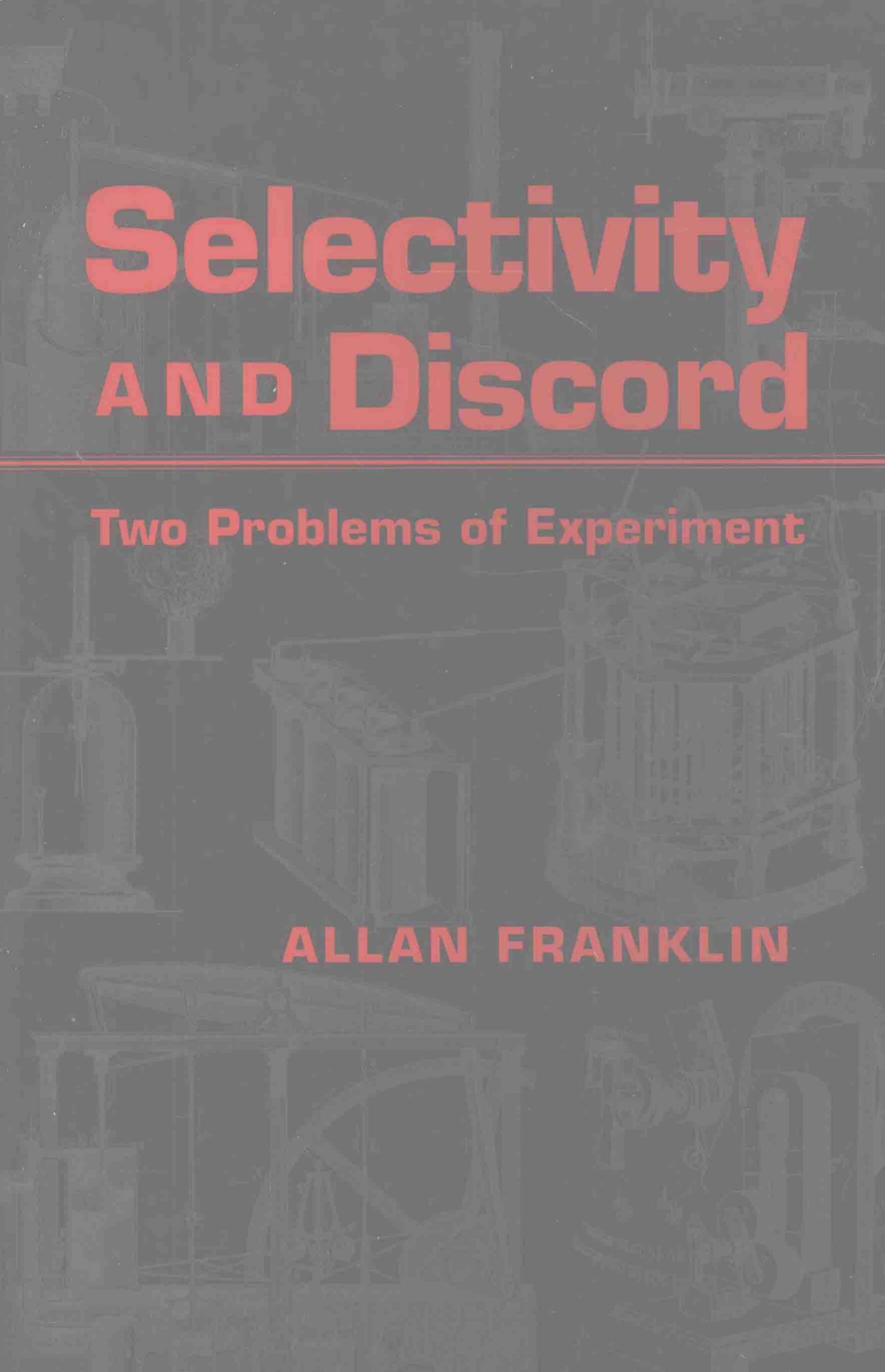

Selectivity and Discord

Two Problems of Experiment

ALLAN FRANKLIN

Selectivity AND Discord

Two Problems of Experiment

ALLAN FRANKLIN

University of Pittsburgh Press

Published by the University of Pittsburgh Press, Pittsburgh, Pa. 15261
Copyright © 2002, University of Pittsburgh Press
All rights reserved
Manufactured in the United States of America
Printed on acid-free paper
10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data
Franklin, Allan, 1938–
Selectivity and discord : two problems of experiment / Allan Franklin
p. cm.
Includes bibliographical references and index.
ISBN 0-8229-4191-0 (cloth : alk. paper)
1. Science—Philosophy. 2. Physics—Experiments. 3. Physics—History.
I. Title
Q175.F795 2002
501—dc21

2002011054

ACKNOWLEDGMENTS

Some of the work for this book was carried out while I was a resident fellow at the Dibner Institute for the History of Science and Technology. I am grateful to Jed Buchwald, the director; Evelyn Simha, the executive director; and their staff, Carla Chrisfield, Rita Dempsey, and Trudy Kontoff, for providing both support and an atmosphere in which it is almost impossible not to get work done. I have benefited enormously from discussions with Rudolf Ganz, one of the leading experimenters on possible low-mass electron-positron states, who also provided a copy of his dissertation. I have also had valuable discussions with George Smith; Roger Stuewer; Noel Swerdlow; my colleagues in the Department of Physics at the University of Colorado, Tony Barker, John Cumalat, and Patricia Rankin; and my colleagues in the Department of History and Philosophy of Science at the University of Melbourne, Maureen Christie, Rod Home, Howard Sankey, and Neil Thomason; and Manuel Thomaz, who joined me as a visitor in the department. Much of the material in this book originally appeared in my papers "The Resolution of Discordant Results," *Perspectives on Science* (1995) 3: 346–420, and "Selectivity and the Production of Experimental Results," *Archive for the History of Exact Sciences* (1998) 53: 399–485. The material in Chapter 8 appeared in "William Wilson and the Absorption of Beta Rays," *Physics in Perspective* (2002) 4: 40–70.

CONTENTS

此为试读，需要完整PDF请访问：www.ertongbook.com

Introduction

Physics and, I believe, all of science is a reasonable enterprise based on experimental evidence, criticism, and rational discussion. It provides us with knowledge of the physical world, and it is experiment that provides the evidence that grounds this knowledge. As the late Richard Feynman, one of the leading theoretical physicists of the 20th century, wrote, "The principle of science, the definition, almost, is the following: *The test of all knowledge is experiment.* Experiment is the *sole judge* of scientific 'truth'" (Feynman et al., 1963, p. I-1). In these postmodern times this might seem to be an old-fashioned view, but it is one I consider correct.

Experiment plays many roles in science. One of its important roles is to test theories and provide the basis for scientific knowledge. It can also call for a new theory, either by showing that an accepted theory is incorrect or by exhibiting a new phenomenon that is in need of explanation. Experiment can provide hints about the structure or mathematical form of a theory, and it can provide evidence for the existence of the entities involved in our theories. Finally, it may also have a life of its own, independent of theory: Scientists may investigate a phenomenon just because it looks interesting. Such experiments may provide evidence for future theories to explain.

In all of this activity, however, we must remember that science is fallible. Theoretical calculations, experimental results, or the comparison between experiment and theory may all be wrong. Science is more complex than "The scientist proposes, Nature disposes." It may not always be clear what the scientist is proposing. Theories must often be articulated and clarified. It also may not be clear just how nature is disposing. Experiments may not always give clear-cut results, and they may even disagree for a time. Sometimes they can be incorrect.

If experiment is to play these important roles in science, then we must have good reasons to believe experimental results. I present here an epistemology of experiment, a set of strategies that provides reasonable belief in experimental results. Scientific knowledge can then be reasonably based on these experimental results.

Not everyone agrees. Harry Collins, for that example, remarks that "the natural world has a small or non-existent role in the construction of scientific knowledge" (Collins, 1981, p. 3).[1] And Barry Barnes has stated that "Reality will tolerate alternative descriptions without protest. We may say what we will of it, and it will not disagree. Sociologists of knowledge rightly reject epistemologies that *empower* reality" (Barnes, 1991, p. 331).[2] This view led Andy Pickering to remark that "there is no obligation upon anyone framing a view of the world to take account of what twentieth-century science has to say."[3] In this book I argue for the view that nature, as revealed by experiment, plays an important and legitimate role in science. I will begin by offering my own version of an epistemology of experiment, a set of strategies used by scientists to argue for the correctness of an experimental result. I have argued elsewhere that such strategies are justified. I also discuss the views of other scholars, some that support my own view and others that do not.

Experimental Results

The Case for Learning from Experiment

AN EPISTEMOLOGY OF EXPERIMENT

It has been two decades since Ian Hacking asked "Do we see through a microscope?" (Hacking, 1981). Hacking's question really asked How do we come to believe in an experimental result obtained with a complex experimental apparatus? How do we distinguish between a valid result[4] and an artifact created by that apparatus? If experiment is to play all of the important roles in science mentioned above and to provide the evidential basis for scientific knowledge, then we must have good reasons to believe in those results. Hacking (1983) provided an extended answer in the second half of *Representing and Intervening.* He pointed out that even though an experimental apparatus is laden with (at the very least) the theory of the apparatus, observations remain robust despite changes in the theory of the apparatus or the theory of the phenomenon. His illustration was the sustained belief in microscope images despite the major change in the theory of the microscope when Abbe pointed out the importance of diffraction in its operation. One reason Hacking gave for this continued belief is that in making such observations, the experimenters intervened—they manipulated the object under observation. Thus, in looking at a cell through a microscope, one might inject fluid into the cell or stain the specimen. One expects the cell to change shape or color when this is done. Observing the predicted effect strengthens our belief in the proper operation of the experimental apparatus and in the validity of the observation itself.

Hacking also discussed the strengthening of belief in an observation by independent confirmation. The fact that the same pattern of dots—dense bodies in cells—is seen with "different" microscopes (e.g., ordinary, polarizing, phase-contrast, fluorescence, interference, electron, acoustic) argues for the validity of the observation. One might question whether "different" is a theory-laden term. After all, it is our theories of light and the microscope that allow us to consider these microscopes as different from each other. Nevertheless, the argument holds: Hacking correctly argues that it would be a preposterous coincidence if the same pattern of dots were produced in two totally different kinds of physical systems. Different apparatuses have different backgrounds and systematic errors, making the coincidence, if it is an artifact, most unlikely. If it is a correct result, and the instruments are working properly, the agreement of results is understandable.[5]

Hacking's answer is correct as far as it goes. It is, however, incomplete. What happens when one can perform the experiment with only one type of apparatus, such as an electron microscope or a radio telescope, or when intervention is either impossible or extremely difficult? Other strategies are needed to validate the observation.[6] These may include:

1. Experimental checks and calibration, in which the experimental apparatus reproduces known phenomena. For example, if we wish to argue that the spectrum of a substance obtained with a new type of

spectrometer is correct, we might check that this new spectrometer could reproduce the known Balmer series for hydrogen. If we correctly observe the Balmer series, then we strengthen our belief that the spectrometer is working properly. This also strengthens our belief in the results obtained with that spectrometer. If the check fails, then we have good reason to question the results obtained with that apparatus.[7]

2. Reproducing artifacts that are known in advance to be present. An example of this comes from experiments to measure the infrared spectra of organic molecules (Randall et al., 1949). It was not always possible to prepare a pure sample of such material. Sometimes the experimenters had to place the substance in an oil paste or in solution. In such cases, one expects to observe the spectrum of the oil or the solvent superimposed on that of the substance; one can then compare the composite spectrum with the known spectrum of the oil or the solvent. Observation of this artifact gives confidence in other measurements made with the spectrometer.
3. Elimination of plausible sources of error and alternative explanations of the result (the Sherlock Holmes strategy).[8] Thus, when scientists claimed to have observed electric discharges in the rings of Saturn, they argued for their result by showing that it could not have been caused by defects in the telemetry, interaction with the environment of Saturn, lightning, or dust. The only remaining explanation of their result was that it was due to electric discharges in the rings—there was no other plausible explanation of the observation. (In addition, the same result was observed by both spacecrafts Voyager 1 and Voyager 2. This provided independent confirmation. Often, several epistemological strategies are used in the same experiment.)
4. Using the results themselves to argue for their validity. Consider the problem of Galileo's telescopic observations of the moons of Jupiter. Although one might very well believe that his primitive, early telescope might have produced spurious spots of light, it is extremely implausible that the telescope would create images that would appear to be eclipses and other phenomena consistent with the motions of a small planetary system. It is even more implausible that the created spots would satisfy Kepler's Third Law (R^3/T^2 = constant).[9] A similar argument was used by Robert Millikan to support his observation of

the quantization of electric charge and his measurement of the charge of the electron. Millikan (1911) remarked, "The total number of changes which we have observed would be between one and two thousand, and *in not one single instance has there been any change which did not represent the advent upon the drop of one definite invariable quantity of electricity or a very small multiple of that quantity*" (p. 360). In both of these cases one is arguing that there was no plausible malfunction of the apparatus (or no confounding background) that would explain the observations.

5. Using an independently well-corroborated theory of the phenomena to explain the results. This was illustrated in the discovery of the $W^{\pm}$, the charged intermediate vector boson required by the Weinberg-Salam unified theory of electroweak interactions. Although these experiments used very complex apparatuses and used other epistemological strategies (for details, see Franklin [1986], pp. 170–72). I believe that the agreement of the observations with the theoretical predictions of the particle properties helped to validate the experimental results. In this case, the particle candidates were observed in events that contained an electron with high transverse momentum and in which there were no particle jets, just as predicted by the theory. In addition, the measured particle mass of 81 ± 5 GeV/c^2 and 80^{+10}_{-6}, GeV/c^2, found in the two experiments (note the independent confirmation), was in good agreement with the theoretical prediction of 82 ± 2.4 GeV/c^2. It was very improbable that any background effect, which might mimic the presence of the particle, would be in good agreement with theory.
6. Using an apparatus based on a well-corroborated theory. In this case, the support for the theory inspires confidence in the apparatus based on that theory. This is the case with the electron microscope and the radio telescope, whose operations are based on well-supported theories, although other strategies are also used to validate observations made with these instruments.
7. Using statistical arguments. An interesting example of this arose in the 1960s, when the search for new particles and resonances occupied a substantial fraction of the time and effort of those physicists working in experimental high-energy physics. The usual technique was to plot the number of events observed as a function of the invariant

mass of the final-state particles and to look for bumps above a smooth background. The usual informal criterion for the presence of a new particle was that it resulted in a three-standard-deviation effect above the background, a result that had a probability of 0.27% of occurring in a single bin. This criterion was later changed to four standard deviations, which had a probability of 0.0064% when it was pointed out that the number of graphs plotted each year by high-energy physicists made it rather probable, on statistical grounds, that a three-standard-deviation effect would be observed.[10]

These strategies, along with Hacking's intervention and independent confirmation, constitute an epistemology of experiment: They provide us with good reasons for belief in experimental results. They do not, however, guarantee that the results are correct. There are many experiments in which these strategies are applied, but whose results are later shown to be incorrect (examples are presented throughout this book). Experiment is fallible. Neither are these strategies exclusive or exhaustive. No single one of them, or fixed combination of them, guarantees the validity of an experimental result. As the episodes discussed in this book show, physicists use as many of the strategies as they can conveniently apply in any given experiment.

GALISON'S ELABORATION

In *How Experiments End*, Peter Galison (1987) extended the discussion of experiment to more complex situations. In his histories of the measurements of the gyromagnetic ratio of the electron, the discovery of the muon, and the discovery of weak neutral currents, he considers a series of experiments measuring a single quantity, a set of different experiments culminating in a discovery, and two high-energy physics experiments performed by large groups with complex experimental apparatus.

Galison's view is that experiments end when the experimenters believe that they have a result that will stand up in court—a result that I believe includes the use of the epistemological strategies discussed earlier. Thus, David Cline, one of the weak neutral-current experimenters remarked, "At present I don't see how to make these effects [the weak neutral-current event candidates] go away" (Galison, 1987, p. 235).

Galison emphasizes that, within a large experimental group, different members of the group may find different pieces of evidence most con-

此为试读，需要完整PDF请访问：www.ertongbook.com

vincing. Thus, in the Gargamelle weak neutral-current experiment, several group members found the single photograph of a neutrino-electron scattering event particularly important, whereas for others, the difference in spatial distribution between the observed neutral-current candidates and the neutron background was decisive. Galison attributes this, in large part, to differences in experimental traditions, in which scientists develop skill in using certain types of instruments or apparatuses. In particle physics, for example, there is the tradition of visual detectors, such as the cloud chamber or the bubble chamber, in contrast to the electronic tradition of Geiger and scintillation counters and spark chambers. According to Galison, scientists within the visual tradition tend to prefer "golden events" that clearly demonstrate the phenomenon in question, whereas those in the electronic tradition tend to find statistical arguments more persuasive and important than individual events. (For further discussion of this issue, see Galison [1997] and the next section.)

Galison points out that major changes in theory and in experimental practice and instruments do not necessarily occur at the same time. This persistence of experimental results provides continuity across conceptual changes. Thus, the experiments on the gyromagnetic ratio spanned classical electromagnetism, Bohr's old quantum theory, and the new quantum mechanics of Heisenberg and Schrodinger. Robert Ackermann (1985) has offered a similar view in his discussion of scientific instruments:

> The advantages of a scientific instrument are that it cannot change theories. Instruments embody theories, to be sure, or we wouldn't have any grasp of the significance of their operation. . . . Instruments create an invariant relationship between their operations and the world, at least when we abstract from the expertise involved in their correct use. When our theories change, we may conceive of the significance of the instrument and the world with which it is interacting differently, and the datum of an instrument may change in significance, but the datum can nonetheless stay the same, and will typically be expected to do so. An instrument reads 2 when exposed to some phenomenon. After a change in theory,[11] it will continue to show the same reading, even though we may take the reading to be no longer important, or to tell us something other than what we thought originally. (p. 33)

Galison also discusses other aspects of the interaction between experiment and theory. Theory may influence what is considered to be a real effect, demanding explanation, and what is considered background. In his discussion of the discovery of the muon, he argues that the calculation of Oppenheimer and Carlson, which showed that showers were to be expected in the passage of electrons through matter, left the penetrating particles, later shown to be muons, as the unexplained phenomenon. Prior to their work, physicists thought the showering particles were the problem, whereas the penetrating particles seemed to be understood.

The role of theory as an "enabling theory" (i.e., one that allows calculation or estimation of the size of the expected effect and the size of expected backgrounds) is also discussed by Galison (see also Franklin [1995b]). Such a theory can help to determine whether an experiment is feasible. Galison emphasizes that elimination of background that might simulate or mask an effect is central to the experimental enterprise, and not just a peripheral activity. In the case of the weak neutral-current experiments, the existence of the currents depended crucially on showing that the event candidates could not all be due to neutron background.[12]

There is also a danger that the design of an experiment may preclude observation of a phenomenon. Galison points out that the original design of one of the neutral current experiments, which included a muon trigger, would not have allowed the observation of neutral currents. In its original form, the experiment was designed to observe charged currents, which produce a high-energy muon. Neutral currents do not. Therefore, having a muon trigger precluded their observation. Only after the theoretical importance of the search for neutral currents was emphasized to the experimenters was the trigger changed. Changing the design did not, of course, guarantee that neutral currents would be observed.

Galison shows that the theoretical presuppositions of the experimenters may enter into the decision to end an experiment and report the result. Einstein and de Haas ended their search for systematic errors when their value for the gyromagnetic ratio of the electron, $g = 1$, agreed with their theoretical model of orbiting electrons. This effect of presuppositions might cause one to be skeptical of both experimental results and their role in theory evaluation. Galison's history shows, however, that, in this case, the importance of the measurement led to many repetitions of the measurement. This resulted in an agreed-upon result that diverged from theoretical expectations: Scientists do not always find what they are looking for.

STALEY VERSUS GALISON

Recently, Galison has modified his views. In *Image and Logic*, an extended study of instrumentation in 20th-century high-energy physics, Galison (1997) has extended his argument that there are two distinct experimental traditions within that field—the visual (or image) tradition and the electronic (or logic) tradition. The image tradition uses detectors such as cloud chambers or bubble chambers, which provide detailed and extensive information about each individual event. The electronic detectors used by the logic tradition, such as Geiger counters, scintillation counters, and spark chambers, provide less detailed information about individual events, but detect more events. Galison's view is that experimenters working in these two traditions form distinct epistemic and linguistic groups that rely on different forms of argument.[13] The visual tradition emphasizes the single "golden" event. "On the image side resides a deep-seated commitment to the production of the 'golden event': the single picture of such clarity and distinctness that it commands acceptance" (Galison, 1997, p. 22). "The golden event was the exemplar of the image tradition: an individual instance so complete, so well defined, so 'manifestly' free of distortion and background that no further data had to be invoked" (p. 23). Because the individual events provided in the logic detectors contained less detailed information than the pictures of the visual tradition, statistical arguments based on large numbers of events were required.[14]

Kent Staley (1999) disagrees. He argues that the two traditions are not as distinct as Galison believes:

> I show that discoveries in both traditions have employed the same statistical [I would add "and/or probabilistic"] form of argument, even when basing discovery claims on single, golden events. Where Galison sees an epistemic divide between two communities that can only be bridged by a creole- or pidgin-like 'interlanguage,' there is in fact a shared commitment to a statistical form of experimental argument. (p. 196).

Staley believes that although there is certainly epistemic continuity within a given tradition, there is also a continuity between the traditions. This does not, I believe, mean that the shared commitment comprises all of the arguments offered in any particular instance, but rather that the same methods are often used by both communities. Galison does not deny that statistical methods are used in the image tradition, but he thinks that they are relatively unimportant. "While statistics could certainly be used within the image tradition, it was by no means necessary for most applications" (Galison 1997, p. 451). In contrast, Galison believes that experiments in the logic tradition "were inherently and inalienably statistical. Estimation of probable errors and the statistical excess over background is not a side issue in these detectors—it is central to the possibility of any demonstration at all" (p. 451). As we shall see, Galison himself presents an example from the visual tradition that exemplifies the use of statistical strategies.

It is interesting to examine the disagreement between Staley and Galison because it illuminates and illustrates issues in the epistemology of experiment. This examination will also show the complexity of demonstrating the validity of an experimental result and the care shown in that demonstration.[15] I will begin with a discussion of what they both regard as a golden event:[16] Anderson's photograph that provided evidence for the existence of the positron (Figure I.1).

> The image in question is a cloud chamber photograph that shows two tracks, one on either side of a 6 mm lead plate inserted into the chamber. The two tracks match up very closely, suggesting a single particle passing through the lead. Differences in the curvatures of the tracks above and below indicate a higher energy below the lead than above, which entails, on the assumption that it is indeed a single particle and that particles do not *gain* energy when passing through lead, that the particle was traveling from the lower to the upper region of the space in the photograph. Knowing the direction and curvature of the path, as well as the magnetic field, Anderson concludes that the particle has a positive charge. But based on the length of the track and the energy indicated by the curvature, it cannot have been a proton, which would have had a much shorter range. The particle, then, must have much lighter mass, on the same order of magnitude as that of a free negative electron. (Staley, 1999, p. 215)[17]

Staley argues that Anderson was, in fact, making a statistical argument premised on the claim that the probability of a background event that might have mimicked the presence of a positron was small even when compared to the single event under consideration. Anderson explicitly makes such an argument. In considering alternative explanations of the photograph demonstrating the existence of the positron, he stated:

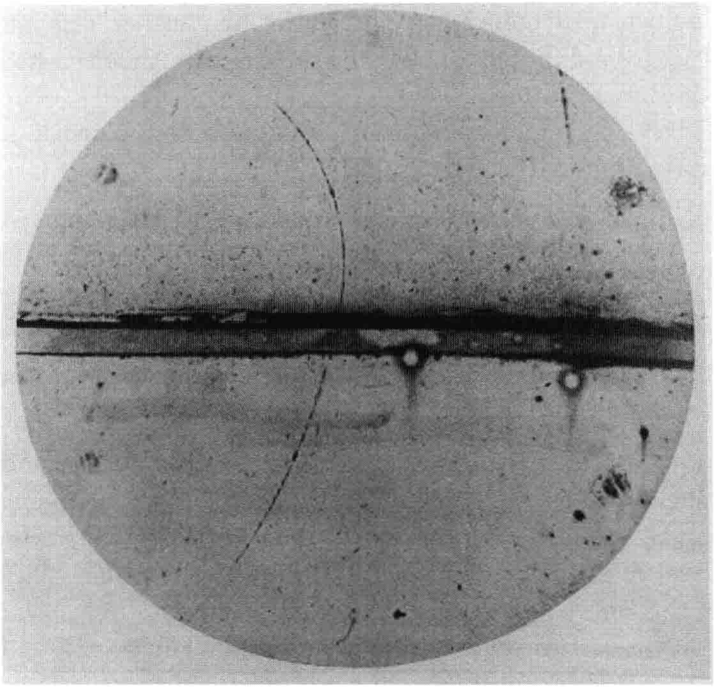

Figure I.1. Anderson's "golden event." The original caption for this figure reads, "A 63 million volt positron ($H\rho = 2.1 \times 10^5$ gauss-cm) passing through a 6 mm lead plate emerging as a 23 million volt positron. [The positron is traveling toward the top of the figure.] The length of this latter path is at least ten times greater than the possible length of a proton path of this curvature." From Anderson (1933, p. 492).

The only escape from this conclusion would be to assume that at exactly the same instant (and the sharpness of the tracks determines that instant to within about a fiftieth of a second) two independent electrons happened to produce two tracks so placed as to give the impression of a single particle shooting through the lead plate. This assumption was dismissed on a *probability basis*, since a sharp track of this order of curvature under the experimental conditions prevailing occurred in the chamber only once in some 500 exposures, and since there was practically no chance at all that two such tracks should line up in this way. (Anderson 1933, p. 491, emphasis added)

As Staley notes, if the probability of a single track is one in 500 exposures, the probability of two such tracks in the same photograph is one in 250,000 exposures, and the probability that they would line up so as to appear to be a single track reduces this probability even further. This was a negligible background indeed, considering the fact that Anderson had only 1,300 exposures.

Galison (1999), in response, notes that Anderson considered four alternative explanations of the photograph:

1. Light positive particle penetrated the lead (ionization ruled out a proton).
2. Simultaneous ejection of positron and electron.
3. Electron *gained* energy in passing downwards through the lead.
4. Two independent electron tracks were perfectly aligned to imitate a positron losing energy. (p. 272)

Galison notes that the first two posit the existence of the positron and thus are not alternative explanations and the third is ruled out by energy conservation.[18] He then asks "Why promote 4) to being *the* unifying epistemological basis of the discovery?" (p. 272). In my view, Galison's "*the*" is an exaggeration. Staley does mention the other alternatives and is here showing that in this particular golden event, the experimenter could, and did, argue on statistical or probabilistic grounds that the background was negligible, and thus that the observation was a real effect. Staley's analysis shows that statistical arguments were one of the arguments used by those in the visual tradition. He is not claiming that it is always the sole argument, or that it is always used. Galison correctly points out that the golden event can be and has been decisive in many instances. He cites Powell et al. (1959): "It is a remarkable feature of those methods in nuclear physics based on recording individual tracks, that the observation of a single event has frequently been of decisive importance in leading to the discovery of phenomena of fundamental importance." Note, however, that Powell and company say "frequently," not "always." What we have seen here is an example of the Sherlock Holmes strategy, in which the elimination of alternative explanations of an experimental result involved the use of statistical arguments.

Staley presents other arguments supporting his view that statistical arguments are not only used within the image tradition, but are often of crucial importance. He presents a discussion of the episode of the discovery of the η meson. In this episode, bubble chamber photographs of the interaction of π^+ mesons with deuterium were examined. Events fitting the hypothesis $\pi^+ + d \rightarrow p + p + \pi^+ + \pi^- + \pi^0$ were analyzed and

此为试读,需要完整PDF请访问: www.ertongbook.com