A BACKWARD LOOK TO THE FUTURE

E. T. Jaynes
Arthur Holly Compton Laboratory of Physics
Washington University
1 Brookings Drive
St. Louis MO 63130, USA

ABSTRACT. We survey briefly some fifty years of thinking about physics and probability with the aim of explaining: (1) What I did not know then, but know now; (2) What I have been trying to accomplish in science and education, and to what extent these efforts have succeeded; (3) What remains unfinished, but where I think the greatest future opportunities lie; and (4) What personal and professional advice I can now give to young people (and wish someone had given me fifty years ago).

Introduction

A meeting like this is an overwhelming experience! I was overwhelmed not only by the sheer number of people who came here from so far; but even more by the kind sentiments expressed. Of course, I had looked forward to seeing again many former students and colleagues and had expected to have chats with each one, to renew our friendship and bring us both up to date about the other's work. But what one can actually do in two days is so helplessly short of what one wants to do! Some of the participants had to leave without our being able to talk at all; there was simply no time for it. Then perhaps this reminiscence can serve as a substitute channel for conveying my thanks and appreciation to all of you.[†]

But a reminiscence may be boring unless it also conveys something useful to the reader. It must be that fifty years of thinking about science and of observing the successes and failures of many ideas, has given one *some* kind of insight that would be helpful to others. Whilst exuding long-term optimism, the following remarks will appear at some places to be of a negative character, deploring recent trends in both science and education. So please understand that my purpose is not to complain, but rather to seek constructive remedies; one must first know what the problems are. The problems created by the blindness of people of my generation are also the opportunities for the next generation, so my advice for the future has necessarily some warnings about recent trends that need to be resisted and reversed. Those who are at the beginning of their careers today will be obliged to deal with these problems; and it is better to be aware of them now than to learn about them gradually over many years, as I did.

[†] In writing this a peculiar problem arose; the constant use of the personal pronoun "I" which one is not supposed to do, and which looks awkward even to me. But since that happens to be the topic here, I did not see any way to avoid it which does not look even more awkward; so I apologize for this in advance and hope it will not distract you enough to obscure the message.

In this I have only followed what seemed to me a good precedent. When my thesis advisor, Eugene Wigner, retired in 1971, he proceeded to unburden himself, in a series of colloquium talks at many places, of some concerns about science and education that had bothered him for many years. But he felt that he should keep his silence until then. So there is also in the following some of the flavor of "*Now it can be told*!"

Teaching or Research?

This has been one of my major concerns throughout my adult life. We have a serious problem in American science education, for which no realistic solution is yet in sight.

I was always much aware of Einstein's advice to a young theoretical physicist: if you are sure that your ideas are right, and therefore you want to be able to work full time on your research without interruptions, get a job as a light-house keeper. On the other hand, if you are not that confident, then as a matter of insurance get a job as a teacher. Then if the research does not pay off, at least you will have the satisfaction of knowing that you are making a useful contribution to society. The fact that, on entering the job market in 1950 I did not seek out a lighthouse and did get a job as an instructor at Stanford University, perhaps indicates my early level of self-confidence.

But it required a few years before I perceived what a science teacher's job really is. The goal should be, not to implant in the student's mind every fact that the teacher knows now; but rather to implant a *way of thinking* that will enable the student, in the future, to learn in one year what the teacher learned in two years. Only in that way can we continue to advance from one generation to the next.

As I came to realize this, my style in teaching changed from giving a smattering of dozens of isolated details, to analyzing only a few problems, but in some real depth. It doesn't even matter very much what those few problems are; once a student knows what it feels like to analyze something in depth, he can do it for himself on whatever other problems may come his way. Equally important, he can recognize in the work of others the distinction between a superficial study and one that is deep enough to be capable of finding new things.[‡]

For Centuries, Universities recognized only the teaching function and "research" was an unknown word. Teachers were not an economic liability to the University; they were paid directly by the students and their income reflected directly the quality of their teaching. So every teacher was under a strong incentive to do better at his job. In the United States, research gained a foothold in the early 20'th Century (in England, about 50 years earlier), and the issue of teaching *vs.* research has been a matter of contention here ever since. My impression is that this is much less of a problem in England.

The following scenario was repeated in essentially the same form at many American Universities; but it was at Stanford that I experienced it first hand. The researchers proceeded to gain influence very aggressively; they created the folklore that a person cannot be a good teacher unless he is first a good researcher. We all know this to be true, necessarily, of the most advanced graduate courses at the boundary of present knowledge; but it is egregiously false for all the other courses where most of the substantive science education

[‡] As a pertinent example, one sees through expositions of Quantum Electrodynamics which proclaim its wondrous success on the basis of a few accurate numbers – without ever asking what aspect of the formalism led to those numbers, or what changes in the formalism would change those numbers. We return to this later.

takes place. Nevertheless, the idea spread among administrators, and in about 1920 the policy was adopted, very quietly, at several Universities including Stanford, that henceforth faculty appointments and promotions would be based solely on research productivity, not on teaching ability or performance.

Thirty years later this news finally leaked out to the undergraduate students at Stanford and, as one would guess, caused an uproar in the Student Newspaper in the 1950's. I wrote a private memorandum to the Physics faculty expressing my own views, observing in particular that good research can be and is done in many other places (Government and Industrial laboratories, Theoretical Institutes); but that if good teaching is not done in our Universities, it does not get done at all.

The Administration responded to the situation by issuing a public statement denying that any such policy existed on promotions. David L. Webster, who had been Chairman of the Stanford Physics Department at the time this policy was instituted, knew better; and without taking part in the local fracas, he revealed the facts for the record in an article which was ostensibly only a reminiscence of the early days of the American Physical Society (Webster, 1957), in which he quoted a few lines from my memorandum. The physics students were quick to seize upon this and spread it over the campus.

Had the matter been left to internal forces in the University, there would have been a movement back to a condition of something like balance between the factions. But just at this time, what an economist would call an "exogenous variable" appeared and took control of the situation. It was the really big research grant sponsored by various alphabetized Government agencies (ONR, AFOSR, AEC, NSF, DOD, *etc.*). This created the situation that a researcher with the right connections – whether competent or not – paid his own way many times over in the money he brought into the University; while a good teacher became, from the standpoint of the Administration, an economic liability. It was easier to get a million dollars for a big research project than it was to get a thousand dollars for upgrading of the educational facilities. So the researcher was back in firmer control than ever; every faculty member was under constant pressure to push his research and neglect his teaching, under pain of jeopardizing his professional career: *Publish or Perish.*

Of the two biggest spenders on research, one offended our teachers by dismissing teaching as a mere "chore" in a memorandum. At a faculty meeting that I attended, the other actually complained out loud that there were far too many courses and other requirements, and they were interfering with the time that students could spend working on HIS research! As if graduate students were there only as a source of cheap labor to further his ambitions. Seeing that there was some opposition to it, the strategy adopted by the big researchers was to accomplish their goal gradually; there was a steady stream of small changes in the Departmental rules, each one cutting down a little more on the amount of course work, language requirements, and the level of general competence required of graduate students. They received less and less education while spending more and more hours manning the vacuum pumps. I protested against this every step of the way; and was outvoted every time.

This Government intervention (always, of course, held to be essential for National Defense and beneficial to the University) has through this unforeseen side effect done more damage to the cause of advanced science education in the United States than has any other factor in the past 40 years. In my opinion, the University science teaching profession is an honorable and socially necessary one; and no good teacher needs to apologize if he does not also turn out volumes of research. Indeed, the quality of our education today determines the quality of our research tomorrow. It is essential that our Universities carry on both functions, at the highest possible level, simultaneously; and some method of administration and financing must be found that will permit this.

But Why The Move?

When in March 1960 I received an inquiry as to whether I might be interested in a faculty position at Washington University, I visited St. Louis and met Edward U. Condon, the Physics Dept. Chairman. We had a long, serious conversation deep into the night, sitting at his kitchen table, in which I told him – in detail, and naming names – of my dismay at what seemed to me the destructive policies in effect at Stanford. He agreed with my judgment, assured me that he and others at Washington University were well aware of this problem and were determined to resist it; and that no faculty member at W. U. would be placing his professional career in jeopardy by being a good teacher.

Ed Condon was the first Academic administrator I had ever known who was a big enough man to think in terms of the long-run interests of the students and the University, instead of the short-run finances of his own little bailiwick, and was currently in office. I was so impressed that I accepted the position he offered me for the privilege of working with him, even though it meant a loss of income, and moved to St. Louis in September 1960. (But over the next ten years I managed to attract to W. U. about as much research money as anybody else, plus the money that built our present library.)

A year after this move it was quite obvious to me that – in agreement with the general observations of David Webster – Washington University was doing a better job of science education than was Stanford; and I think the record of subsequent accomplishments of our students of that period will bear this out. To my great satisfaction, many of our students became excellent teachers and turned out important research, and a surprising number proceeded quickly to become Department Heads, Deans, and on to higher positions in Academia, Government, and Industry.

Similarly, for many years Reed College in Portland, Oregon was famous for its dedicated undergraduate physics teaching; and it contributed a disproportionate number of the Ph.D. physicists in the United States. The long-run effects of good and indifferent science teaching are not hard to discern.

Today this does not seem to be quite the front-burner issue that it was thirty years ago, but perhaps this signifies only that the dedicated science teachers have been vanquished entirely, and no longer exist at all in major Universities; those who are concerned with the alarming decrease in number of students choosing science majors please take note. In any event, the same factors are in operation and there is nothing obsolete in the maxim: *Publish or Perish*. The pages of our scientific journals continue to be cluttered with bad research whose only function is to gain someone a promotion to Associate Professor.

It seems to me that this was a tragedy that did not have to happen, and is wasteful to everyone concerned. Of course, there are natural differences in talent and inclination of scientists. A few are outstandingly successful at both teaching and research, but most of us tend to do better at one or the other; and we ought to be free to choose. Similar divisions of labor occur naturally in nearly every field. For example, among musicians the talents for composing and for performing seldom occur in the same person, and most tend toward one activity or the other. There is no particular stress between these different "factions" in

Music Departments, because they are all coexisting in the same circumstances and see their functions as complementary rather than in competition. But such stress could be created overnight if some outside agency decided to bestow lavish support only on performers but not on composers; *Perform or Perish* would become, very quickly, the operative principle.

Having recognized the problem, what is the solution? In the future, some readers of these words are going to find themselves in Academic Administrative positions and will face this question. The obvious, simplistic answers are to establish quotas (which nobody really wants), to build up a separate endowment dedicated to support of teaching (which is unrealistic); or to remove the dependence on Government grants (which is even more unrealistic). The latter would seem to require that Universities build up their endowments to the point where they are financially independent, and no longer accept Government grants. But unless the University became so wealthy that money was no consideration at all, the experimental researcher would then become a major economic liability and we would have the same problem in reverse. So here is an opportunity for some really new creative thinking. In the meantime, we can only hope that Universities may find Administrators with the wisdom of David Webster and Ed Condon.

Probability Theory

Turning to technical matters, several people have noted that each of my research papers goes through a twenty-year incubation period, during which it is either ignored or attacked; then somebody starts to take note of what is in it, and it becomes studied intensely. My first work on statistical mechanics appeared in 1957; in 1977 philosophers started their nit-picking analysis and commentary about the meaning of every word in it. The Jaynes-Cummings paper appeared in 1963 and was ignored until 1983, when it suddenly became the most cited work in the quantum optics literature. My exposition of the "fast second law" with quantitative applications in biology appeared only in 1989, so we cannot expect anybody to notice it until the year 2009, while my resolution of the Gibbs paradox is not scheduled to be discovered until the year 2013; and so we need not discuss them here.

Evidently, if we want to find any gratifying successes, we must turn to my earliest works, where more than twenty years have passed. Indeed, today I am far more self-confident about probability theory than was possible forty years ago when I started lecturing on probability and information theory. The reason is that the essential evidence is now in; we are sustained no longer by faith and hope, but by proven theorems and accomplished facts on the level of new useful numerical results. There is no doubt that the formulations of probability theory and statistical mechanics as extensions of logic are here to stay; they will be the universally accepted basis of those fields 100 years hence. Too many things are coming out right to allow any other outcome.

Without giving a lecture course, let me convey a quick impression of the nature of these developments, as they stand today. Theoretically, everything has a single very simple rationale and mathematical form, which makes it easy to teach. The basic rules, from which all else follows, are nothing but the standard product and sum rules of probability theory; but now they are derived uniquely as principles of logic, from very elementary qualitative requirements of consistency and rationality.

Functionally, probability theory as extended logic includes as special cases all the results of the conventional "random variable" theory, and it extends the applications to useful solution of many problems previously considered to be outside the realm of probability theory. The now much strengthened theoretical foundation and continued pragmatic success – and the failure of critics to uncover any defects in it or offer any usable alternatives – justify this confidence in it.*

When applied to problems of parameter estimation or hypothesis testing, probability theory as logic is generally called *Bayesian inference*, on historical grounds explained elsewhere, and it is accomplishing a major house-cleaning in the field of statistics. The "orthodox" methods of inference as taught by statisticians since the 1930's consist of about a dozen intuitive devices (confidence intervals, unbiased estimators, significance tests, *etc.*), without any connected theoretical basis. Each is usable in some small range of problems for which it was invented; but each produces contradictions and absurd conclusions when applied out of its proper range. Now all of these are basically methods for reasoning from incomplete information; that is, for information processing. Yet the orthodox practitioners never thought of probability in terms of information.

Orthodox methods of inference also faced insuperable technical difficulties (nuisance parameters, lack of sufficient or ancillary statistics, *etc.*) which made it impossible to extract all the information from one's data in many problems. Indeed, some textbook authors recommended procedures which amounted to throwing away practically all the relevant information in the data, merely to achieve an unbiased estimate.† When there is cumulative error, an author at the Bureau of Standards advocated estimating the slope of a linear relation from the first and last measurements alone – thus throwing away all the evidence of the intermediate data points – which is easily shown to lead to an order of magnitude less accuracy in the estimate.

When these *ad hoc* devices are replaced by their Bayesian counterparts, all these difficulties disappear. The resulting algorithms have no restricted range of validity; you can apply them to arbitrary extreme conditions and the results continue to make sense. This confirms the theoretical expectation; Bayesian methods are the exact rules for conducting inference, while orthodox methods are only intuitive approximations to them. One expects that an approximation will be valid only in some restricted domain.

In Bayesian inference, nuisance parameters are not only eliminated effortlessly from our final algorithms; now, instead of creating problems, their use becomes a means for improving the precision of our results. In effect, this procedure warns probability theory to be on the lookout for some disturbing but uninteresting effect contaminating the data, and make proper allowances for it. The disturbing effect is not treated as mere "noise"; to do so would be to lose all the highly cogent information we have about its nature.‡

---

* Every objection to this formulation that we have seen arises, very obviously, from the fact that the critic has completely misunderstood what the ideas are; and so is attacking a straw man of his own making. The situation is much like that of Darwin's Theory of Evolution; as the well-known biologist Stephen J. Gould has noted recently, those who still attack Darwin's theory only reveal that they do not understand what the theory is.

† Yet, as R. A. Fisher emphasized already sixty years ago, the criterion of bias has no theoretical justification (it is not even invariant under a change of variables), and a biased estimate can be far more accurate than an unbiased one.

‡ For example, if we want to estimate only the frequency of a sine wave plus noise, then the amplitude and phase are nuisance parameters, which affect the data but whose effects must be eliminated somehow in the final algorithm for processing the data into an estimate of the frequency. Then any prior information about how the amplitude might be varying with time becomes highly cogent for determining the accuracy of our frequency estimates.

New insight into these matters has continued to the present day. A major advance in our understanding and technique occurred five years ago in the thesis of Larry Bretthorst, when it was found that by introducing a nuisance parameter and integrating it out again, one can in some cases achieve orders of magnitude improved resolution in spectrum analysis, over the previous practice of taking the fourier transform of the data. This is having its first impact in analysis of NMR free induction decay data, where the decay rate no longer places the limitation on the accuracy with which we can determine the oscillation frequency; we expect that other applications will profit from it just as much.

Orthodox methods of inference had another distressing property: on the one hand, to get conclusions from data they offer no way to take into account our prior information about the parameters of interest; yet on the other hand they require us to make additional assumptions about the frequency distribution of errors that are arbitrary; that is, not justified by any of our information. In contrast, it is now a proven theorem that, when we apply strict Bayesian principles with initial probabilities assigned by maximum entropy, our subsequent inferences depend only on the data and the circumstances (maximum entropy constraints) about which we had prior information; there is no room for gratuitous assumptions because circumstances about which we have no information automatically cancel out and contribute nothing to our final conclusions. Those simple product and sum rules have concealed, for all these years, a remarkable power and efficiency for information processing.

This is an example of a property, only recently discovered, which is making Bayesian inference exciting today. Because of it we now understand a fact that has been puzzling to workers in probability theory since Augustus de Morgan discovered it in 1838; that Gaussian sampling distributions almost always lead to the most successful inferences *whether or not the frequency distributions of our errors are in fact Gaussian*. To see why this is true it was necessary to think of probability theory as extended logic, because then probability distributions are justified in terms of their demonstrable information content, rather than their imagined – and as it now turns out, irrelevant – frequency connections.*

But in addition to this great power, Bayesian methods are also very simple in principle and easy to apply; therefore they enable us to find useful solutions to problems of reasoning that are so complex that they could not even be formulated in orthodox terms. For many examples, with fully worked-out numerical details, see Bretthorst (1988). I am now trying desperately to complete a large book, *Probability Theory: The Logic of Science*, which is to have all the theoretical background of this, plus many more applications and comparisons with orthodox results. Many of you have seen some preliminary fragments of it, which are issued to interested persons from time to time, as they become available.

New Adhockeries

In recent years the orthodox habit of inventing intuitive devices rather than appealing to any connected theoretical principles has been extended to new problems in a way that makes it appear at first that several new fields of science have been created. Yet all of them are concerned with reasoning from incomplete information; and we believe that we have theorems establishing that probability theory as logic is the *general* means of dealing with all such problems. We note three examples.

* The point here – perfectly obvious in retrospect but not noticed for 150 years – is that the actual distribution of the errors is known from the data; whatever distribution we might have expected before seeing the data is made irrelevant by that information.

Fuzzy Sets are – quite obviously, to anyone trained in Bayesian inference – crude approximations to Bayesian prior probabilities. They were created only because their practitioners persisted in thinking of probability in terms of a "randomness" supposed to exist in Nature but never well defined; and so concluded that probability theory is not applicable to such problems. As soon as one recognizes probability as *the general way to specify incomplete information*, the reason for introducing Fuzzy Sets disappears.

Likewise, much of Artificial Intelligence (AI) is a collection of intuitive devices for reasoning from incomplete information which, like the older ones of orthodox statistics, are approximations to Bayesian methods and usable in some restricted class of problems; but which yield absurd conclusions when we try to apply them to problems outside that class. Again, its practitioners are caught in this only because they continue to think of probability as representing a physical "randomness" instead of incomplete information. In Bayesian inference all those results are contained automatically – and rather trivially – without any limitation to a restricted class of problems.

The great new development is Neural Nets, meaning a system of algorithms with the wonderful new property that they are, like the human brain, *adaptive* so that they can learn from past errors and correct themselves automatically (WOW! What a great new idea!). Indeed, we are not surprised to see that Neural Nets are actually highly useful in many applications; more so than Fuzzy Sets or AI. However, present neural nets have two practical shortcomings; (a) They yield an output determined by the present input plus the past training information. This output is really an *estimate* of the proper response, based on all the information at hand, but it gives no indication of its accuracy, and so it does not tell us how close we are to the goal (that is, how much more training is needed); (b) When nonlinear response is called for, one appeals to an internally stored standard "sigmoid" nonlinear function, which with various amplifications and linear mixtures can be made to approximate, to some degree, the true nonlinear function.

But, do we really need to point out that (1) Any procedure which is adaptive is, by definition, a means of taking into account incomplete information; (2) Bayes' theorem is precisely the mother of all adaptive procedures; the *general* rule for updating any state of knowledge to take account of new information; (3) When these problems are formulated in Bayesian terms, a single calculation automatically yields both the best estimate and its accuracy; (4) If nonlinearity is called for, Bayes' theorem automatically generates the exact nonlinear function called for by the problem, instead of trying to construct an approximation to it by another *ad hoc* device.

In other words, we contend that these are not new fields at all; only false starts. If one formulates all such problems by the standard Bayesian prescription, one has automatically all their useful results in improved form. The difficulties people seem to have in comprehending this are all examples of the same failure to conceptualize the relation between the abstract mathematics and the real world. As soon as we recognize that probabilities do not describe reality – only our information about reality – the gates are wide open to the optimal solution of problems of reasoning from that information.

Quantum Theory

Here no such gratifying successes can be reported, because the fate of my ideas about quantum theory and electrodynamics will be determined, necessarily, by Nature and not by any arguments invented by me or by anybody else. There is no way any of us can feel very confident of what the future may bring here. Nevertheless, we do have a little more than faith and hope to sustain us. I am convinced, as were Einstein and Schrödinger, that the major obstacle that has prevented any real progress in our understanding of Nature since 1927, is the Copenhagen Interpretation of Quantum Theory. This theory is now 65 years old, it has long since ceased to be productive, and it is time for its retirement (along with mine).

Just for that reason, this is where the great opportunities for the next generation lie. Let us examine the logical impasse now facing us, and see why I think the aforementioned Bayesian principles may be the key to resolving it, just as they did in cases just cited. But here the situation is far more complex and subtle, so unlike the above relatively trivial examples, it will require much more deep thinking to see exactly how to carry this out.

The more its defenders insist that all is well and there are no contradictions in present quantum theory, the more blatantly those contradictions stare us in the face and tie our hands, making it impossible to proceed. We are familiar with a proposition of elementary logic: that from a false proposition all propositions, true and false, may be deduced. There is a corollary, noted by Källén (1972): from an inconsistent theory any result may be derived. Indeed, we know that to get correct predictions out of Quantum Electrodynamics (QED) required a great deal of art and tact, found only after twenty years of efforts (1927-1946); for the right experimental numbers to emerge one must do the calculation (*i.e.*, subtract off the infinities) in one particular way and not in some other way that appears in principle equally valid.

In space-time, the Feynman propagators have violent singularities on the light-cone, far worse than the delta functions that arise in the Green's functions in other parts of mathematical physics. They guarantee that any integral of the form $\int S_F(x-y)\, f(y)\, d^4y$ which might be thought of as a first-order perturbation, diverges if in the integration the separation $(x-y)$ crosses the light-cone. Surely, in a properly formulated theory, there would be no infinities to subtract; yet another 45 years of efforts have not found that formulation – except perhaps in Schwinger's source theory, which seems to be ignored by workers in the field.* Then in what sense can one claim that QED is a great success?

But this is not limited to field theory. As we have noted in some detail elsewhere (Jaynes, 1991), throughout the history of quantum theory, whenever we advanced to a new application it was necessary to repeat this trial-and-error experimentation to find out which method of calculation gives the right answers. Then, of course, our textbooks present only the successful procedure as if it followed from general principles; and do not mention the actual process by which it was found. In relativity theory one deduces the computational algorithm from the general principles. In quantum theory, the logic is just the opposite; one chooses the principle to fit the empirically successful algorithm.

But after all, how can one build rationally from a theory whose basic principles are in this condition: Present quantum theory uses relativistic wave equations, but tries to solve

* Schwinger's recent writings have a cryptic character, almost as if he were trying to conceal what is really happening. However, we think that he achieves the correct finite numbers by his use of finite sources with time-like separation. Then with x in the future source, y in the past one, his regions of integration never reach the light-cone. If so, then we see at once that the singularities in the Feynman propagators are actually contributing nothing to the finite experimental numbers of the theory; only the finite terms inside the light cone will be needed in the future correct theory. This speculation makes such good physical sense that we hope it turns out to be true.

them with propagators that – quite aside from the divergences – violate relativity by failing to vanish outside the light-cone, and run backward in time! What can this possibly mean?

On a more elementary level, present quantum theory claims on the one hand that local microevents have no physical causes, only probability laws; but at the same time admits (from the EPR paradox) instantaneous action at a distance! Today we have in full flower the blatant, spooky contradictions that Einstein foresaw and warned us about 60 years ago, and there is no way to reason logically from them. This mysticism *must* be replaced by a physical interpretation that restores the possibility of thinking rationally about the world.

We see the effects of this in the fact that today, a large portion of research in theoretical physics has been reduced to wheel-spinning; random fiddling with the mathematics of the old theory, without giving a thought to its physical foundations. One would think that the folly of this might have been learned from the example of Einstein; yet his repeated warnings go unheeded even as his worst fears are realized before our eyes.

I believe the answer to this must be that our present formalism contains two different things. It represents in part properties of the real world, in part our information about the world; but all scrambled up so that we do not see how to disentangle them. But at least we can see that the spooky things will cease to be spooky as soon as we think of the formalism in terms of inference from incomplete information. Then what is traveling faster than light, and backwards in time, is not a physical influence; but only a logical inference. David Hestenes thinks that his reformulation of the Dirac equation accomplishes this separation into the subjective and objective features of the theory; in our view this is an attractive possibility, but not yet a demonstrated fact. Before we could judge this, one needs to work out the full, explicit solutions to the standard QED problems (Spontaneous Emission, Lamb Shift, Vacuum Polarization, Anomalous Moment, Bethe-Heitler formula, Mott scattering, *etc.*) in the Hestenes formalism and see what this interpretation is saying at every step of the derivations.

What has held up progress in this field for so long? Always our students are indoctrinated about the great pragmatic success of the quantum formalism – with the conclusion that the Copenhagen interpretation of that formalism must be correct. This is the logic of the Quantum Syllogism:

> The present *mathematical formalism* can be made to reproduce many experimental facts very accurately.
>
> **Therefore**
>
> The *physical interpretation* which Niels Bohr tried to associate with it must be true; and it is naïve to try to circumvent it.

Compare this with the Pre-Copernican Syllogism:

> The mathematical system of epicycles can be made to reproduce the motions of the planets very accurately.
>
> **Therefore**
>
> The theological arguments for the necessity of epicycles as the only perfect motions must be true, and it is heresy to try to circumvent them.

In what way are they different? The difference is only that today everybody knows what is wrong with the Pre-Copernican syllogism; but (from the frequency with which it is still repeated) only a relatively few have yet perceived the error in the Quantum Syllogism. In quantum theory, we are still using pre-scientific standards of logic in a dozen different places, as noted in more detail elsewhere (Jaynes, 1989, 1990).

Both Quantum Electrodynamics (QED) and my "neoclassical" theory (Jaynes, 1973) contain some elements of truth, but also some elements of intolerable nonsense, and thus far nobody – least of all, me – has seen how to unscramble them. But there is great new hope in the fact that their elements of truth and nonsense are so different, and in the spectacular recent advances in experimental technique reported at this meeting by Herbert Walther and Pierre Meystre, which are now making direct contact with these issues.

From a physical rather than formal standpoint, the worst elements of nonsense in QED concern the infinite zero-point energy and vacuum fluctuations. We have given physical arguments (Jaynes, 1990) which seem conclusive to me, showing that (a) these fluctuations are not real and are not necessary to account for the experimental facts; and (b) their removal from QED would also remove at least some of the infinities that have plagued the theory from the start, and a great deal of surplus formalism not actually used.

The underlying idea of neoclassical theory (NCT) is that photons are not real physical objects existing in the electromagnetic field; all the field phenomena of propagation, diffraction, and interference are accounted for correctly by the classical EM theory of Maxwell, Poynting, Lorentz, Larmor, and Poincaré. The alleged particle-like properties of radiation are consequences only of the laws of interaction with matter, and appear only in the reaction of matter to the fields. Those who do not believe this continue to be astonished at the fact that, whenever a new interference effect is observed, it is always puzzling and mysterious from the standpoint of quantum theory; but it is just what classical EM theory would have predicted. Another example of this was presented at this meeting by Leonard Mandel, in which the QED uncertainty is just that due to the unknown phase of the classical idler oscillations.

The worst elements of nonsense in NCT concern the apparently different physical status of longitudinal and transverse fields, although they are partly interconverted by a Lorentz transformation.† In addition it faces unsolved computational problems in dealing with free particles [which we view as stable wave packets, held together by the self-interaction forces of the Dirac equation as discussed in Jaynes (1991), that are ignored in the "spreading wave packet" solutions of elementary quantum theory], and really works nicely only for bound states of atoms, where we know the wave functions quite accurately from previous solutions of the Schrödinger or Dirac equation (because then the Coulomb forces are large compared to the self-interaction forces). Some critics have claimed that NCT gives some field correlation effects wrong; I think they have misapplied it and suggest that, as noted in Jaynes (1973), it is QED that gives some field correlation effects wrong (For that work, the mandatory twenty year incubation period is nearly over; we expect that in 1993 somebody will start to take note of what is in it).

But NCT also contains one bit of fundamental truth that is not in QED. This is the simple explanation of the basic reason for the relation $E = h\nu$, which is in present quantum theory only an empirical relation for which – astonishingly – nobody seems ever to have

† This may be a removable difficulty, because NCT is still far from being cast permanently in concrete; many modifications are still possible while preserving the basic ideas.

sought any theoretical explanation.‡ We found part of the answer in that this relation follows from an action conservation law of the Hamiltonian; and that, viewed in this way, it need not be a universally valid principle, so there is a possibility of an experimental test.

Presumably, that elusive "future correct theory" of electrodynamics, toward which we have been groping for a Century, will be some kind of mixture of these theories and others, but with some great unifying principle not yet seen (but which will be perfectly obvious to all as soon as it is seen). Surely, it will have to recognize this element of truth, and find some way of incorporating it into quantum theory, while discarding the elements of nonsense of both theories.

In the meantime, those who apply the naïve ideas of present quantum theory confidently to astrophysical problems – conditions extrapolated dozens of orders of magnitude beyond anything that we know to be true – are in my view engaging in speculation that is not necessarily wrong in aim, but is just a Century premature. Attempts at unified force theories also seem premature for the same reason; until we get elementary quantum theory into some kind of rational order, unrestrained speculation from it is highly unlikely to lead to the truth.

One of the principles of scientific inference – which has always been well understood by the greatest scientists – is that it is idle to raise questions prematurely, when they cannot be answered with the resources available. For Isaac Newton it would have been foolish to raise questions that were not foolish for Erwin Schrödinger 250 years later; for Gregor Mendel it would have been foolish to raise questions that were not foolish for Francis Crick 100 years later. By "foolish" we mean "without hope of success". Of course, we all enjoy indulging in a little free speculation about the future of science; but for scientists to expend their serious professional time and effort on idle speculation can only delay any real progress.

Dealing with Critics

Looking back over the past forty years, I can see that the greatest mistake I made was to listen to the advice of people who were opposed to my efforts. Just at the peak of my powers I lost several irreplaceable years because I allowed myself to become discouraged by the constant stream of criticism from the Establishment, that descended upon everything I did. I have never – except in the past few years – had the slightest encouragement from others to pursue my work; the drive to do it had to come entirely from within me. The result was that my contributions to probability theory were delayed by about a decade, and my potential contributions to electrodynamics – whatever they might have been – are probably lost forever.

But I can now see that all of this criticism was based on misunderstanding or ideology. My perceived sin was not in my logic or mathematics; it was that I did not subscribe to the dogmas emanating from Copenhagen and Rothamsted. Yet I submit that breaking those dogmas was the necessary prerequisite to making any further progress in quantum theory and probability theory. If not in my way, then necessarily in some other.

‡ Contrast this with the reasoning of Einstein, when he noted that the equality of gravitational mass and inertial mass was only an empirical relation without any theoretical explanation; in 1946 he wrote: "It is, however, clear that science is fully justified in assigning such a numerical equality only after this is reduced to an equality of the real nature of the two concepts." To find that explanation he was led to General Relativity. Neoclassical theory was made in quite conscious imitation of Einstein's example; we expect that a full understanding of $E = h\nu$ will be of just as fundamental a nature as is General Relativity, far beyond anything in present quantum theory.

In any field, the Establishment is not seeking the truth, because it is composed of those who, having found part of it yesterday, believe that they are in possession of all of it today. Progress requires the introduction, not just of new mathematics which is always tolerated by the Establishment; but new conceptual ideas which are *necessarily* different from those held by the Establishment (for, if the ideas of the Establishment were sufficient to lead to further progress, that progress would have been made).

Therefore, to anyone who has new ideas of a currently unconventional kind, I want to give this advice, in the strongest possible terms: *Do not allow yourself to be discouraged or deflected from your course by negative criticisms* – particularly those that were invented for the sole purpose of discouraging you – unless they exhibit some clear and specific error of reasoning or conflict with experiment. Unless they can do this, your critics are almost certainly wrong, but to reply by trying to show exactly where and why they are wrong would be wasted effort which would not convince your critics and would only keep you from the far more important, constructive things that you might have accomplished in the same time. Let others deal with them; if you allow your enemies to direct your work, then they have won after all.

Although the arguments of your critics are almost certainly wrong, they will retain just enough plausibility in the minds of some to maintain a place for them in the realm of controversy; that is just a fact of life that you must accept as the price of doing creative work. Take comfort in the historical record, which shows that no creative person has ever been able to escape this; the more fundamental the new idea, the more bitter the controversy it will stir up. Newton, Darwin, Boltzmann, Pasteur, Einstein, Wegener were all embroiled in this. Newton wrote in 1676: "*I see a man must either resolve to put out nothing new, or become a slave to defend it.*" Throughout his lifetime, Alfred Wegener received nothing but attacks on his ideas; yet he was right and today those ideas are the foundation of geophysics. We revere the names of James Clerk Maxwell and J. Willard Gibbs; yet their work was never fully appreciated in their lifetimes, and even today it is still, like that of Darwin, under attack by persons who, after a Century, have not yet comprehended their message (Atkins, 1986).

The recent reminiscences of Francis Crick (1988) make many other important points that we might otherwise have included here, because they apply as well to physics as to biology. For example, he notes that in studying brain function, theoretical work "tended to fall into a number of somewhat separate schools, each of which was rather reluctant to quote the work of the others. This is usually characteristic of a subject that is not producing any definite conclusions." Exactly the same could be said of several areas of physics; we would add only that, even when definite conclusions are proclaimed by one school, they are not often justified or permanent.

Conclusion

From all of the above discussion one generalization stands out: today we are desperately in need of some fundamental reforms in science education. *Scientists need to be trained, not only in the details of their specialty, but also in the principles of scientific inference in general*. As our incomplete information becomes more massive and quantitative, the same disciplined reasoning that Louis Pasteur was able to carry out in his head, now needs to be put into quantitative form.

I submit that this is precisely the Bayesian inference that Arnold Zellner and I have been teaching for several years; and it needs to be consolidated into undergraduate courses that are considered just as essential as is general literacy, for a scientist in any field. A little familiarity with and respect for the principles of rational inference would have muted the old claims of absolute truth and finality for quantum theory, and the sensational recent ones for such things as Chaos, Artificial Intelligence, Neural Nets, and the Big Bang theory.* I have the impression that physicists are in most need of this, because their training has completely neglected probability theory, and so they have become the most naïve of all scientists in the matter of judging what conclusions are and are not justified by some observed facts. Perhaps this is a legacy from the pre-scientific logic that we teach in quantum theory; the Bayesian-trained economists that I know would never commit such errors.

But it must not be supposed that our problems arise only in advanced science education; they commence already in kindergarten. From the educationists one hears constantly such phrases as: "skills for effective living", or "socially desirable responses", "adequate social behavior". Why is it important to be literate? Not because one will need to read with comprehension; but rather because "personal adjustment demands some speech and reading facility." Why is it important for a child to learn arithmetic? Not because he is going to have to know how to add and subtract numbers correctly for all the rest of his life. Of course not! It is important that he learn arithmetic because "a sense of failure here might affect his personality development". This is not material for standup comedians: every one of these quotations was found in course catalogs for Schools of Education. Such phrases as "adequate knowledge" or "rational behavior" do not occur at all. We need look no further to understand what needs to be corrected in American elementary education.

Perhaps it is clear that each of my little scoldings is at the same time an opportunity for the next generation to do something constructive about it. I am optimistic about the future because I know that my former students are capable of this, and suspect that their students are also. But the solutions will not come overnight, and they may need encouragement. Young scientists ought to study these problems most earnestly – just because of the fact that the solutions are not yet obvious (else my generation would have solved them).

Thank you again for all your kindness, and patience with my ramblings. Now I shall return home and do my best to complete those promised books, with all the details that these remarks have only hinted at, rather vaguely.

REFERENCES

Aspect, Alain (1986) "Experimental Tests of Bell's Inequalities with Pairs of Low Energy Correlated Photons", in *Frontiers of Nonequilibrium Statistical Mechanics*, G. T. Moore and M. O. Scully, Editors, Plenum Press, New York, pp. 163-183. See also the exchange between Aspect and Barut in the same volume.

Atkins, P. W. (1986), "Entropy in Relation to Complete Knowledge", Contemp. Phys. **27**, pp. 257-259.

* Recently it has been claimed that directional irregularities in the cosmic microwave background constitute decisive evidence in favor of the Big Bang theory. But the principles of scientific inference tell us that a theory is confirmed by observing things which it predicts *that are otherwise unexpected*; not by observing things that would be expected whatever one's theory of the beginning of the Universe.

Bell, J. W. (1987), *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press.

Bretthorst, G. L. (1988), *Bayesian Spectrum Analysis and Parameter Estimation*, Lecture Notes in Statistics, Vol. 48, Springer-Verlag, Berlin.

Crick, Francis (1988), *What Mad Pursuit*, Basic Books, Inc., New York.

Einstein, Albert (1946), *The Meaning of Relativity*, Princeton University Press, pp. 56-57.

Jaynes, E. T. (1973), "Survey of the Present Status of Neoclassical Radiation Theory", in *Proceedings of the 1972 Rochester Conference on Optical Coherence*, L. Mandel & E. Wolf, editors, Pergamon Press, New York.

Jaynes, E. T. (1989), "Clearing up Mysteries: The Original Goal", in *Maximum Entropy and Bayesian Methods*, J. Skilling, Editor, Kluwer Academic Publishers, Dordrecht, Holland, pp. 1-27.

Jaynes, E. T. (1990), "Probability in Quantum Theory", in *Complexity, Entropy and the Physics of Information*, W. H. Zurek, editor, Addison-Wesley Pub. Co., Redwood City CA, pp. 381-404.

Jaynes, E. T. (1991), "Scattering of Light by Free Electrons as a Test of Quantum Theory", in *The Electron*, D. Hestenes & A. Weingartshofer, Editors, Kluwer Academic Publishers, Dordrecht, Holland; pp 1-20.

Källén, Gunnar (1972), *Quantum Electrodynamics*, Springer-Verlag, New York.

Schwinger, Julian (1970), *Particles, Sources, and Fields*, Volume 1, Addison-Wesley Publishing Co., Reading MA. Volume 2 (1973).

Webster, David L. (1957), "Reminiscences of the Early Years of the Association", Am. Jour. Phys. **35**, pp. 131-134.