be distinguished by an unbounded but finite number of observations of the behavior of a billiard ball on a table with a convex obstacle in the center. To play off the themes of Kant's antinomies in the *Critique of Pure Reason*, especially the third one on causality, the revised view is that determinism is not false, but transcendental. I elaborated on this idea in several articles in the 1990s. The view I now advocate is that either universal determinism or universal indeterminism are transcendental, i.e., beyond experience. So a choice between them is a transcendental metaphysical one. Moreover, the invariance I mentioned is undoubtedly present in almost every variety of complex natural system from chaotic billiard balls to quivering *Aplysia* or your favorite bacteria. That transcendental determinism will not disappear as an attractive doctrine is supported by its presence even in theories of quantum mechanics, notably the Bohmian approach.

My interest in the foundations of probability continued during this period, but I will not enter into the technical details of various papers. I summarized my thoughts, and answered some questions I now think I understand, relatively recently in the long Chapter 5 of *Suppes* (2002).

Foundations of Psychology

I continued work on most of the psychological topics discussed in the earlier period prior to 1979, but I will only mention here two new areas of psychological research I had not contributed to or even really thought about before. One is the study of eye movements and the other is the study of electrical and magnetic activity in the brain.

Eye Movements

I cannot now remember how I first got interested in eye movements. But it was about 1980. Somewhat rashly I agreed to host an eye-movement laboratory run by Jim Anliker, which was supported by one of the government agencies. Jim was an experienced scientist, with a doctorate in psychology, who had devoted at least the previous ten years to eye-movement experiments. But he had published very few papers, and did not have much interest himself in building stochastic models of the obviously probabilistic behavior of eye movements. So it was a desirable collaboration, with each of us contributing, but also depending on the other. We did some experiments I still like in which we recorded the eye movements of students doing arithmetic exercises. The actual behavior was much more complex than the hallowed explanations of how students work such exercises and use the algorithms they were taught. Two detailed papers on this work were published in 1982 and 1983. I was then invited by Eileen Kowler, another specialist in eye movements, to write a review paper on eye-movement studies of both arithmetic and reading. This time no new experiments were conducted, but I assiduously studied the large literature on eye-movements in reading, which I previously knew only in a superficial way. I organized my ideas to formulate some stochastic models of eye movements in reading in this article (1990) and in a second one (1994) that grew out of my participating in a conference on autism in Sweden. A couple of years later, Julie Epelboim received a postdoctoral fellowship to study mathematical models in psychology with me. Besides Jim Anliker, she was the only experimental psychologist, well-trained in the specialized lore of eye-movement research, with whom I have worked. We published in 2001 a paper I still like a lot on eye movements in solving geometric problems with diagrams. The eye movements revealed directly the subtle interplay in the shifting back and forth between the texts of the problems and the accompanying diagrams. The model we developed was used to estimate the size of working memory in solving such problems.

Brain Experiments

I was invited to give a paper and participate in a conference at the University of Minnesota on high performance computer applications in the behavioral sciences in May of 1996. I don't really remember the details of my own paper, but I was fascinated by the lecture of Professor Sam Williamson. This decisive event led to research that

has occupied me since 1996 more than any other area of research and seems likely to continue to do so in the future.

Here is a record of what I wrote about that lecture and what I thought about in the days following it (May and June, 1996).

Saturday, 11 May 1996

A lecture on current work on Magnetic Imaging (MEG) by Sam Williamson, a physicist at NYU who has been one of the early workers in developing the use of SQUIDS for recording the magnetic field generated by electrical activity in the brain. I asked Sam various questions in a skeptical mode about what could and could not be done. I also continued to chat with Sam and got a better sense of the nature of the magnetic recordings. This is the first time that I had heard about any of the details of magnetic recording with SQUIDS (superconducting quantum interference devices).

22 May 1996

On the way to Europe, I stopped with Christine in New York and during the afternoon of May 22, I met Sam Williamson at NYU's Center for Neural Science.

4 June 1996

The following memo of record is directly transcribed from the writing of this note on 4 June 1996 on United Airlines Flight 955 from London to San Francisco.

Memo of Record

...The purpose of this memo is to record my meeting with Sam Williamson at NYU – Sam is a physicist at the NYU Center for Neural Science. We met from 2 PM to 3 PM on Wed., May 22 – Christine and I were on our way to Europe and we stopped overnight in New York. I called Sam in hope of meeting him after hearing a talk by him on MEG (magnetoencephalography) – the technique using SQUIDS for measuring magnetic fields arising from neuronal activity. The fields can be measured externally but close to the scalp.

In the course of our intense one-hour conversation, I suggested to Sam that it might be possible to recognize internal speech, initially just individual words such as “Yes” and “No”, by analyzing the magnetic wave produced by the neuronal activity by an approach similar to that used for standard speech

recognition. My conjecture is that the wave pattern of internal speech must, in a precise sense to be determined, be approximately isomorphic to the actual speech wave.

A main basis for the conjecture is that the time course and sense of prosodic emphasis in the conscious awareness of the internal speech seems, as would be expected, very similar to that of the corresponding actual speech.

Although the physical details are bound to be subtle and complex, there is much to support this conjecture. We produce too much speech of which we are consciously aware, and it is too easy to produce corresponding internal speech for the relation between the two to be impossibly complex. The subjective sense of similarity is of the greatest importance.

Let us suppose that the area of the cerebral cortex where internal speech takes place can be localized, because it is in fact a relatively local phenomenon. If, on the other hand, production of such speech involves a distributed network of neurons, then internal speech recognition will be much more difficult.

Assuming, then, localization of activity, the subject must, during training trials, tell us what was said silently on each trial. Put better, confirm that what was said was the word or phrase requested by the experimenter. A variety of machine-learning techniques can be applied to classify correctly what was silently said on the basis of the magnetic wave form recorded.

In principle, if the technique works for a very small number of words, it should be relatively straightforward to extend it to a large vocabulary, as is the case with current research on actual speech – in fact, pretty good commercial products exist. As in the case of actual speech, continuous internal speech will be much more difficult.

Of course, a different kind of machine learning may be required to learn the “meaning” of the magnetic waves. But given the abundance of current techniques, the difficulties should be not too great, granted it is technically possible to record meaningful magnetic waves.

If internal speech can be recognized, many further research questions immediately arise. Perhaps the first of great interest would be this. When we do an associative word search of memory, do the magnetic waves of the neuronal activity reflect a meaningful

use of internal speech. The conscious subjective impression is certainly affirmative, but is it so?

Here are some more questions briefly stated. Could the execution of simple arithmetic algorithms be similarly interpreted in a meaningful way from observation of the magnetic field when done purely mentally? For example, if I silently say “Nineteen plus fifteen is thirty-four”?

A large bundle of questions of a different sort concern internal conscious recognition of mental visual images and their manipulation or more detailed construction or enhancement. Internal speech recognition seems much the easier problem, but the massive work on visual pattern recognition should apply rather directly, as far as general technique of recognition is concerned.

If, like the work I did in the 70’s on speech recognition over the telephone, we can succeed with “Yes”, “No” and the ten digits, we will have taken the crucial first step. It remains to be seen if the current level of technology, and our understanding of how to use it, is up to the task.

Much has happened since 1996, and my understanding of the many complex aspects of signaling in the brain has vastly improved. The first important finding was that in the initial experiments at Scripps Institute in La Jolla with magnetoencephalography (MEG), we were fortunate that the standard electroencephalography (EEG) was also used by the placing of such sensors in the standard EEG 10-20 configuration. What we found to our surprise in our first analyses, as reported in the initial paper (Suppes, Han, and Lu 1997), was that we were able to recognize brain waves better by using the EEG results, rather than the 140 some MEG sensors. From the standpoint of continued pursuit of the research, this was a lucky break, because it is relatively easy and cheap to run EEG experiments. The technical details of MEG experiments are advanced lessons in modern superconducting devices: for example, SQUIDS must be kept at a temperature close to 0° Kelvin. In contrast, EEG research requires equipment that has not been drastically modified for more than fifty years. Moreover, it is widely used, not only in research, but in thousands, indeed hundreds of thousands of medical offices and clinics throughout the world.

So after these initial experiments, which had some success, I and my collaborators rapidly published five articles in the Proceedings of the National Academy of Sciences, of which I am a member, having been so elected in 1978. These articles marked the first burst of activity. Since then we have settled down to a more elaborate program of experimentation, with the computational analysis and modeling requiring vastly more time and effort than the experiments themselves.

This is not the place to describe the work in technical detail, which is properly reported in a series of short articles that have been appearing over the last few years, and also in Chapter 8 of Suppes (2002), but I would like to give a sense of the kind of prospects we see for the research we have embarked on. One way of describing the effort is that we are like cryptographers trying to understand how the brain encodes language and may use it internally. In support of such internal use is the surprising finding that when very simple visual images are shown to experimental subjects there is rather good evidence that within a few hundred milliseconds they will activate in the brain an image of the word corresponding to the concept. For example, if we show subjects a circle or a triangle, then this surprising finding has relevance to an old philosophical controversy. It is a controversy about how the brain or the mind represents abstract ideas, such as the general concept of a particular color or shape. I quote from our summary of this reference to the eighteenth-century dispute in Suppes, Han, Epelboim, and Lu (1999b, p. 14663).

Early in the 18th century, Bishop Berkeley (1710) famously criticized John Locke's Theory of abstract ideas (1690). David Hume (1739, p. 17) later summarized succinctly Berkeley's argument.

“A great philosopher [Berkeley] has disputed the receiv'd opinion in this particular, and has asserted, that all general ideas are nothing but particular ones, annexed to a certain term, which gives them a more extensive signification, and makes them recall upon occasion other individuals, which are similar to them.”

Berkeley's views are well supported by our results. After visual display of a patch of red or of a circle, the image is represented in the cortex by the brain wave of the word *red* or *circle* within a few hundred ms of the display and somewhat quicker than is the representation in the cortex of the spoken word *red* or *circle*. To the skeptical response that we do not really know it is the word *red* or *circle* that is being represented in the cortex, as opposed to the particular visual image, we respond that everything we have learned thus far about the one-dimensional temporal representation of words, presented either auditorily or visually, supports our inference, the spatial unidimensionality of the temporal representation used for recognition, above all. Perhaps just as important, the filtered brain waves representing the spoken color or shape words conform closely to the brain waves of the many

other words whose brain waves we have identified in our earlier work.

A second slightly earlier finding, also surprising, was that we could use brain waves of one subset of subjects to recognize the brain waves of a disjoint subset of persons. This provides evidence of an important kind of invariance in the brain-wave representations of words between individuals (Suppes, Han, Epelboim, and Lu, 1999a). Having such an invariance is certainly not something that can be derived a priori, but it greatly facilitates the efficiency of the complex system of communication of passing messages or other linguistic expressions from one person to another. If my physical procedures for processing a given word or sentence are quite different from another person's, one can anticipate it will lead to more difficulties in fast decoding of what was said by each of us to the other. It is easy to expand on this theme. It is not my purpose here to do so.

More recently (2006), in the search for very high rates of recognition, that is, rates on the order of ninety percent for at least two or three hundred individual trials, with the null hypothesis being that the recognition rates would by pure chance be only fifty percent, we have moved from words to more general auditory and visual images. We have made a big effort to find stimuli that could easily be discriminated, that is, easily discriminated in terms of the brain waves generated by subjects listening or looking at the auditory or visual stimuli. Most of this work has not yet been published. It had taken us in a direction I had not really anticipated at the beginning. In fact, our best results, the very best results, not surprisingly really, are for the binary contrast between a highly selected auditory stimulus and a highly selected visual stimulus. The outstanding instance of this so far is the auditory word *Go* juxtaposed in trials to the visual stimulus of a bright red stop-sign. One of the impressive things about these experiments is that the rate of trials can be quite fast. We randomly display one of the two stimuli every second so that in twenty minutes we already have 1200 trials. For several subjects we have tested, we have recognition rates above ninety-five percent on test trials for which no parameters of any kind have been estimated. These are very high recognition rates under classical tests of the null hypothesis. The probability of the null hypothesis being true, i.e., the p -value of the null hypothesis, is an extremely small probability and the results are highly significant. For example, it is not uncommon to have $p < 10^{-50}$ and sometimes even $p < 10^{-100}$. These are, by ordinary psychological standards, extraordinary statistical results, which are highly desirable, in order to have a firm basis for further analysis.

I end this discussion with two additional examples. The first one is about comparing stimuli of the sort referred to. On one trial the subjects sees a bright blue circle, and on the next, with the same luminance, a bright red triangle: other trials use the other two combinations, blue triangle and red circle. In analysis of these data, we are able to discriminate, if we hold shape constant, the color, and

if we hold color constant, the shape. These findings support rather strongly the psychological experiments on attention of some years ago (Treisman and Gelade, 1980). But at the same time they exhibit some contextual effects that cannot be ignored. So that, for example, which color is being held constant can make a difference. The important point is that the concept of an invariant, so nicely formulated for physical quantities like length, is more complex when the invariance is for less idealized properties of concrete physical stimuli. Some contextual effects seem inevitable, and were the source of many of the difficulties in earlier psychological research, especially in the theory of perception. What is heartening, on the other hand, is the robustness of what can still be achieved without having a simple exact invariant.

The second example is also related to many kinds of psychological and philosophical arguments of the past. Are imagined brain images, when compared to ones resulting from the presentation of physical stimuli, as in our experiments, of the same sort or different? So, for example, if I take the red stop-sign mentioned earlier and if I now ask a subject to imagine a red stop-sign, will the brain image generated be similar to the brain image generated from looking at the red stop-sign? The answer is positive. We have been able to classify the stimulus-driven brain image using the imagined brain image even though, as one would expect, the recognition rates are not as good, but still significant statistically. Second, and perhaps more important, is the earlier question, can we classify and recognize such images at all and the answer to that is again positive. We do best, of course, when comparing one imagined image, for example, of the auditory stimulus *Go* against the imagined red stop-sign, that is, imagined image compared to imagined image of another sort. The recognition rates, again, are not as good as for the strong, and as Hume would say, vivid impressions of a stimulus, but still quite significant. (In this discussion, the use of the word *image* is not meant to imply that brain images have direct perceptual qualities.)

Well, it is easy for me occupy too many pages with the discussion of our current brain research. The last examples would seem to suggest that we have given up on language, but this is not so. We think that there is still much that we can do and have hopes of being able to accomplish. One of the examples is showing a structural isomorphism between sentences and their generated brain waves. The mapping here is in terms of words to brain images of words, so that, for example, the word *Paris* in a sentence maps to a similar brain image when the word *Paris* occurs in different sentences and in different grammatical roles. It is easy to take this intuitive idea and give a formal definition of structural isomorphism for the correspondence between sentences and their brain images. Exactly the same problem arises for such a correspondence between words and their phonemic or syllabic structure, and the corresponding brain images. Here too we have some successes, but also problems to deepen our understanding. The work done so far is not yet published.

I should also mention one other approach, mainly theoretical, on which Acacio de Barros and I have spent considerable time, with as yet nothing published. In the recent work we have been joined by Gary Oas, head of physics instruction at EPGY. A great variety of theoretical and experimental research in biology and in physics shows that collections of neurons often behave as coupled electromagnetic oscillators that may be synchronized by an appropriate external stimulus, or even by attention to an internal computation or imagined brain image. We are now working on modeling these neural oscillators by weakly coupled phase oscillators satisfying Kuramoto's nonlinear equations (1984). Our initial efforts have focused on modeling with such oscillators the relatively simple but experimentally quite successful behavioral stochastic models of learning, initiated by a classic paper Estes (1950). Our current work specifically uses behavioral models and experiments from Estes (1959), Suppes and Atkinson (1960), and Suppes and Ginsburg (1963). We introduce sets of oscillators to represent sets of stimuli, other sets of oscillators to represent sets of responses, and similarly for reinforcements. Conceptually, we think of this modeling network of oscillators to be a special case of modeling more general associative networks. Some quite positive specific results are about to be published, but I shall not try at this point to describe what we have done in more detail.

It is easy to mention in closing many things we have not touched, nor has anyone else as yet successfully. Probably the most interesting example would be detailed recognition of the brain waves retrieving memories. For instance, retrieving words from long-term storage when listening to speech or reading, or even the more complicated case of producing speech, that is, of speaking.

I want to make a final philosophical remark of a general sort about our brain research. Neural experiments, as opposed to purely psychological ones on mental phenomena, are almost certain to encourage an empirical brand of nominalism. The reason is that the brain itself is not going to recognize, in any direct sense, an abstract idea or abstract concept. There must always be some physical representation. As the axiom goes in physics: all information and all computations are physical in nature. There is no nonphysical information and there is no nonphysical computation. So in order to plan anything, to think about anything, or to feel anything, concrete physical representations of brain activity are needed, just of the kind that nominalists such as Ockham, and in their own way, in the controversy mentioned earlier, Berkeley and Hume supported. The radical nature of this nominalism has not been adequately absorbed in modern philosophy of mind. Clever philosophers will find ways to work around it and to hold on to their traditional and cherished ideas. My own prediction is that as detailed work of the kind described, and ever better levels of precision and depth are achieved in the future, there will inevitably be a change. The seemingly permanent abstract

mentalese of philosophy and folk psychology will be challenged by a deep skepticism about the actual empirical validity of this traditional approach. It works only when not too many details are called for. This doesn't mean, on the other hand, that the language of folk psychology and of common sense will be abandoned. Just as we haven't abandoned the talk about ordinary objects and events like tables, chairs, and water running in the kitchen, because of all the many results of modern physics. We will retain folk psychology and ordinary mentalese, but that doesn't mean we believe they can express anything like the exact scientific account of what is going on, or that they are adequate for analyzing the structure and functions of the brain.

Philosophy of Language

During this period I have written a number of different papers in the philosophy of language, but I think I will restrict my remarks here to one sustained effort that I enjoyed a great deal and hope to continue, even though the last paper was written in 2001. This effort is a series of nine papers written on the learning of robotic natural language. Here the effort was not with learning on the part of humans, but how to teach robots to understand ordinary natural language, for example, such commands as *Put the bolt in the red box*, or *Get the screw*, *Put the nut on the screw*, *Get the large washer*, or more elaborately *Get the red screw behind the plate, which is left of the washer*.

The important feature of this work is that the robots' learning did not start absolutely from scratch. It was assumed they already had machine-language instructions for carrying out the robotic manipulations required. A demo was implemented on an "academic" but real robot by my former student and now successful entrepreneur in China, Dr. Lin Liang. What was important was learning how to interpret ordinary language in the several machine-language levels used in implementation. This is a nontrivial task. To show that we were not simply oriented toward English, and also because my two collaborators on most of this work were Li Liang who is Chinese and Michael Böttner who is German, we always did experiments in at least Chinese and German as well as English. Indeed, in one paper we studied the learning of interpretations of the machine language in ten different natural languages. What does learning an interpretation mean? Here is the answer. We had a set of formal axioms for learning that we implemented. As you might expect, they are associative in nature. Then we had axioms about changes of state in long-term memory using semantic categories and the semantically interpreted internal language already available to the robot. We also needed and had axioms for denotational learning and here we had denotational learning computations. Given the attention that has been devoted to denotation in the philosophy of language and the philosophy of mind, what is interesting here, and probably most distinctive, are the specific denotational learning axioms. Seldom, if

ever, in the ordinary philosophical enterprise have such denotational learning concepts been explicitly used and tested.

Finally we had various computations using the concept of congruence I introduced in my presidential address to the Pacific Division of the American Philosophical Association (1972), where I urged that a geometrical notion of meaning, using various strong and weak concepts of congruence, was the way to proceed without being obligated to adopt some single notion of synonymy. Such a notion of congruence is absolutely critical to semantic paraphrase and must play a role, in my judgment, in any extended development of robotic learning of natural language. All nine papers were written with the collaborators already mentioned. What I liked was the success we had in using classical concepts from psychology and from philosophy of language in developing the very specific theory embodied in the kind of axioms mentioned above.

Education and Computers

During the period running from 1979 to 1992, really all my work in this area was focused on the use of computers for instruction, what was called at that time *computer-assisted instruction* and now is described in various ways, for example, *computer-based instruction*. The important point is that there was much to be done, and I was able to round up resources to do a good deal of research on such teaching, not only in elementary mathematics where I had started back in the 1960s, but also in teaching of language, such as various levels of teaching of English from elementary-school to freshman courses in college, and especially the teaching of foreign languages. Much of the work over a considerable period was summarized in the book I edited in 1981 entitled *University-Level Computer-Assisted Instruction at Stanford: 1968–1980*. It provides an overview of what was done at the university level. In fact, this 930-page book is perhaps the most detailed analysis of such work published by anyone during the period when it began in the 1960s until the present time. Indeed, a book of this length will probably not be published in the future, because such extensive details will be reported on the web in electronic form.

Over the next few years after 1992, I published very little on the use of computers for educational purposes, but in 1992 a new effort began. This was the Education Program for Gifted Youth (EPGY) that I, with others, organized at Stanford. I have served as director since the beginning, and continue to do so today. The aim is to provide online courses for precollege students. We are not just focused on the last few years of secondary school but begin instruction in kindergarten in the case of mathematics. Building on my earlier work,—first a series of textbooks for elementary-school mathematics in the early 1960s, then corresponding computer-based courses of the same sort—, I revised the material extensively and quickly created a K-7 mathematics course for EPGY, which has

been, up until now, the course with by far the largest enrollment. I list among my publications the CDs produced for the various levels of that course until 2005. However, that CD-type of publication has now ceased, and the course is offered on the web through a browser. It is clear that future developments will be directly online on the internet.

I now have something that I particularly value that I did not have in the extensive earlier work. This is a centralized data base from all the sites on which the courses are being used, so we have, as a result, massive files that can be used to guide revisions of lectures and exercises, by analyzing student responses, and to test various psychological models of student learning and performance. The resulting publications are now principally being placed on the web site of EPGY. I look upon such electronic distribution as the primary method for this kind of publication in the future.

I have put a lot of effort into EPGY from 1992 to the present. We have extensive courses in mathematics, physics, and English for gifted students, many of whom begin as early as four-years old. But we also offer advanced courses in mathematics and physics for secondary-school students who are able first to complete a calculus course, either in their own local school or with EPGY, well before they graduate from high school, so that the last several years before college can be spent studying what are really online versions of undergraduate mathematics and physics courses at Stanford, and for which they get a Stanford University transcript.

During this period, up until 1990, I was also CEO of Computer Curriculum Corporation, a commercial company dedicated to offering computer-assisted instruction in schools. In 1990 the company was sold to Paramount Corporation, primarily know for movies but at that time also the owner of a large number of educational publishing houses. Returning actively in 1992 to the development of new computer-based courses within the framework of Stanford had academic aspects that I very much appreciate. Here I mean *academic* in the following sense. There is no real push to create something that must be profitable, and it is possible to spend much more time on research questions. Something that is perhaps even more important than anything aimed directly at published research is having the opportunity to explore uses of computers that would be considered impractical in a commercial setting. Let me just give two recent examples on which I have been working the last several years.

Beginning as long ago as the 1960s, I introduced elementary theorem proving on the computer to very bright elementary-school students in Palo Alto and other places. The proofs focused on are those for elementary theorems of arithmetic, which follow from the elementary axioms for the ordered field of rational numbers. In the beginning, only the very first parts of the natural axioms in this subject were used by the students. This was repeated in several

different forms over many years, each year increasing the number of students and often having the course used by older students. The course was then suspended for at least ten years, and I am pleased to have restarted it once again in a more elaborate form. So, we are now offering for elementary-school grades 4 to 7 a quite elaborate development of the theorems based on the rational operations of addition, multiplication, subtraction, and division, and including, for example, the definition of absolute value and the tricky theorems about absolute value and inequalities that are so useful later in proving the standard ϵ - δ theorems in analysis.

As part of this effort, we also offer what is I think one of the best features introduced earlier, but now done much more thoroughly. The students are given a great many exercises in which they are not told whether the exercise has a sentence that is to be proved by using theorems already proved, or is to be shown false by giving a counterexample. The point of the exercises is to make the students think conceptually and strategically whether a given formula can be proved or is a counterexample. If it is a counterexample, they must give specific numerical values to the variables to show that it is. Through such exercises, students learn some useful methods of problem analysis they might otherwise not. So as I write this, we are completing the new version of something begun more than forty years ago in its first version (1964). This will be the best and most complete version, and undoubtedly will have the most students using it. It will establish, in a better way than in the past, the ability of able students in elementary school to learn to give rigorous elementary mathematical proofs, with validity checked explicitly by appropriate computer programs.

The second example is connected with one of my most cherished ideas in the philosophy of language. I have already mentioned paraphrase. Both philosophers and psychologists have neglected this concept. If I hear a lecture, which I often do,—in fact one of the main pleasures and also problems of being in a place like Stanford is that there are too many lectures that I would like to hear each week, and now that I have a strong interest in neuroscience far too many—, so it is a real problem to choose what I want. In any case, when I go to one of these lectures and someone asks me afterward, “What was the lecture about?” I can, of course, not possibly repeat in serious literal detail what was said. I give a paraphrase that varies in the coarseness of its summary, and that does not use the exact words and phrases of the lecturer, but my own. This is the glory of paraphrase, one of the great syntactic and semantic features of human speech not nearly remarked upon enough.

What are the mechanisms of paraphrase, how do we do it? Well, there are many things to be said, but I will not try to go into the open research questions that I think should be of current interest. Rather, I will describe a direct application in our current EPGY teaching of elementary-school language arts, English as a second language, and

Chinese as a second language. For students taking language arts in American schools in English, or students learning English or Chinese as a second language, one problem is getting them to practice writing large numbers of sentences in English or Chinese, as the case may be, to enhance their command of the language. It is a fact of life, much regretted by many of us, that it is simply the case for a large number of reasons that students' writing of English in American schools is not evaluated very thoroughly and concretely by teachers. Certainly not as much as was often the case in the past. The same is true of foreign-language instruction. So the problem is to see how much we can do with sophisticated computer programs. The results, of course, are not going to be perfect. But the finer points of grammar are not well controlled by elementary-school teachers of language arts either. Our objective is to do at least as well and at the same time to give students extensive practice in writing. So, we have constructed a large sequence of limited environments, which we can manage rather well in terms of highly detailed analysis of what the student is doing. We present on the computer screen, in any of the kind of cases mentioned, a fixed vocabulary of between twenty to fifty words. Students can click and drag on these words to write sentences. Clicking and dragging is faster, particularly for young students, than typing. We have written very large and very specific context-free grammars for the given restricted vocabulary—which can be quite accurate because of the restriction.

We also do computations of paraphrase to check the semantic correctness of students' written answers to exercises. What students are ordinarily doing is answering questions or giving a brief description of some kind. We do not expect the student to give a unique answer. That would be inappropriate for any rich use of language. Moreover, in most cases the given vocabulary list requires students to have to write something very different from what they have read. So, for example, they must learn to use in proper anaphoric fashion the many kinds of pronominal and other anaphoric constructions important in English. The computation of a paraphrase being correct is the right semantic approach for this purpose. We have not, in any complete sense mastered the subject, but we are continually improving our computer programs and learning ourselves from the mistakes students make.

Because formal rules of paraphrase have not been widely discussed in the literature of mathematical and computational linguistics, in contrast to context-free grammars and the rules that can generate them, some brief comments at least are in order. The rules of paraphrase are, at bottom, semantic rather than grammatical in spirit, for instance, rules of deletion and generalization, which work this way.

John and Mary are jogging in the park.

Question: *Where are they?*

Paraphrased answer: *They are in the park.*

In this answer, the information about jogging was deleted as unnecessary in an adequate paraphrase.

Alice is eating a plum, and Bob is eating an apple.

Question: *What are they eating?*

Paraphrased answer: *They are eating fruit.*

These exercises are very simple, but it is easy to generate more subtle and difficult cases. Many of the main heuristic principles of paraphrase are easy to state. Here are two examples, much in the spirit of Paul Grice's maxims of conversation (1989).¹

- (1) Delete information not relevant to the question asked or point being focused on in a conversation.
- (2) Do not add information not present in the original text or speech being paraphrased.²

I'll end with what I consider is the biggest failure of technological development of educational use of computers, one that I forecast in the 1960s would be widely used by now. This is speech interaction between student and computer. Now, of course, it is easy enough to talk to the student with the kind of programs I have just been discussing and we have a great deal of audio instruction in the courses. But we do not, at present, have extensive use of speech recognition for responses of students. This, it seems to me, is a significant failure, one that we actually will be able to master once we make a sustained effort. It is a failing that is more general than just our relatively small effort at EPGY. One of the great surprises that by now (2006) operating systems for computers that include sophisticated recognition software for speech interaction have not been developed and more widely used than is actually the case. Certainly, in this century, one of the significant computer developments, from the standpoint of broad use, will be the reduction of the use of the keyboard and the natural use of the human voice to interact with hardware devices everywhere. As we do so, even the psychology of the way we think about our computers will change.

¹ The first two Grice maxims (1989, p. 26) are the following:

1. Make your contribution as informative as is required (for the current purposes of the exchange).
2. Do not make your contribution more informative than is required.

Surprisingly, Grice has only one short indexed reference to paraphrase in the 1989 volume.

² Systematic paraphrasing has a long and distinguished history that reaches back to at least late ancient times. One of the most influential examples has been Themistius' self-conscious paraphrase of Aristotle's *De Anima* and other works, written in the fourth century A.D. His paraphrastic commentary on the *De Anima* is now available in an English translation (1996).

Philosophy and Science

In what I said about the earlier part of my career under this heading, I mentioned that I had worked both in philosophy and science, though from the standpoint of research, I primarily thought of myself as a philosopher of science. I think this is probably not really the most accurate characterization. I continue to do and continue to have great interest in the philosophy of science, but it is certainly also true that, in many respects, more of my energy in the last quarter of a century has been devoted to scientific activities. A good record of these activities is to be found in my too lengthy 2002 book, *Representation and Invariance of Scientific Structures*. This large book summarizes much of the work I have done in many different areas. I think of it as written for technically minded philosophers of science, but the details go into individual sciences where I have done research at various times. The two main areas are psychology, above all, and physics.

I could take another line and say that a distinction between philosophy of science and science is in itself incorrect. In many ways I am sympathetic with such a summary of Quine's view, namely, that philosophy should mainly be philosophy of science and philosophy of science should mainly itself be science. This is a way of saying that philosophy is not privy to any special methods different from the methods used in the sciences. I certainly very much agree with these ideas, but also think that there are special aspects of problems that are of particular philosophical interest and often cultivated only by philosophers. Current philosophers of physics do not expect to develop special theories of space and time, but rather, make philosophical commentaries on the work done by physicists. It is one of the mistakes in the philosophy of mind not to have a similar attitude toward psychology.

The two big topics from psychology that I cover in the 2002 book and that reflect my perennial interests are visual perception, especially visual perception of space, and learning theory. Both of these have vast empirical and theoretical literatures, extensive already in the nineteenth century, much of the work is not really known by most psychologists and certainly by even less philosophers. Many current questions are of permanent conceptual interest, with a strong natural philosophical aspect. I will just mention some that I discuss in detail in the book. First, concerning perception, there is the complicated and subtle query of whether or not visual space is Euclidean in character. I will not even try to summarize the kind of answer I think is appropriate, but refer to Chapter 6. Another subject in psychology, treated even more extensively, is, the foundations of learning theory and, especially the question of whether or not it is possible to build from elementary concepts of learning, such as association and similarity, the complex cognitive concepts that we think of as higher order. In spite of the fact that we know very clearly that the answer in the case of mathematics to such questions is affirmative, many psychologists

refuse to believe that really at bottom its association, as I sometimes like to say, all the way down and all the way up. But consider the mathematical case. An arbitrarily complex recursive function can be computed by a universal Turing machine with a small number of states and a small number of output symbols. It is one of the mathematical triumphs of the twentieth century to show that such a reduction of computation to very simple devices is possible. That lesson has still not sunk in as far as it should, and as widely as it should, in either psychology or philosophy. These matters are discussed extensively in Chapter 8 of the 2002 book.

The two topics from physics that I know something about in detail and can therefore write about with confidence are special relativity and quantum mechanics. I include detailed discussions of the foundational literature on special relativity, but not extended to the more complicated case of general relativity. A natural foundational question is that of giving elementary axioms for special relativity and this I try to survey and give one example of. As in all such large subjects, there is extensive literature going back to the early years of the twentieth century. I am referring here not to Einstein's famous and important 1905 paper, but actually the attempts to give axiomatizations in the synthetic and qualitative spirit familiar in geometry of the same time, i.e., axioms formulated very much in the style of Hilbert's influential nineteenth-century book on foundations of geometry. The other topic is entanglement in quantum mechanics. I said enough about that already, so I need not say more here, except that I give a summary of earlier results in Chapter 7 of the 2002 book.

Although the book is long, there are many subjects, dear to my heart and about which I know something, that I had to leave out for various reasons. The most important omission was to not include a chapter on statistics following the long chapter on probability. I originally planned such a chapter, which would move from the foundations of probability (Chapter 5), and questions about probability theory in general, to the much more particular questions of statistics and analysis of data. I have spent a great deal of my academic life analyzing data and I certainly fully intended to have quite a bit to say about statistics in the 2002 book. But I kept putting off the writing of that chapter and when I finally needed it, not enough had been done to include in the actual publication. This is a shame, because one of the great developments of the twentieth century was the move from probability, thoroughly developed in many ways by nineteenth-century mathematicians starting with the magnificent work of Laplace at the end of the eighteenth century, to statistics. As a formal subject, statistics really only began in the twentieth century, but much was done during the hundred years of that century. It is still the case that not very many books in philosophy of science discuss with any thoroughness the foundations of mathematical statistics and how those foundations carry over and affect applications. Now, I certainly admit there is not some Chinese

wall between probability theory and statistics. The concepts run back and forth between the two subjects, and yet, there is still a clear and important distinction that I did not represent well.

The other big omission was not to have much more to say about statistical mechanics, which ironically is not really about statistics, in the sense just used, but an application of probability theory to one of the most important scientific cases. I said toward the end of the book that the subject was too difficult to present rigorously. I think now, in retrospect, I could have written a short chapter, certainly, not one that covered many topics, but, of course, I didn't try to do anything like that in quantum mechanics either. The only sustained argument close to the spirit of statistical mechanics was the analysis in Chapter 4 of entropy as a complete invariant of many isomorphic stochastic processes.

Maybe I don't even think that statistical mechanics is the second most important, maybe it was the omission of what I like to call the ergodic theory of free will. I have written some papers about the application of the sort of ergodic theory discussed at the end of Chapter 4 under questions of invariance, to questions of free will. I am firmly convinced that many of the tangles of compatibilism and incompatibilism can be resolved by application of beautiful and original ideas in modern ergodic theory showing, as I like to put it, the inability to distinguish between stochastic or deterministic models of many natural phenomena. It is recognized that in almost all of the important physical cases where continuous quantities are measured, there must be errors of measurement. In practice these errors of measurement are, essentially without exception, bounded away from zero no matter how good the experiments are. Given finite errors of measurement and accompanying kinds of statistics, there are beautiful proofs, especially by Ornstein and Weiss (1991), of the following sort of theorem. Consider the motion of a billiard ball with a convex object in the middle of the table to disturb its motion. For such a setting, two theories, the deterministic one of classical mechanics plus measurement errors and a stochastic Markov theory, will be mathematically inconsistent as theories of the billiard-ball motion. But we cannot distinguish between the empirical correctness of them, no matter how many observations we take. This makes as I like to say, both determinism and indeterminism in a universal sense transcendental, and creates the right kind of arena in which to discuss intentionality and free will. It is not possible to go into my analysis of such matters here, but just to express my regrets that I did not include it in the 2002 book.

Pragmatism

There is one point about philosophy and science that I have not really emphasized in any definite way, but that has become important to me. This is my much stronger explicit interest in pragmatism. In earlier years, I was put off by its superficially developed

philosophical doctrines. They seemed to be lacking in depth and, perhaps to make a joke, any serious model theory. But this was a mistake on my part. With the modern move away from foundations as an explicit aim of most philosophical work in the sciences or mathematics, I have come to see that pragmatism now fits in very well.

It is fair to say, as I have emphasized earlier in these pages, my thinking about the relationship between philosophy and science for a very long time was in terms of constructing explicit formal structures that gave a detailed sense of how a particular part of science would look when given the kind of explicit treatment characteristic of that given structures in modern mathematics. My 2002 book gave a good many examples of this. I am not against those examples now, but already as I was writing the final version, I found myself moving toward a more pragmatic view of science. I will give just two examples.

The first concerns how I ended up treating the variety of approaches to the foundations of probability in Chapter 5. This is the longest chapter in *Representation and Invariance of Scientific Structures*. I started out, when I was writing a semi-final draft of the chapter, say five years before publication, revising material from much earlier years, that I would be particularly sympathetic to a Bayesian approach. The more I got into it, the more I realized this was not really the way I now felt. One of the things that changed by mind was when I explicitly noticed that the qualitative axioms that I liked, in terms of thinking about the formal foundations of subjective probability, for instance, the qualitative axioms for a weak ordering, were not restricted to subjective ideas about probability. I could not imagine why I hadn't thought about it more clearly earlier, but, in any case, in the final version of Chapter 5, I made a number of remarks that such qualitative approaches were also very natural for objective propensity interpretations of probability. I gave in the chapter several examples. These examples were more in terms of qualitative axioms to construct a density, such as a discrete density for the geometric distribution, or the corresponding exponential distribution for continuous phenomena, as in the case of radioactive decay.

Then I found something very reinforcing. I have always liked Fred Mosteller's down-to-earth approach to statistics and the wisdom he conveys to those of us like me who are not as well educated as he is in all matters statistical. Well, I found in examining Fred's wonderful treatise with D. L. Wallace (1964/1984) on authorship of the *Federalist Papers* that it reflected, in a way that I felt extraordinarily sympathetic to, a pragmatic approach, which I summarized at the end of Chapter 5. Mosteller reports that Bayesian friends asserted that much of the analysis seemed really Bayesian, and objectivist friends said the same. So he scarcely knew how to classify the statistical approach he and Wallace used.

Reflecting on this example caused me to go back to something that I had looked at before, namely, what about the attitude of physicists to probability, especially, in that decisive case of modern physics—the probabilities that occur in quantum mechanics? So, I put in this same last section of Chapter 5 quotations from some of the most distinguished physicists who worked on quantum mechanics in the early days. Their wholly pragmatic attitude toward probability is evident. They didn't really see it as necessary, in any sense whatsoever, to make a commitment to a foundational view, but they understood very well that the computational aspects of probability were exactly what they needed for the new theoretical treatment of quantum phenomena. I'm not going to repeat here what I say there, but I am trying to give a sense of how pragmatism has more and more dominated my own thinking.

The second example, of quite a different kind, is a paper I wrote in 1999 on pragmatism in physics, where I was concerned with earlier historical episodes. I began with the history of ancient astronomy, a subject in which surely I am a rank amateur, but from which I could not resist drawing some general parallels. I felt particularly encouraged by reading Noel Swerdlow's attractive book on Babylonian astronomy (1998). I remember asking Noel, do you read the cuneiform tablets? His answer was, "Of course not." What I loved about that answer is that Noel is one of the distinguished historians of ancient and medieval astronomy. In our conversations, if I speak carelessly and make a mistake about Ptolemy, he tells me so at once. I love the fact the he had ventured into writing this excellent analysis of the Babylonian attitude to planetary motion, and yet he himself did not read the original texts. So, I tried to push a pragmatic theory, looking at the broad history from ancient astronomy to Kepler, to show how many concepts that were important to Babylonians for making omens and the like, and later, many aspects of Greek thought as well, were simply pushed out of the way and ignored. But the varied and detailed observations made by the Babylonian astronomers and used by Ptolemy more than five hundred years later, are even of some use today. Ptolemy's own central work was preserved in the tradition of a millennium and a half span leading up to Kepler and including, of course, the less important work of Copernicus. In this long period two important things were preserved: the observations reaching back to Babylonian times, and many of the Ptolemaic methods of computation, which Copernicus himself continued to use and were only changed by the new astronomy, as Kepler called what he introduced. The whole subject was then given a much greater state of perfection by Newton, with the introduction of gravitational dynamics. But much of what Kepler and Newton did rested on the shoulders of these observational and calculational giants of the distant past. It is this that is pragmatic—keeping the useful and letting go of the rest. Let me reformulate this last remark in a more purely pragmatic way. So what usually happens in the history of science is that which is true and

useful is kept, that which is false and useless is dropped. This leaves two other cases of course. What about false and useful? Well, those traditions can last a very long time, and it is easy to cite instances of something that is not literally true but is close enough as an approximation to be very useful. (Approximations are central to much important science.) And then there is the fourth case of true and useless. There are many banal claims that are true but not useful and they get lost in time, except for the arcane interests of a few benighted scholars. This is too optimistic for a good many political and social historians who do not believe in progress, but scarcely any serious historian of astronomy or physics can or does hold their pessimistic position. Even more universally accepted is the recognition of the massive progress made in astronomy and astrophysics since ancient times.

I emphasize in saying this, I am not endorsing a pragmatic account of truth as usefulness. For me they run on separate tracks that are often correlated, but one is not definable or reducible to the other.

In 1990 I received the U.S. National Medal of Science, which I was surprised and pleased to get. At that time very few of the medals of science had been awarded for work in the social and behavior sciences. I cite the brief statement attached to the award, written, I am sure, by some of my friends, because I am proud of the summary of four decades of my work, and it may provide a quick overview for some readers:

“For his broad efforts to deepen the theoretical and empirical understanding of four major areas: the measurement of subjective probability and utility in uncertain situations; the development and testing of general learning theory; the semantics and syntax of natural language; and the use of interactive computer programs for instruction.”

References

Bell, J.S. (1966). On the problem of hidden variables in quantum mechanics. *Reviews of Modern Physics* **38**: 447–52.

Berkeley, G. (1710). *Principles of human knowledge*. Dublin: Jeremy Pepyat.

Estes, W. K. (1950). Toward a statistical theory of learning. *Psychological Review* **57**: 94–107.

Estes, W. K. (1959). Component and pattern models with Markovian interpretations. In R. R. Bush and W. K. Estes (Eds.) *Studies in mathematical learning theory*. Stanford: Stanford University Press, pp. 9–52.

- Grice, P. (1989). *Studies in the way of words*. Cambridge, London: Harvard University Press.
- Holland, P. W. and P. R. Rosenbaum (1986) Conditional association and unidimensionality in monotone latent variable models. *The Annals of Statistics* **14**: 4, 1523–1543.
- Hume, D. A. (1739). *A treatise on human nature*. London: John Noon.
- Krantz, D. H., R. D. Luce, A. Tversky, and P. Suppes (1971). *Foundations of measurement, vol. I: additive and polynomial representations*. New York: Academic Press.
- Kuramoto, Y. (1984). *Chemical oscillations, waves, and turbulence*. Berlin; New York: Springer. Republished by Dover Publications, 2003.
- Locke, J. (1690). *An essay concerning human understanding*. London: Thomas Basset.
- Luce, R. D., D. H. Krantz, A. Tversky, and P. Suppes (1990). *Foundations of measurement, vol. III: representation, axiomatization, and invariance*. New York: Academic Press.
- Milonni, P. W. (1993). *The quantum vacuum: an introduction to quantum electrodynamics*. London: Academic Press, Inc.
- Mosteller, F. and D. L. Wallace (1964/1984). *Applied Bayesian and classical inference: the case of the federalist papers*. Springer Series in Statistics. New York: Springer.
- Ornstein, D. S. and B. Weiss (1991). Statistical properties of chaotic systems. *Bulletin of the American Mathematical Society (New Series)* **24**: 11–116.
- Suppes, P. (2002). *Representation and invariance of scientific structures*. Stanford: CSLI Publications.
- Suppes, P. (2006). Transitive indistinguishability and approximate measurement with standard finite ratio-scale representations. *Journal of Mathematical Psychology*, **50**: 329–336.
- Suppes P. and R. C. Atkinson (1960). *Markov learning models for multiperson interactions*. Stanford: Stanford University Press.
- Suppes P. and R. Ginsburg (1963). A fundamental property of all-or-none models, binomial distribution of responses prior to conditioning, with application to concept formation in children. *Psychological Review* **70**: 139–161.

- Suppes, P. and M. Zanotti (1981). When are probabilistic explanations possible? *Synthese* **48**: 191–199. Reprinted in *Foundations of probability with applications. Selected papers, 1974–1995*. Cambridge: Cambridge University Press, 1996.
- Suppes, P., D. H. Krantz, R. D. Luce, and A. Tversky (1989). *Foundations of measurement, vol. II: geometrical, threshold and probabilistic representations*. New York: Academic Press.
- Suppes, P., Z.-L. Lu, and B. Han (1997). Brain-wave recognition of words. *Proceedings National Academy of Sciences* **94**: 14965–14969.
- Suppes, P., B. Han, J. Epelboim, and Z.-L. Lu (1999a). Invariance between subjects of brain-wave representations of language. *Proceedings of the National Academy of Sciences USA* **96**: 12953–12958.
- Suppes, P., B. Han, J. Epelboim, and Z.-L. Lu (1999b). Invariance of brain-wave representations of simple visual images and their names. *Proceedings of the National Academy of Sciences USA* **96**: 14658–14663.
- Swerdlow, N. M. (1998). *The Babylonian theory of the planets*. Princeton: Princeton University Press.
- Themistius (350 A.D./1996). *On Aristotle's On the Soul*. Translated by Robert B. Todd. New York: Cornell University Press.
- Treisman, A. M. and G. Gelade (1990). A feature-integration theory of attention. *Cognitive Psychology* **12**: 97–136.
- Zeilinger, A., (2005). The message of the quantum. *Nature* **438**: 743.