PATRICK SUPPES

FROM THEORY TO EXPERIMENT AND BACK AGAIN

In this article I consider two substantive examples of the way in which there is continuing interaction in science between theory and experiment. The picture of theory often presented by philosophers of science is too austere, abstract and self-contained. In particular, the picture of theory that is painted is much too removed from the shock effects of new experiments. Perhaps even more to the point, in many parts of science the actual formulation of theory is much driven by the latest experiments.

The first example comes from scientific research I am currently doing on language and the brain. I begin by describing the work in broad terms. I then present the response of new experiments and new theoretical statistical analysis of the data to answer claims that the recognition rates for brain-wave representation of words and sentences is not significant, because of the large amount of information available. Here the use of the concept of an extreme statistic is used to answer this criticism in a detailed way. Discussion of this example will end with some brief remarks on how this use of more detailed statistical methods is now generating new experiments, and having an impact on the design of the experiments.

The second example deals with experiments and physical theory on the entanglement of particles, and the consequent nonlocality of standard quantum mechanics. After some general remarks on this area of research in quantum mechanics and its philosophical importance for our basic physical concepts, I turn to the theoretical work of Greenberg, Horne and Zeilinger and their proposed "GHZ-type" experiments.

First the purely theoretical result, formulated in probability-one terms, is stated. Then the question is asked, how can such probability-one theoretical results be tested, given the inevitable inefficiences of particle detectors.

This prompts a new theoretical effort to derive inequalities, like those of Bell for other experiments, to deal with GHZ-type experiments. What comes out of the analysis is that better experimental results should be achievable with very careful design and use of current photon detectors. But the proof

M.C. Galavotti (ed.), Observation and Experiment in the Natural and Social Sciences, 1–41.

of this is rather detailed and relies on theory in critical ways at several points. These examples are but current illustrations, but the lesson is meant to be universal. The continual interaction between theory and experiment occurs in nearly every developed branch of science.

LANGUAGE AND THE BRAIN[1]

Some historical background

Aristotle said that the distinguishing feature of man as an animal is that he is a rational animal, but, in more biological and psychological terms, it is that of being a talking animal. Language is, in ways that we have not yet fully explored, the most distinguishing mark of man as an animal. Its processing is centered, above all, in the brain, not just for the production of speech, but for the intentional formation of what is to be said or for the comprehension of what has been heard or read. So it is the brain's processing of language that is the focus of this section. I begin with a historical sketch of the discovery of electrical activity in the brain.

An early reference to electricity being generated by muscles or nerves of animals comes from a study by Francesco Redi (1671), who describes in this way an experiment he conducted in 1666: "It appeared to me as if the painful action of the *torpedine* (electric ray) was located in these two sickle-shaped bodies, or muscles, more than in any other part." Redi's work was done in Florence under the Medici's. These electrical observations were fragmentary and undeveloped. But the idea of electrical activity in the muscles or nerves of various animals became current throughout the eighteenth century (Whittaker 1951, Galvani 1791). Yet it was more than 100 years after Redi before the decisive step was taken in Bologna by Luigi Galvani. He describes his first steps in the following manner:

> The course of the work has progressed in the following way. I dissected a frog and prepared it… Having in mind other things, I placed the frog on the same table as an electric machine. When one of my assistants by chance lightly applied the point of a scalpel to the inner crural nerves of the frog, suddenly all the muscles of the limbs were seen so to contract that they appeared to have fallen into violent tonic convulsions. Another assistant who was present when we were performing electrical experiments thought he observed that this phenomenon occurred when a spark was discharged from the conductor of the electrical machine. Marvelling at this, he immediately brought the unusual phenomenon to my attention when I was completely engrossed and contemplating other things. Hereupon I became extremely enthusiastic and eager to repeat the experiment so as to clarify the obscure phenomenon and make it known. I myself, therefore, applied the point of the scalpel first to one then to the other crural

---

[1] This section is taken from my forthcoming book *Representation and Invariance in Scientific Structures,* Stanford, CA: CSLI Publications.

> nerve, while at the same time some one of the assistants produced a spark; the phenomenon repeated itself in precisely the same manner as before.
>
> (Galvani 1791/1953, pp. 45–46)

Galvani's work of 1791 was vigorously criticized by the well-known Italian physicist Alessandro Volta (1745–1827), who was born in Como and was a professor of physics at the University of Pavia. Here are his words of criticism, excerpted from a letter by Volta to Tiberius Cavallo, read at the Royal Society of London:

> The name of animal electricity is by no means proper, in the sense intended by Galvani, and by others; namely, that the electric fluid becomes unbalanced in the animal organs, and by their own proper force, by some particular action of the vital powers. No, this is a mere artificial electricity, induced by an external cause, that is, excited originally in a manner hitherto unknown, by the connexion of metals with any kind of wet substance. And the animal organs, the nerves and the muscles, are merely passive, though easily thrown into action whenever, by being in the circuit of the electric current, produced in the manner already mentioned, they are attacked and stimulated by it, particularly the nerves.
>
> (Volta 1793/1918, pp. 203–208)

Galvani was able to meet these criticisms directly and in 1794 published anonymously a response containing the detailed account of an experiment on muscular contraction without the use of metals (Galvani 1794). The original and important nature of Galvani's work came to be recognized throughout Europe. The prominent German physicist Emil Du Bois-Reymond (1848) summarized in the following way Galvani's contribution:

> 1. Animals have an electricity peculiar to themselves, which is called Animal Electricity.
> 2. The organs to which this animal electricity has the greatest affinity, and in which it is distributed, are the nerves, and the most important organ of its secretion is the brain.
> 3. The inner substance of the nerve is specialized for conducting electricity, while the outer oily layer prevents its dispersal, and permits its accumulation.
> 4. The receivers of the animal electricity are the muscles, and they are like a Leyden jar, negative on the outside and positive on the inside.
> 5. The mechanism of motion consists in the discharge of the muscular fluid from the inside of the muscle via the nerve to the outside, and this discharge of the muscular Leyden jar furnishes an electrical stimulus to the irritable muscle fibres, which therefore contract.
>
> (Du Bois-Reymond 1848/1936, p. 159)

A next event of importance was the demonstration by Carlo Matteucci (1844) that electrical currents originate in muscle tissue. It was, however,

PAOLO LEGRENZI

NAÏVE PROBABILITY

The theory of mental models and extensional probability

Suppose that someone tells you: "If the director is in the office, then her secretary is in the office too". You start to think about the different possibilities compatible with the conditional. You think of the possibility of the director in the office, and so her secretary is in the office too. You think about what happens if the director is not in the office: in one possibility, the secretary is in the office; in another possibility, the secretary is not in the office, either. You have envisaged the three possibilities that are compatible with the truth of the conditional assertion, which we summarize as follows, using "¬" to denote negation:

Director in office	Secretary in office
¬ Director in office	Secretary in office
¬ Director in office	¬ Secretary in office

Following philosophers and logicians, we refer to such possibilities as the "extensions" of the conditional assertion, i.e., possibilities to which it refers. And when individuals infer probabilities by considering the extensions of assertions, we shall say that they are reasoning *extensionally.*

You can tackle the same problem in a different way. You know that directors are unlikely to spend as much time in the office as their secretaries. This stereotype may have occurred to you as you were thinking about the problem, and you might have based your inference on it. When you think in this way, you do not consider the extensions of assertions, but rather you use some index – some evidence or knowledge – to infer a probability. We use "non-extensional" as an umbrella term to cover the many ways in which people can arrive at probabilities without thinking about extensions. Of course, you might think about a problem both extensionally and non-extensionally.

M.C. Galavotti (ed.), Observation and Experiment in the Natural and Social Sciences, 43-55.

Given a problem about a set of events, you can consider its partition, that is, the exhaustive set of possible conjunctions of individual events. In the problem about the director and the secretary, there are four such possibilities, which comprise this "partition" for the problem:

Director in office	Secretary in office
Director in office	¬ Secretary in office
¬ Director in office	Secretary in office
¬ Director in office	¬ Secretary in office

Once you know the probabilities for each possibility in a partition, you know everything that is to be known from a probabilistic standpoint. So let us introduce some probabilities, which for convenience we state as chances out of a hundred:

		Chances
Director in office	Secretary in office	50
Director in office	¬ Secretary in office	0
¬ Director in office	Secretary in office	30
¬ Director in office	¬ Secretary in office	20

You can now deduce the probability of any assertion about the domain, including conditional probabilities, such as:

> The probability that the director is not in the office given that the secretary is in the office: 30/80

The mental model theory postulates that each mental model represents a possibility, and that its structure and content capture what is common to the different ways in which the possibility might occur. For example, when individuals understand that either the director or else the secretary is in the office, but not both, they construct two mental models to represent the two possibilities:

director	
	secretary

where each line represents an alternative model, "director" denotes a model of the director in the office, and "secretary" denotes a model of the secretary in the office. Likewise, a conjunction, such as:

The director is in the office and the secretary is in the office

has only a single mental model:

director secretary

Granted that individuals construct mental models to represent the possibilities described in assertions, they can reason by formulating a conclusion that holds in their mental models, and they can test its *validity* by checking whether it holds in all possible models of the discourse. They can establish the invalidity of a conclusion by finding a counterexample, i.e., a model of the discourse in which the conclusion is false.

The theory makes a fundamental assumption, which is known as the principle of *truth:*

> Individuals represent assertions by constructing sets of mental models in which, first, each model represents a true possibility, and, second, the clauses in the assertions, affirmative or negative, are represented in a mental model only if they are true in the possibility.

Consider an exclusive disjunction in which only one of the two clauses is true:

The director is not in the office or else the secretary is in the office.

The mental models of the disjunction represent only the two true possibilities, and within them, they represent only the two clauses in the disjunction when they are true within a possibility:

¬ director
secretary

The first model represents the possibility that the director is not in the office, but it does not represent explicitly that it is false that the secretary is in the office. The second model represents the possibility that the secretary is in the office, but it does not represent explicitly that it is false that the director is not in the office (i.e. the director *is* in the office).

The mental models of conditionals are simple. For a conditional, such as:

LÀSZLÓ E. SZABÓ

FROM THEORY TO EXPERIMENTS AND BACK AGAIN ... AND BACK AGAIN ...

COMMENTS ON PATRICK SUPPES

"The picture of theory often presented by philosophers of science is too austere, abstract and self-contained" Professor Suppes writes. While, as it turns out from the two substantive examples considered in the paper, a closer analysis of the experimental details, the method of data processing and the most important features of the measuring equipments can be fruitful in understanding the basic concepts and the metaphysical conclusions drawn from the theoretical description of the experimental scenario.

Since my field of interest is closer to quantum mechanics, I would like to focus on Suppes' second example based on de Barros and Suppes (2000) general analysis of the realistic GHZ experiments, where experimental error reduces the perfect correlations of the ideal GHZ case. The following important question motivated their analysis: "How *can one verify experimentally predictions based on correlation-one statements, since experimentally one cannot obtain events perfectly correlated?*" De Barros and Suppes' analysis makes use of inequalities which are said to be "*both necessary and sufficient for the existence of a local hidden variable*" for the experimentally realizable GHZ correlations. In applying their analysis to the Innsbruck experiment, however, they only count events in which all the detectors fire. While necessary for the analysis of that experiment, they recognize that this selective procedure weakens the argument for the nonexistence of local hidden variables.

In Szabó and Fine (2002) we pointed out that their analysis does not rule out a whole class of local hidden variable models in which the detection inefficiency is not (only) the effect of the random errors in the detector equipment, but it is a more fundamental phenomenon, the manifestation of a predetermined hidden property of the particles. This conception of local hidden variables was first suggested in Fine's *prism model* (1982) and, arguably, goes back to Einstein.

M.C. Galavotti (ed.), Observation and Experiment in the Natural and Social Sciences, 57–62.

Both, de Barros and Suppes' analysis and our polemics, confirm, however, Suppes' thesis about the continuing interaction in science between theory and experiment.

Theory => Experiment

De Barros and Suppes approach the problem in the following way. Without loss of generality, the space of hidden variable can be identified with $O = \{+, -\}^6$, the set of the $2^6 = 64$ different 6-tuples of possible combinations of the values of $\sigma_{1x}, \sigma_{1y}, \ldots \sigma_{3y}$. Then the GHZ contradiction amounts to the assertion that no probability measure over O reproduces the expectation values.

De Barros and Suppes demonstrate this by concentrating on the product observables *(A, B, C* and *ABC)* for which they derive a system of inequalities that play the same role for GHZ that the general form of the Bell inequalities do for EPR-Bohm type experiments; namely, they provide necessary and sufficient conditions for a certain class of local hidden vaxiable models. Their inequalities axe just

$$\begin{aligned}
-2 &\leq E(A) + E(B) + E(C) - E(ABC) \leq 2 \\
-2 &\leq E(A) + E(B) - E(C) + E(ABC) \leq 2 \\
-2 &\leq E(A) - E(B) + E(C) + E(ABC) \leq 2 \\
-2 &\leq E(A) + E(B) + E(C) + E(ABC) \leq 2
\end{aligned}$$

and clearly this is violated by

$$\begin{aligned}
E(A) = E(B) = E(C) &= 1 \\
E(ABC) &= -1
\end{aligned}$$

Experiment => Theory

In the realistic experiments, due to inefficiencies in the detectors or to dark photon detection, the observed correlations were reduced by some factor e; that is:

$$E(A) = E(B) = E(C) = 1 - \varepsilon$$

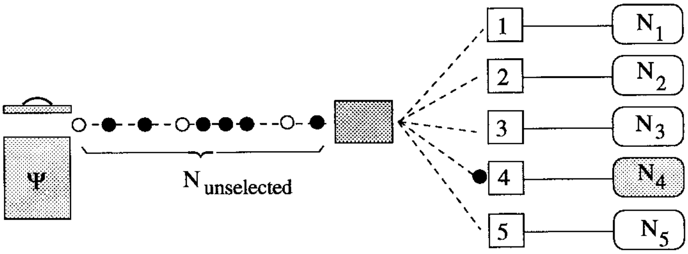
$$E(ABC) = -1 + \varepsilon$$

Theory => Experiment

Then, it follows immediately from the inequalities that, "*the observed correlations are only compatible with a local hidden variable theory*" if $\varepsilon > 1/2$. De Barros and Suppes (2000) translated this condition into the language of the dark-count rate and the detector efficiency.

Experiment => Theory

Estimating the realistic values of the dark-count rate and the detector efficiency, they found that the Innsbruck experiment is not compatible with a local hidden variable theory.



$$N_{\text{unselected}} \neq \sum_i N_i$$

$$tr\,(WP_i) = \frac{N_i}{\sum_i N_i}$$

Figure 1: *In a typical quantum measurement, quantum mechanical "probabilities" are equal to the relative frequencies taken on a* sub-ensemble of *objects producing any outcome.*