

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 known 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

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).

philosophical doctrines. They seemed to be lacking in depth and, perhaps to make a joke, a very 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 that 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 Pemyat.

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, P. Suppes and A. Tversky (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.