KARIN KNORR CETINA

Epistemic Cultures

How the Sciences Make Knowledge

Ch. 3

2003

Harvard University Press
Cambridge, Massachusetts
London, England

3

Particle Physics and Negative Knowledge

3.1 The Analogy of the Closed Universe

I shall now seek to characterize epistemic cultures more systematically, switching from laboratories to laboratory processes and to the ways different sciences understand and enact empirical research. I want to begin with high energy physics; the chapter examines the observable order of HEP's research policies as these are displayed in physicists' experimental activities, in their meetings (where these activities are exhibited among fellow physicists), in their explanations to me (which I draw from for "easy" illustrations not requiring whole transcripts or research histories), and in their conversations.

When defining laboratories I said that the empirical machinery of HEP is a sign-processing machinery. It moves in the shadowland of mechanically, electrically, and electronically produced negative images of the world—in a world of signs and often fictional reflections, of echoes, footprints, and the shimmering appearances of bygone events. In what follows I show how this world that is marked by a loss of the empirical operates in terms of a negative and reflexive epistemics (the notion of epistemics is used here to refer to the strategies and practices assumed to promote the "truth"-like character of results). An analogy appropriately describes the "truth-finding" strategies of experimental HEP, which I want to present up front. This is the analogy of the brain as an informationally closed system. The neurophysiology of cognition is based on theories developed in the nineteenth century proposing that states of arousal in a nerve cell in the brain represent only the *intensity* but not the

nature of the source of arousal. Maturana and Varela (e.g., 1980) applied these results to the experimental study of perception. They concluded that perception must be seen as a cognitive process that is energetically open but informationally closed. Perception is accomplished by the brain, not the eye, and the brain can only construe what it sees from signals of light intensity that arrive at the retina. In order to form a picture of the nature of the source of these signals, the brain refers to its own previous knowledge and uses its own electrochemical reactions. Phrased differently, during perception the brain interacts only with itself, not with an external environment. It reconstructs the external world from internal states, and in order to accomplish this the brain "observes" itself. Consciousness, according to this theory, is a function of a nervous system capable only of recursive self-observation.

I want to argue that, like the brain, high energy physics operates within a *closed* circuitry. In many ways, it operates in a world of objects separated from the environment, a world turned inward, or, better still, a world entirely reconstructed within the boundaries of a complicated multilevel technology of representation. A detector is a kind of ultimate seeing device, a type of microscope that provides the first level of these representations. The representations themselves show all the ambiguities that afflict any world composed of signs. Yet particle physics is perfectly capable of deriving truth effects from sign-processing operations.

The spotlight in this chapter is on the rough parameters of these procedures to gain access to the world and stability in outcomes. Section 3.2 provides more details of the semiological understanding of inaccessible objects and their variously distorted traces. In 3.3 I discuss the notion of the "meaninglessness" of pure measurement and the practice of molding real "data" from the intersection of theory, detector models, and measured components. Section 3.4 describes the replacement, in these experiments, of the "care of objects" with the "care of the self"—their switch from an emphasis on "observing the world" to an emphasis on observing (controlling, improving, recording, understanding . . .) their own components and processes. Sections 3.5 and *3.6* complete the view of experimental HEP as an internally referential system by describing the turn toward negative and liminal knowledge for epistemic purposes—and by exemplifying relevant strategies such as unfolding, framing, and convoluting. Readers interested in selective differences in laboratory processes between molecular biology and HEP might focus on Sections 3.3 to 3.5.

3.2 A World of Signs and Secondary Appearances

In HEP experiments, natural objects (cosmic particles) and quasi-natural objects (debris of particles smashed in particle collisions) are admitted to experiments only rarely, perhaps for intermittent periods of several months in an experiment that used to last ten and more years and now lasts more than twenty. The proposals for UAl and UA2, the two large collider experiments at CERN, were approved in 1978, after several years of preparatory work, and both experiments were dismantled in 1991, although analysis of UAl and UA2 data continued. During the upgrading period in which the detectors were rebuilt, which lasted from the early 1980s to the end, the experiments had four "runs" (data-taking periods) between 1987 and 1990 lasting about four months each. Thus, researchers deal with the objects of interest to them only very occasionally, while most experimental time is spent on design, installation, testing, and other work outlined below.

What is more, these objects are in a very precise sense "unreal"—or, as one physicist described them, "phantasmic" *(irreale Gegenstände);* they are too small ever to be seen except indirectly through detectors, too fast to be captured and contained in a laboratory space, and too dangerous as particle beams to be handled directly. Furthermore, the interesting particles usually come in combination with other components that mask their presence. Finally, most subatomic particles are very short-lived, transient creatures that exist only for a billionth of a second. Subject to frequent metamorphosis and to decay, they "exist" in a way that is always already past, already history.

These phantasmic, historical, constantly changing occurrences can be established only indirectly, by the footprints they leave when they fly through different pieces of equipment. When a detector "sees" or "senses" their presence, it registers their passage through a substance, which used to be a liquid contained in a bubble chamber detector (e.g., Galison 1987), but now is often solid or a mixture of gases. Detectors

are made of many subdetectors wrapped in shells around a beam pipe from which the particles emanate during collision. These subdetectors use different technologies, such as scintillating fibers, wire chambers, silicon crystals, etc. The interaction of particles with detector materials, through liberations of electrons and the emitting of light by electrons, results in the first level of particle traces in a series of three levels. The work on this level is done by the particles themselves. The experiment designs and builds the apparatus in which the particles register. Physicists, however, do not start with the particles, they start with representations of the detector, that is, "offline" manipulations of the signals extracted from detectors after data have been taken. This level of representation *reconstructs* the events in the detector and slowly molds these signals into a form that echoes the particles of interest to physicists. ("Online" manipulations are manipulations during data taking.) Finally, the third level involves representations of physics: from the reconstruction of events *in* the detector, physicists create "variables" that are no longer interpreted in terms of the signs that register in detector materials, but are designed and analyzed in terms of distributions and models in physics (e.g., expected distributions for certain kinds of particles).

Step two in particular includes major "chains" of complicated substeps, which I cannot present here in detail. Suffice it to say that physicists' classification of the major segments of these chains as the work of "production" and "reconstruction" indicates their representational concerns, continued in step three through procedures of "choosing the right variables" to represent physics processes. What I want to illustrate further are the complications that arise in collider experiments, particularly those which work from proton-antiproton collisions, from the fact that their signs marking interesting events are muffled and smeared by signs from other occurrences in the detector. In these experiments the universe of signs and traces is overlaid by a universe of simulations and distortions of signs and traces. These effects derive from uninteresting *parts* of events, from other *classes* of events, or from the *apparatus* itself—they refer to the "background," the "underlying event," the (detector and electronics) "noise," and the "smearing" of distributions. All of these phenomena are a threat to the scientists' ability to recognize interesting events. They may falsify the signature of events, misrepresent their character, or jeopardize their identification. They deceive detectors, and hence analysts, about the presence of events, the shape of their distributions, and the identity of the information they provide. They worsen the results and "resist" (e.g., Pickering 1991) physicists' attempts "to get results out," causing infinite problems to researchers.

The most insidious of these antiforces of the experiment surely is the *background:* competing processes and classes of events that fake the signal. The physicists in the proton-antiproton collider experiments observed for this study see themselves as "buried in background": in the words of a participant working on the top analysis in UA2, "The nature of the problem is to deal not really with the signal so much as the background. You have to deal with the horrible case that you didn't want to see." Their task, as they see it, is to get the proverbial needle out of the haystack. The signs of the events of interest are muted by the background. If you think of these signs in terms of footprints, it is as if millions and even billions of different animals stampeded over a trail, and among their imprints one seeks to discern the tracks of a handful of precious animals. In the original collider experiments at CERN, the counter-rotating bunches of protons and antiprotons met every 7.6 microseconds {*7.6* millionths of a second). During the upgrading of these experiments, collisions occurred every 3.7 microseconds. Each collision produces an "event": a cascade of new particles that spill out into the surrounding detector. But only very few of these events give rise to the particles physicists define as "interesting." In the search for the Z^0 boson at CERN in the early 1980s, less than one event was retained out of every 10,000,000,000 interactions (Barger and Phillips 1987: 31). In the search for the top quark, during the upgrading of UA1 and UA2, for example, it was expected that approximately 40 top events would appear in six million selected electron triggers (electron candidates), a number already vastly reduced from the number of interactions.

The importance of the background manifests itself in a variety of ways in an experiment. First, the background is from "last year's" or the "last ten years'" physics (events studied at lower energy co-occur with new ones at higher energy). In other fields, topics that are "understood" are black-boxed and set aside or become incorporated into new techniques, but in HEP some of these topics keep popping up as "the enemy"—as the background that has to be dealt with if one wishes to learn anything about "new physics." In this sense, the old physics does lend a hand to the new in the familiar sense (Merton 1973), but it also haunts it. Second, backgrounds are specific to specific detectors. Therefore, the background for a particular class of events needs to be studied and taken care of anew with respect to each detector constellation. Third, the task of dealing with the background, of sifting out "interesting" events from uninteresting ones, penetrates all levels of experimental activities. It manifests itself in the way subdetectors are constructed—not only to receive particles, but also to provide a defense against certain backgrounds. It is embodied in a hierarchy of selections, which started in UA2 with a three-level trigger system (a multistage, electronic threshold-setting process through which only some events in numerous interactions are retained and selected out of the detector), continued through "safe," more precise, additional selections in the filter and production program, and ended with "analysis cuts," which are the selections made by "individual" physicists to separate the signal from the background. The flow of activities in the whole experiment can be expressed in terms of a series of selection processes. Finally, interest in this enemy is shown in the manifold distinctions and classifications applied to the background, and in a vocabulary of killing and suppression that I will illustrate in Chapter 5.

The background may be the most insidious antiforce in the experiment, but it is not the only one. Physicists also have to deal with a wide variety of *noise;* random, unpredictable, and undesirable signals in a detector and in the electronics of the apparatus that mask the desired information. A third category of undesired signal is the "underlying event," the result of hadron colliders producing large, inelastic cross-sections[1]—event-production rates dominated by low-energy particles ("spectator" jets) that do not participate in the interaction proper but nonetheless pass through the detector, where they may overlap with events of interest, make the energy of the initial partons (components of nucleons) uncertain, and add confusing extra tracks. A fourth category of obstacle, "smearing," refers to a distortion of physical distributions that makes these distributions wider. The resolution of a detector refers to the degree to which it can give distinguishable responses to two separate particles of two very small energies, or separate tracks for two particles running through the material closely together. Since this resolving power cannot be infinite, the resulting distributions are wider than they should be for physical reasons alone. A second reason for the smearing of distributions is related to the uncertainty of surrounding processes and quantities in physics. For example, the transverse momentum distribution of a W boson is seen to smear the mass measurement of the W because it is not well known.[2]

A final point to add is that different categories of distortion may also interact with each other and work together to co-produce a particular effect. The background events of a particular type produced in UA2's scintillating fiber detector were called "ghosts"—fake tracks of particles misinterpreted as real tracks. They could be caused by pieces of real particle tracks overlapping with one another or with fake information from cross-talk between layers of fiber. Ghosts could also derive from pure electronic noise or from smearing—smearing "fattened" projections, making the resolution of this detector worse, which meant that they picked up more ghosts. Thus, smearing, noise phenomena, and real particles all contributed to the making and the size of this background. At one time, UA2 was swamped with ghosts—"you end up with 50 real tracks producing 2,000 ghosts"—a problem that was hard to handle not only because of the technical difficulties involved, but also because of the sheer amount of computer time needed to solve the problem.

The term *ghost* is a vivid native pointer to the shadow world of signs and appearances in high-energy collider experiments. Through the anti-forces of the experiment, the world of signs is joined by a level of simulated and distorted signals and their by-products—an extra world that sticks to the surface of valid representations as a coating of paint sticks to the surface of an object.

3.3 The "Meaninglessness" of Measurement

I now want to turn to the "meaninglessness" of measurement, an issue that adds to the picture high energy physics presents of a self-enclosed

system and sets it apart from many other sciences. In many fields, measurements, provided they are properly performed and safeguarded by experimenters, count as evidence. They are considered capable of proving or disproving theories, of suggesting new phenomena, of representing more or less interesting—and more or less publishable—"results." This view holds irrespective of the fact that measurements are theory-laden, prone to raise arguments in crucial cases, and sometimes subject to re-interpretation. Within the framework of their dependence on a certain paradigm and tradition, measurements count as self-sufficient quantities; they are granted a powerful role in validating knowledge, and they are considered irreplaceable as witnesses and arbiters in scientific disputes. They provide, one might say, end-of-the-line verdicts; verdicts which experimental work leads to through intermediary and final steps, from which this work takes its clues, and at which point it pauses and starts afresh. In high energy collider physics, however, measurements appear to be defined more by their imperfections and shortcomings than by anything they can do. It is as if high energy physicists recognized all the problems with measurements that philosophers and other analysts of scientific procedures occasionally investigate. As if, in addition, they had pushed one problem to its limit and drawn a conclusion that other sciences have not drawn: Purely experimental data "means nothing by itself." Not only are there few quantities that can be measured relatively directly, but even those are not to be taken as they are. Experimental numbers are dependent upon a particular detector configuration and on the criteria applied in extracting information from the detector. Another detector, another set of criteria, yields other measurements. A UA2 postdoctoral fellow (now a member of ATLAS) reacted indignantly to my insinuation that one might "just measure" the mass of the W: *"You cannot read off a detector how big the mass of a particle is like you can read the time off a watch!"*

The master problem, then, is the detector. As an example, consider the strong force-coupling constant, Alpha S, in effect a measure of the probability of the emission of a force-carrying particle. Alpha S depends on the momentum transferred between two quarks, which in UA2 was assumed to be equal to the squared mass of the W. What one could measure experimentally was the ratio between the number of W plus 1 jet event divided by the W plus 0 jet events. As the physicist who had done this work in his doctoral thesis explained, however, there is a problem:

> As an experimental quantity this number is totally dependent on the detector configuration and on the criteria used in jet-identification. *It is a purely experimental number which says nothing in itself. It is absolutely meaningless.* Because if you took another detector with a greater acceptance, for example, a detector that is almost completely hermetic ((one that covers almost the whole area around the beam pipe such that no particles can escape)), the measured quantity would be much higher. The experimental quantity must be put in relation to theory. If the theory is correct, then one has to find the same Alpha S with the new detector, with a greater acceptance, with the greater value for the ratio.

Thus, what is interesting is not the experimental value, but "the theoretical ratio in relation to the experimental ratio for a given detector configuration." To get this value, one must first determine the experimental ratio described above; second, one had to assemble a Monte Carlo calculation that included all the necessary theoretical calculations *and* simulated the detector. The Monte Carlo also simulated the "fragmentation" (the breakup of quarks and gluons into jets) and the underlying event, etc. From this simulation one obtained the same ratio as the experimental one, in theory. The theoretical ratio was a function of, among other things, the coupling constant. It increased when the coupling of relevant particles increased. The experimental ratio, on the other hand, was a constant. The "real" Alpha S derived from intersecting the experimental value with the "Monte Carloed" curve of the theoretical ratio.

The same procedure, in principle, is necessary to obtain "directly measurable" quantities such as the W mass, as described by a physicist in UA2 and ATLAS who had used the W channel:

> **KK:** ((How do you measure the W mass?))
>
> JJ: (()) If one looks at the experimental spectra ((distributions)) of the W mass, one gets an impression of where the W mass lies. But one has to run an MC to describe the data. Because it takes into account not only the decay properties of the W boson, but also *how my detector reacts to it.* One can see the Jacobean peak in the spectrum of the missing transverse momentum, and where it is in the spectrum. But to know why the peak looks like it does, why it is smeared to the degree it is, and what this means for the uncertainty of my mass measurement, you can only find that out by running an MC.

An experimental measurement in HEP experiments is a sort of *amputated* quantity; a quantity that, without the nonmeasured parts that are missing from it, is not worth much as an experimental result. It is not a final figure that can stand on its own, but a position in a structure of relations to which other positions must be added before the whole becomes useful. With respect to the analogy of the closed universe, this epistemic strategy means that measurements are placed firmly—and obviously—*inside* the ranks and components of the experiment rather than outside of it, as attempted in other fields (see Cicourel 1964, 1974, for the social sciences). They are not cast as external evaluations of internal propositions, but rather as elements and stages that are held in check and turned into something useful only through the interaction of these elements with other features of the experiment.

3.4 The Structure of the Care of the Self

The conception that data are contingent upon the measurement apparatus and are representations of this apparatus (rather than of the world) reveals the detector as a mediating device (Wise 1993) interposed *between* the experiment and the phenomenal word. The detector is the outpost of the experiment; it has to bear the impact of incoming particles that strike its materials. Like the retina, which is hit by the photons that make up raw light and which converts these into the nerve signals that the brain interprets as visual images, the detector is hit by particles and converts their impact into electrical currents and pulses that must be interpreted as physical processes. The analogy between the experiment and the brain as interpreter suggests that considerable internal processing is needed to perform this task. The energy and track reconstructions mentioned before reflect this processing. Reconstructions are based on the premise that one knows the detector and all other components of the measurement machinery. Equally, the idea that measurements must be intersected with detector simulations assumes that the complete measurement apparatus can and must itself be properly understood. HEP collider experiments seek this understanding. They substitute a concern with their own internal production circuitry for a concern with real-time objects, found in other sciences such as molecular biology. To borrow a phrase from Foucault (1986), they substitute the care of objects with *the care of the self.* By this I mean the preoccupation of the researchers with the experiment itself, *with observing, controlling, improving, and understanding its components and processes.* Confronted with a lack of direct access to the objects they are interested in, caught within a universe of appearances, and unwilling to trespass the boundaries of their liminal approach (see Section 3.5), they have chosen to switch, for large stretches of the experiment, from the analysis of objects to the analysis of the self.[3]

3.4.1 SELF-UNDERSTANDING

The "care of the self" becomes obvious, for example, merely by looking at an experiment's expenditure of time. More time in an experiment is spent on designing, making, and installing its own components, and in particular on examining every aspect of their working, than on handling the data. Another indicator is the importance credited to self-analysis in practice and discourse at all points of experimental activities. For example, energy reconstruction is considered a relatively "uncontroversial" step in data processing, but only insofar as one believes that one "understands" the calibrations that go into it. These in turn are part of the "understanding of the behavior of the detector" that comprises a major portion of the care of the self. The detector is an apparatus that is self-created and assembled within the experiment. Nonetheless, the behavior of this apparatus, its performance, blemishes, and ailments are not self-evident to the physicists. These features must be learned, and the project of understanding the behavior of the detector spells this out.

What exactly does one mean by understanding the behavior of the detector? As a physicist working on the inner detector in ATLAS and UA2 put it, it means "knowing when some physics process of some kind happens (in the detector), what comes out of it." To understand means "being able to do a perfect mapping" of it, and "trying to unfold what has happened between an input and an output" of results. Understanding the behavior of a detector begins when the its first components arrive and

undergo several stages of test-bench and test-beam studies, all presenting their own difficulties and time requirements. But it involves much more. For example, learning how a calorimeter (which is central to electron identification) responds—that is, learning "what signal you get out of the detector for a given input"—in UA2 required several steps. First, it required an understanding of the basic response of the detector, which entered into the fixed calibration constants. Second, it required an understanding of these calibrations, which meant examining how the detector response changed over time. With regard to these changes, one needed to learn the long-term deterioration of the response, a slow, continuous "aging" process. Finally, it required an understanding of the short-term "instabilities" in the response by which detectors are also affected, the nonuniform "jumps" in the behavior of phototubes, analog-digital converters, and the basic electronic noise. Figure 3.1 illustrates these different levels of understanding involved in learning the workings of a calorimeter.

Understanding thus refers to a comprehensive approach to learning what happens in every relevant part of the material, what happens over time, and why these things happen. This approach is maintained even when understanding is *not* necessary for the successful completion of ongoing work. To provide an example, the organic-chemical structure of scintillators in a calorimeter dissolves slowly over time, which leads to the deteriorating light response that physicists call "aging" of the detector. During the routine measurement of the degree of aging, after the first (1987) data-taking run in UA2, it was found that the light response in certain compartments of the calorimeter had actually changed for the better; the calorimeter had "gotten younger." This was unexpected and seen as a problem that had to be understood. The effect was soon attributed to the fact that the calorimeter had been "packed" into nitrogen during the run, a gas that was used for the first time in the experiment. To understand the problem, test setups were prepared at CERN and independently in Milan.

Both setups found a 6-8 percent gain in the light response of the scintillators over time, which corresponded to the gain found in the experiment. They also found that if nitrogen was switched off, the response returned to normal. This was taken as an important confirmation of what happened in the real calorimeter—its response also returned to normal when two of its slices were tested after the run with the nitrogen

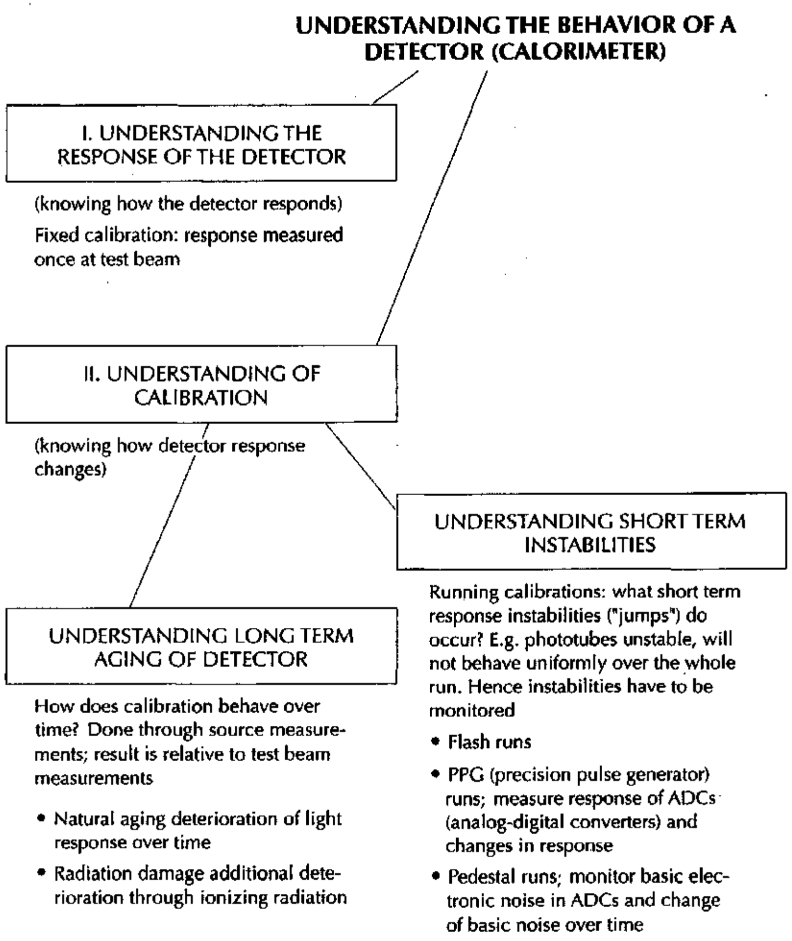


Figure 3.1 Some components of understanding the behavior of a detector.

switched off. It was also decided to use one of these slices and expose it to nitrogen at test beam. Since the test beam could be set to different energies, one could learn more about nonlinearities in the response. Efforts at "understanding" how, for example, the diffused gas became trapped differently in different parts of the calorimeter, continued despite

the fact that this knowledge was not needed to start the next run or perform analysis on the data. In fact, it had been decided almost immediately not to use the nitrogen again because of expected nonuniformities in the light gain—signs of which had been seen at the first source measurements of aging after the run (the light gain was high in the hadronic part of the calorimeter, but nonexistent in the endcaps). It was also known that the response would return to normal after the nitrogen was switched off, which meant that the original calibrations that provided the baseline for measurements of changes, and the enormous work that went into them, were not lost. With respect to the afflicted run, its luminosity was such that mass precision measurements of the kind to which an error of 6-8 percent caused by the nitrogen mattered could not be performed—the statistics were too low. Thus, learning exactly the size of the effect of nitrogen on the light gain, for particles of different energy in various regions of the detector at different points in time, was of no relevance to data taking. Nonetheless, participants of the CERN calorimeter group and from MILAN studied the problem well into the summer of 1988; they went to some length to obtain as differentiated and precise a picture as they could possibly produce.

Understanding the behavior of the detector is only one component of the overall project of understanding. All important elements and forces in the experiment are subject to understanding, among them a considerable number of algorithms, especially those dedicated to particle (track) identification. All the track recognition algorithms that had been written for UA2 before the first data-taking period were called "guesswork games." Comments about them ran along the line that "there surely is a lot more understanding to be gained" of what they do and how they perform. Another example of an item to be understood is the background and the "cuts" one needs to deal with the background—variables that separate the signal from the background (see also Galison 1987). In the following excerpt from a conversation, the cuts refer to the top analysis:

KK: Looking in detail means looking at histograms/statistics ((of the cuts))?

AP: () these cuts are foreseen, but you don't know how tight to make them and we don't know what the background would look like in the

> top sample yet. We know what the background looks like in the old W sample. So in fact before we can really do a top analysis the very first thing is to select Ws (). And then you can really tune the cuts to select those electrons. But then when you go to the top you are dealing with the softer electrons (). So you have to study how well you are doing, the first step would be to apply the W cuts, we even know that those are probably too hard. So you would have to look at your signal ((asking)) well what happened to the signal as you change the cuts (), what happens if I change this cut and make it bigger and bigger, and what have I learned about the cut and how much signal do I (?...). And that sort of thing.

In the following exchange about physics analysis, the need for understanding cuts was summarized as follows:

> KK: How did you do this ((making the cuts))?
> SH: You have to do essentially a study for every cut you put on. You have to study what effect it has on the background, is it buying you something, and how much? Ideally, you want to *understand* each cut individually.

There were also in UA2 a multitude of special test runs, for example, runs in which the trigger setting was changed so that "minimum bias events," which are used to simulate the underlying event, could be studied. Finally, when something goes wrong, when measurements diverge from expectations and problems occur, then understanding refers to an unfolding (see Section 3.6) of what has happened.

3.4.2 SELF-OBSERVATION, SELF-DESCRIPTION, AND RE-ENTRY

The care of the self has a threefold structure: *self-understanding* (discussed above), *self-observation,* and *self-description* (see Figure 3.2).

Self-observation is a form of *surveillance,* present at many levels of these experiments, but especially at its later stages, and during runs. The most clearly specified forms of self-observation involve increasingly more sophisticated levels of online and offline *monitoring,* "More sophisticated monitoring" means that the computer performed the watching, while humans watched computer messages projected onto screens or checked computer printouts. Hundreds of such messages were displayed

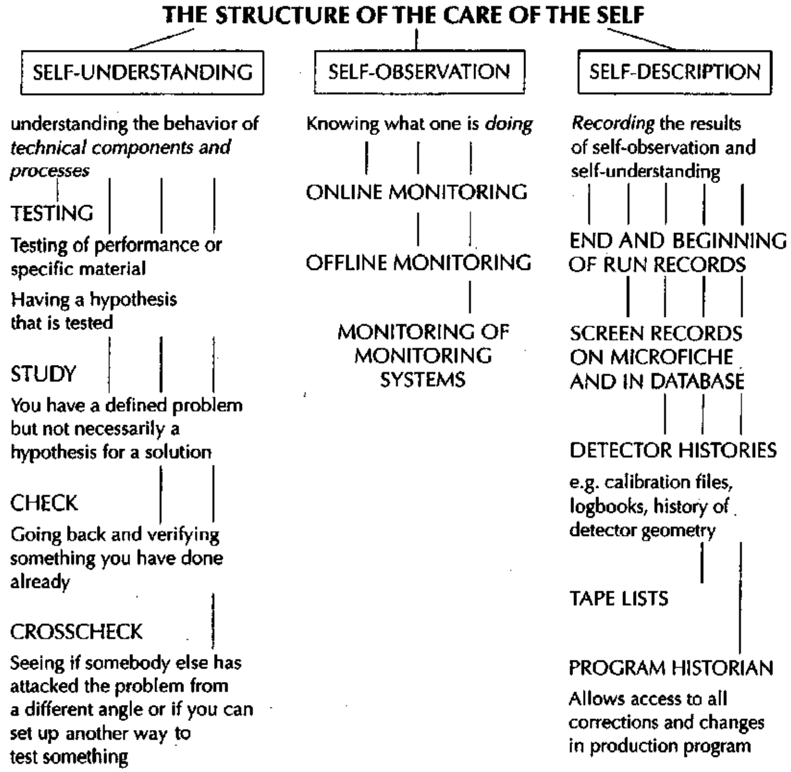


Figure 3.2 The tripartite structure of the care of the self in high energy physics.

on up to ten screens in the "counting room" of the UA2 during data taking. The main messages changed every few moments; they indicated whether the detector was putting something on the tape, which events were being taken, which monitoring tasks were running, and what problems the computer detected. Offline, the data were looked at in a more comprehensive fashion by the production program, which reconstructed every event, while experts looked at varying amounts of the printouts and plots produced by the data. Data-taking efficiency was monitored through a calibration program, performed every few days, that detected dead channels, increasing noise levels, etc. There was also a multitude of special data runs in which some functions were turned off to cross-check these functions against each other. Finally, besides online and offline monitoring there were "checks of the checking system"; different observers watched the same occurrence concurrently and thus were able to detect whether other checking systems worked. They also used special graphic "event display" computers "to see what your experiment is actually doing out there, whether it is doing what you have told it to do, and whether, having told it that, did you really mean something else."

Self-understanding and self-observation are continued through meticulous efforts at self-description, including not only logbook keeping but many forms of computerized recording and track keeping. These efforts generate "histories" of the experiment. In UA2 they included "bookkeeping" information so that people could find their way around the many data tapes: for example, tape lists, records on tapes, and "end-of-run records." To illustrate, the latter alone contained

> The type of record, run number, start time of the run and end time, which tape it was written on, what software triggers and hardware triggers were enabled, number of events in the run, length of tape used, number of words in UA2, which gates, what type of trigger (whether we are running on beam or whether it was a clock trigger), which triggers were enabled at each level, which processing was enabled, which detectors were active, what the pre-scaling factors were on the level 1 triggers, what all the thresholds are in the level 2, what all those parameters were on level 3, etc.

"So in principle we know exactly what was going on online," the physicist coordinating the offline operations said. Besides tape records and equally elaborate run records, there were "detector histories" which included the logbooks physicists keep on all the tests and monitoring tasks they have performed. An important component in relation to detectors was the calibration files—also a record—of the energy scales through which raw signals are transformed into physics quantities and their changes over time. When "running production," one had to know "all the calibration files for all of the experiment at all times." For example, one had to know the pedestal values (the basic noise fluctuations), which changed often and were measured several times daily. They were considered almost as *photographs* of the status of the experiment at the moment the tape was written. Finally, there was the tool called the "historian." This was a program that maintained the main production program, keeping all the old versions on files. As native references to "history keeping" and similar terms suggest, physicists were well aware of the effort they maintain not only to know what they are doing, but also to store this knowledge and keep it available for future reference.

Records such as those described above fulfill many purposes. One is "backtracking" in error searches. Some "old" members of these experiments (in UA2 members who had already participated in it before the upgrade, and in ATLAS those who had worked on UA1 or UA2 before) perform similar functions: they were praised for a memory that, when questions came up, could supply reasons for past technical problems and decisions. Most important, perhaps, and in line with the conception of HEP experiments as internally referential systems, history keeping serves the purpose of re-entering the self-understanding gained through tests, studies, checks, cross-checks, and continuous self-observations into experimental analyses. For example, the understanding physicists acquire about the behavior of the detector re-enters physics analysis in the form of calibration constants, changes and corrections to these constants, quality factors, acceptance calculations, efficiencies, and uncertainties. Re-entry completes a circle, channeling the results of the care of the self back into experimental work.

3.5 Negative Knowledge and the Liminal Approach

Having argued that to assure success HEP experiments turn toward the care of the self, I now want to add that they also turn toward the study of "liminal" phenomena, things which are neither empirical objects of positive knowledge nor effects in the formless regions of the unknowable, but something in between. *Limen* means "threshold" in Latin. The term has been used in the past to refer to the ambiguous status of individuals during transitional periods of time (Turner 1969). I shall use the term to refer to knowledge about phenomena on the fringe and at the margin of the objects of interest. High energy physics incorporates liminal phenomena into research by enlisting the world of disturbances and distortions, imperfections, errors, uncertainties, and limits of research into its project. It has lifted the zone of unsavory blemishes of an experiment into the spotlight, and studies these features. It cultivates a kind of negative knowledge. Negative knowledge is not nonknowledge, but knowledge of the limits of knowing, of the mistakes we make in trying to know, of the things that interfere with our knowing, of what we are not interested in and do not really want to know. We have already encountered some forces of this kind in the background, the underlying event, the noise, and the smearing of distributions. All of these are limitations of the experiment, in the sense that they are linked to the features of the detector, the collider, or the particles used in collisions. High energy collider physics *defines* the perturbations of positive knowledge in terms of the limitations of *its own* apparatus and approach. But it does not do this just to put the blame on these components, or complain about them. Rather, it teases these fiends of empirical research out of their liminal existence; it draws distinctions between them, elaborates on them, and creates a discourse about them. It puts them under the magnifying glass and presents enlarged versions of them to the public. In a sense, high energy experimental physics has *forged a coalition* with the evil that bars knowledge, by turning these barriers into a principle of knowing.

In Christian theology, there was once an approach called "apophantic theology" that prescribed studying God in terms of what He was *not* rather than what He was, since no positive assertions could be made about His essence. High energy experimental physics has taken a similar route. By developing liminal knowledge, it has narrowed down the region of positive, phenomenal knowledge. It specifies the boundaries of knowledge and pinpoints the uncertainties that surround it. It delimits the properties and possibilities of the objects that dwell in this region by recognizing the properties of the objects that interfere with them and distort them. Of course, if one asks a physicist about "negative knowledge" he or she will say that the goal remains to catch the (positive, phenomenal) particles at loose, to measure their mass and other (positive, phenomenal) properties, and nothing less. All else is the ways and means of reaching this goal. There is no doubt that this goal is indeed what one wishes to achieve, and occasionally succeeds in achieving, as with the discovery of the vector bosons at CERN in 1983 (Arnison et al. 1983a,b; Bagnaia et al. 1983; Banner et al. 1983). My intention is by no means to deny such motivations or their gratification, but what is of interest as one works one's way into a culture is precisely the ways and means through which a group *arrives at* its gratifications. The upgrading of liminal phenomena—the torch that is shone on them, the time and care devoted to them—is a cultural preference of some interest.[4] For one thing, it extends and accentuates what I call HEP's negative and self-referential epistemics. For another, the majority of fields, among them molecular genetics, does not share this preference. And, lastly, it is quite remarkable how much one can do by mobilizing negative knowledge.

There are two areas in which the liminal approach is most visible: the area of *errors* and *uncertainties* and the area of *corrections.* Let me start with the latter. The idea of a correction is that the limits of knowing must enter into the calculation of positive knowledge. For example, "meaningless" measurements can be turned into meaningful data by correcting them for the peculiarities and limitations of the detector. "What you really want to know," as a physicist summed up this point, "is, given that an event is produced in your detector, do you identify it." Corrections can be characterized as "acceptances", or as "efficiencies." An acceptance tells physicists "how many events my detector sees of what it should see"; it is the number of observed events divided by the number of produced events.[5] An overall acceptance calculation requires a detector response model: it requires that all the physics processes that end up in a detector are generated, and a full detector simulation is created to ascertain what the detector makes of these processes. In UA2, the detector response model also included such components as a simulation of the underlying event and detector performance measures, such as its geometrical acceptance (which describes how many events are lost through the incomplete coverage of detectors with dead angles and cracks), its resolution (which refers to the smearing of distributions described earlier), and its response curves (which determine the reaction to energy inputs that deviate from those of the test beam used to determine the basic calibration constants). Computing the overall acceptance is a complicated and elaborate effort, which brings together nearly all the results of the self-analysis of the experiment into its final operative stage. A physicist from an outside institute working in UA2 summarized this as follows:

> **KK:** ((How would you characterize the acceptance))?
> NO: Inside the UA2 apparatus we produce an object we want to measure. To do that we have to exploit to the best of our knowledge how UA2 works. This knowledge is gained by test beam studies etc., it is a complete structure of layered information. () (This information) is the complete opposite of the basic piece of a single efficiency. *It contains all we know and all our bias, both in the experimental set-up and in the procedure of how we analyze this complex structure.* [Emphasis added.]

Acceptance calculations are joined by "the basic pieces" of efficiency corrections, which tell physicists, once they have a particle in the sensitive region of the detector, whether they have *identified it* as the particle they are looking for. Many components of these experiments have efficiencies: triggers, special conditions set, particle reconstruction programs, fit methods, and, above all, cuts. Overall efficiencies are computed from many such components. For example, the paper published on the top quark (Akesson et al. 1990:182) lists six components to the overall efficiency used to find a W electron as part of the signature of the top quark:

$$E(W) = E(\mathrm{cal}) \times E(v) \times E(\mathrm{trk}) \times E(\mathrm{Si}) \times E(\mathrm{ps}) \times E(P)$$

where $E(\mathrm{cal})$ — the efficiency for finding an electromagnetic cluster from an electron candidate in the central calorimeter; E(v) = the vertex-finding efficiency (the vertex is the point of interaction in an event and the point where the particle tracks coming from the event should coincide); E(trk) = the track-finding efficiency for isolated high energy tracks; E(Si) = the efficiency of a cut made with the help of the silicon detector to reduce background; $E(\mathrm{ps})$ = the track-preshower matching efficiency for finding the impact point of candidate electron tracks with the position of electromagnetic showers; £(P) = the efficiency of a cut eliminating candidates with an energy greater than 1 GeV in the second hadronic compartment of the calorimeter. In a 1991 study of heavy particle decays in two jets (of

particles heavier than the W or Z°), the overall efficiency investigated included eight components, ranging from the trigger efficiency to that of a mass cut.

Correction factors like efficiencies may themselves be corrected for by other factors—there is, in a sense, a hierarchy of corrections. For instance, if not all particle tracks in a reference sample used to calculate a tracking efficiency are genuine, the efficiency may be corrected by determining, through another detector, the fraction of genuine tracks. Such secondary corrections can be seen as part of the *background evaluation*, a third major correction task in the experiment. A hierarchy of "cuts" (selection procedures) is normally implemented to "get rid of" the background, but the background that appears with exactly the same signature as the signal cannot be removed through cuts. It must be evaluated through other means, such as through Monte Carlo simulations, and then be subtracted from the signal.

Corrections of corrections indicate how the liminal approach extends to self-knowledge—it, too, includes its very own limitations, in terms of which it can and must be specified. Besides secondary corrections, and the original ones, "errors" and "uncertainties" are the other major components of an experiment that exemplifies the liminal methodology. Physicists distinguish between *statistical errors* (the variation obtained from repeated measurements, a new limitation) and theoretical and experimental *systematic errors*, which indicate fundamental mistakes one makes in ones' procedures. If the first is variation due to random error, the second is similar to using a ruler that is too short for measuring the length of an object and will always yield biased results, no matter how many times the measurement is repeated. As one physicist, then a member of the CERN group in UA2 and now a member of an outside institute working on ATLAS, put it:[6]

> The systematic error is just a way to measure our ignorance . . . (it) is our way to try and estimate what we've done wrong. And if we knew what we did wrong we could correct it rather than to allow for it in the error.

I want to illustrate this by an example that also shows how the existence of different theories is transformed into an uncertainty calculation, or into liminal knowledge. Structure functions describe the density of quarks and gluons (collectively called *partons)* within the proton. They are used to obtain the number of expected events produced in a proton-antiproton collision and the kinematics of the collision. Structure functions are based on extrapolations of experimental measurements from low energy fixed target experiments (which exist) through "parametrizations"—that is, through fitting curves that match the available data points. They predict the values of unmeasured densities in HEP experiments. In early 1991, about 45 sets of structure functions were available, involving different assumptions in their calculations (see also Plothow-Besch 1990).

The systematic errors in this case are "uncertainties"[7] deriving from the fact that one cannot say which of these structure functions is correct. They represent, as a senior physicist performing the analysis of structure functions said, different theories:

> One structure function might lead you to this value, another to that, etc. If these values would result from measurements you could construct a broad Gaussian out of this with an average and a sigma. If you treated it as a Gaussian, you would say that there is a 68 percent chance that the real value is within (± 1 sigma) . . . But these are not measurement errors, these are different theories, and for the moment we have no way of telling which is right and which is wrong. All of these values are equally probable .. .

Physicists normally select, she added, a number of these functions for their cross-section measurements, regarding the variation between different functions (the spread between the curves) as the theoretical systematic error or the uncertainty associated with them. She made it clear that this procedure must be improved by making it possible for physicists to use *all* the structure functions on the market, provided they were extrapolations from recent data, so they could calculate their uncertainties from the complete set (the program for this became available in 1991; see Plothow-Besch 1990). The point was to enlarge the basis for the uncertainty calculation by taking into account all functions (theories, interpretations) physicists had imagined to this point, and to update the basis with every new interpretation. The interesting point

for us is that these practices are an attempt to gather up the dissension in a field with respect to a particular issue and to transform it into a measure of (liminal) knowledge.

To scientists in other fields, the procedure of turning variations between answers into a precise measure of current knowledge limitations is quite stunning. Would sociologists care to consider the variability between different theories on a subject as a source for making a calculation of their theoretical error? Different theories in sociology—or in molecular biology—give rise to scientific arguments, and to the formation of different groupings of scientists divided along their theoretical preferences, but never to error calculations. Would it make sense in these fields to require that the dispersion of different theories should somehow be ascertained so that we would know, if not what is right, then at least how far we might go wrong? Of course, sociologists and biologists do not make primarily quantitative predictions. But this difference is hardly enough to account for the disinclination to exploit liminal phenomena in these fields.

I can only add two further illustrations of the turn toward liminal knowledge in HEP experiments, one concerning refinements between different analyses, the other concerning the kind of analysis that is routinely produced. Refinements between an analysis and a later one—done, for example, after more data have been accumulated in an experiment—frequently involve changes in error and uncertainty calculations, such that a previously large error may be estimated more precisely and become a correction accompanied by a smaller error. Thus refinements bring about shifts in the weight and distribution of liminal items rather than their elimination. As a physicist searching for SUSY (supersymmetric) particles in UA2 put it:

> Some of the things which we just put as large uncertainties we've tried to make more precise estimates of the effects and then, correct for it, and then assign a smaller uncertainty. Essentially, that has been the sense of, besides the increased statistics, of our efforts to improve the precision of the mass measurement.

Furthermore, greater precision with regard to uncertainties in an improved analysis may mean that one identifies *more* of them. In the analysis of the W mass, seven corrections and uncertainties were listed in the original published paper (Alitti et al. 1990). The improved analysis, however, listed ten uncertainties: "Some subtle effects not considered previously are now taken into account," as the presenter of the work at a collaboration meeting said. When I asked participants about this they said "it makes them uncomfortable" to see such a long list of errors. Nonetheless, an improvement in knowledge might just as well be an imaginative extension of the universe of errors and obstacles to knowledge as it might be an extension of positive knowledge.

The type of results routinely produced in the respective experiments, my final point in this section, illustrates a similar epistemic strategy. I am referring to "limit analyses," which identify the boundaries of a region within which a certain physical process can be said to be unlikely. Limit analyses offer ways out of the problem of negative results. If the top that one searches for in the data is not there, it is at least possible to say "up to a mass of 64, or 70 GeV for which we have searched the terrain the top is unlikely to occur." Limit analyses usually produce limits on the number of events. The number of events reflects back almost immediately on the mass.[8] "So if you look for something and you don't see it (in your data) then you find out at which mass would you start to produce large enough numbers of these things so that you *should* have seen them," said a physicist looking for squarks as a function of squark mass. The limit indicates the boundary beyond which this large enough number of particles should have occurred. From the limit, one concludes that the particle one is seeking cannot be lighter (have less mass) than the mass indicated by this boundary. In the search for supersymmetric particles, for example, no events were seen in the data sample. The mass limit was determined by generating an expected signal and seeing how much one could vary the mass of the supersymmetric particles while remaining compatible with the zero events seen in the data.

Limits are perhaps the most frequently produced results of experiments. Even in experiments designed to produce precision mass measurements of known particles, such as the LEP experiments, limits may be the more frequent output; LEP is said to produce "a stream of papers where they put limits on all sorts of things." The point of interest here is that limit analyses are yet another way to mobilize negative knowledge and

turn it into a vehicle for positive knowing. The limits experiments produce are the thresholds from which new experiments start their searches.[9]

3.6 Moving in a Closed Universe: Unfolding, Framing, and Convoluting

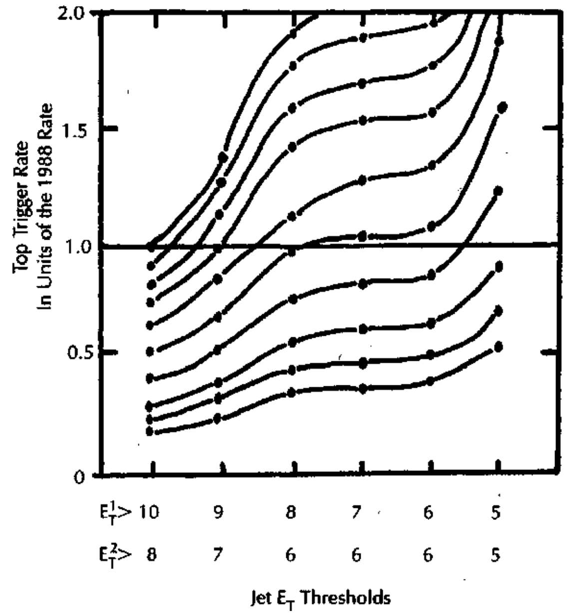
In this last section of the chapter, I want to bring into focus some of physicists' ethnomethods of moving in an internally referential system and of articulating such a system. I shall use simple examples, limiting myself to the practices of unfolding, framing, and convoluting. By *unfolding* I mean the continuing unraveling of the features of physical and technical objects, of their details, composition, hidden sequences, and behavioral implications, through the reflexive redeployment of the approach to the data points generated. Let me illustrate with an anecdote. In 1989, the head of online operations was worried about the major online computer in the experiment, the device that wrote the incoming data to tape. The computer was old. Would it make it through the next data run or should it be replaced? What the physicist wanted was a curve that unfolded the deterioration of the device:

> I tried to get some data out of Digital, the firm making these computers, by asking them, since it's an old computer of the kind of which they built (a) 100 if they could make a statistical analysis and see what is the uptime ((the percentage of time the computer is running)) of the computer after 1 year, after 2 years, after 10 years, after 15 years.

He imagined a graph that would give him the probability of the computer's uptime as a function of its age. Suppose in this graph the probability of uptime remains at 100 percent during the first years and then decreases by 10 percent each year during the middle age of the instrument before dropping off more sharply and finally flattening out in a tail. Would that curve make it any clearer when one should get rid of the device? No, since one would still have to decide on the level of uptime that was acceptable. But in many other ways it should. First, a steady decrease of performance by 10 percent per year up to a certain age is itself important information that might prompt one to make the cut at the point where the curve begins to drop off more sharply. Second, the curve can be used to make further calculations, for example of expected data losses at specific levels of downtime, of subsequent error increases, and of reductions in the probability of finding the result one is looking for. Third, these probabilities in turn provide starting points and baselines for the exploration of additional considerations, such as the expense of a new computer versus the expense of repair costs and the training time involved in the switch. And so on. Thus, unfolding a problem by articulating it in a detailed curve is a machine for generating further unfoldings, for internally articulating the information through redeploying the procedure and adding further loops. One can also say that the ethnomethod of unfolding "unsticks" a stuck decision system by transforming it into a knowledge system (see also Knorr Cetina 1995).

If the object is not the kind whose characteristics one can unfold, then variation is used to achieve a similar effect. This can be illustrated by a typical figure resulting from a trigger study performed in an effort to find a more efficient top trigger algorithm for use in the 1989 collider run (see Figure 3.3). The plot gives a vivid impression of how variation unfolds variables, in this case top trigger rates for different missing momentum cuts as a function of jet transverse energy thresholds (for the details, see Incandela 1989).

The second epistemic strategy I want to look at, *framing,* is to consider objects or pieces of information in the light of other such components, which serve to check, control, extend, or compensate the former. Through framing, different components of an experiment or of the field are related to one another. The relationship may be confrontational (the components may be played one against the other) or complementary (the second object may achieve what the first could not accomplish). Framing is used on all levels of an experiment. For example, detectors are built to verify each other, to act as reference counters in efficiency measurements, to eliminate each other's fake signals, to complete the partial track segments any one of them can produce individually, or to produce track samples to be used as "templates against all other detectors." Besides detectors, different analysis modules (e.g., different fits to data points), different data runs, and whole experiments frame each other. Runs, for example, are constantly produced with some functions switched off so that data produced by other runs may be evaluated, in regard to the



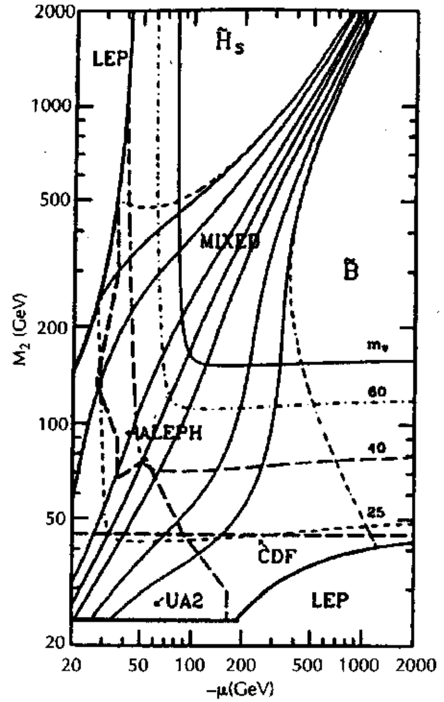
Trigger Rates normalized to the 1988 top trigger rate are shown for different missing P_T cuts as a function of jet E_T thresholds. The E cut for the electromagnetic cluster in the event is 11.5 GeV. (The 1988 top trigger rate was ≈7.0% of the combined W^l and Z° trigger rates.)

Figure 3.3 Figure resulting from the use of variation in a trigger study. (From J. Incandela, *pp Note 566,* 10 March 1989.)

working of different pieces of the apparatus, the composition and behavior of liminal phenomena such as noise, the underlying event, etc.

One of the most interesting framing strategies is perhaps the institution of "sister experiments"; experiments are set up at the same time in the same laboratory with the purpose of, among other things, comparing each other's results. Typical examples of sister experiments at CERN were UA1 and UA2, the four currently running LEP experiments and the ATLAS and CMS experiments at the LHC. There is more than one reason why sister experiments come about, most importantly perhaps the desire of more physicists to work at a new machine than can comfortably be grouped into a single collaboration. Nonetheless, the way they are organized is similar to other components in the system—as quasi-independent epistemic resources that sustain and control each other. The participants of the sister experiments at CERN often seem to know surprisingly little about the progress of their siblings. There are few occasions for systematic contacts, and participants learn much of what they do know through informal channels. Sister experiments compete in the pursuit of the same or similar goals, but they do not simply replicate each other's results (compare Collins 1975): they work and run in parallel, not at successive points in time, and their design and apparatus differs significantly in crucial respects. UA1, for example, had a magnetic field and a muon detector, whereas UA2 did not. Framing is not based on a principle of similarity in all significant details but on a principle of differences within the limits of similar overall goals. Their partial independence allows sister experiments to be featured as controls. Their differences allow them to compete, to test themselves against each other, and to drift apart in various directions.

Framing also occurs among experiments that succeed each other in time and are located at different places. Its most obvious manifestations are exhibits in experimental papers, which often include comparable distributions from other experiments (see Knorr Cetina 1995). Much like the variations deliberately created to generate bands of uncertainty around possible true values, these distributions create ranges or successive boundaries of experimental values into which the values of new experiments fit. Figure 3.4 gives an example of a multi-experimental display of results. By this framing strategy, physicists convey a more urgent lack of trust of single experiments and measurements than we see in other fields, a disbelief that these experiments can stand on their own as they do elsewhere, supported exclusively by the special linkage they have created between themselves and nature. High energy collider experiments stand first and foremost *in relation to each other,* as these figures imply. The exhibits epitomize a belief in the epistemic relevance of a strategy that joins individual findings through contrast, comparison, and resemblance. The display that includes data from different experiments becomes a display of the state of the field (see also Chapter 7). In a sense, the strategy extends the liminal approach of emphasizing limitations and possible errors and of gaining confidence through this emphasis.



Contours of the LSP mass (short-dashed: 25 GeV, long-dashed: 40 GeV, dot-dashed: 60 GeV, solid: m_w) in the (μ, M_2) plane for tan $\beta=2$. Also shown are the contours (solid) of constant gaugino purity ($N_{11}^2 + N_{12}^2 = p$), with $p = 0.99$ (almost pure gaugino), 0.98, 0.9, 0.5, 0.1, 0.02, and 0.01 (almost pure higgsino), as well as the contours of the almost pure bino $\tilde{B}$ and (anti)symmetric higgsino $\tilde{H}_{A,S}$ (short-dashed), with purity $p = 0.99$, as defined in the text. Regions where $0.01 < N_{11}^2 + N_{12}^2 < 0.99$ are marked as "MIXED." The allowed region is constrained by several experimental data. The results from ALEPH and CDF are preliminary.

Figure 3.4 Published results from one experiment in particle physics (UA2) "framed" by the results of other experiments. (Reprinted from *Physics Letters B,* vol. 262, no. 1, 13 June 1991: Y. Okada, M. Yamaguchi, and T. Yanagida, "Renormalization-Group Analysis on the Higgs Mass in the Softly-Broken Supersymmetric Standard Model," p. 61; copyright 1991, with kind permission from Elsevier Science–NL, Sara Burgerhartstraat 25, 1055 KV Amsterdam, The Netherlands.)

Let me conclude by turning now to what is perhaps a special case of framing—the case where the components are of a very different nature and where the logic of the procedure is based on the conjunction and interaction of these components, or, to use a native concept, on their *convolution.* Perhaps one could say that such intersections between different universes of meaning—between, for example, "Monte Carloed" self-understandings through which theoretical numbers are run and experimental measurements—are a way to push the framing method to its limit, by showing that only the convolution of the different frames can lead to real results. Monte Carlos in general (numerical integration methods for simulating the properties of systems through random walks) play a significant role in the convolution of experimental components, but I cannot examine them here in detail; their manifold uses in HEP experiments require a proper analysis in their own right (for their history, see Galison 1987, 1993; Merz 1998). Suffice it to say that *convolution* is a term used for folding together, in Monte Carlos, mathematical calculations, phenomenological terms describing the "breakup" (fragmentation) of quarks and gluons into jets, simulations of the underlying event, and the modeling of detectors and other components. But it is also a term that describes well the general strategy of mixing together resources and quantities that come from very different origins in an attempt to come to grips with the limitations of specific data or approaches. Structure functions, for example (see Section 3.5), which are often associated with the "theory part" of value calculations, have an experimental portion (data from deep, inelastic lepton nucleon fixed target experiments) and a theoretical portion (the Altarelli-Parisi evolution equations, which make the transition from fixed target to collider experiments), and, as a third component, different fits through which appropriate mathematical functions are fitted to the original data. The fragmentation of quarks and gluons into jets, which is also simulated in these Monte Carlos, has a phenomenological portion—a theoretically motivated but not exactly calculable "model," which is exemplified by the "Field-Feynmann" fragmentation supplemented by a model of soft gluon radiation; and a data component—model parameters like the number of gluons and their angle of radiation are adjusted to calorimeter data until they coincide reasonably well.

Mixtures are also constantly created between data and Monte Carlos. For example, when one cannot play different parts of the detector off against others to measure the tracking efficiency of an algorithm, one can take a real event and "stick in its middle" (superimpose on it) a single Monte Carlo track. The question then asked is whether the algorithm can

find the test track that one knows is there. Another example is the analysis of Alpha S briefly described in Section 3.3. Alpha S is determined for a certain momentum transferred between two quarks in an interaction that, in the analysis performed, was set equal to the W mass squared. Through experimentation one could determine a ratio between different W events. The same ratio was generated through Monte Carlos in the above fashion: including the detector model and all the other components that finesse the end result and lead to a determination of systematic errors. Since the ratio is proportional to Alpha S in the Monte Carlo, one ran a great number of these to obtain different values of the ratio as a function of Alpha S. The result was a curve that was then entered in the same plot as the experimental value for the ratio. Where the curve for the expected values of the ratio intersected with the experimental value, the "true" value of Alpha S could be read off the plot.

In the original case, the "convolution" of theoretical functions with detector simulations and other components represents, at least in part, convolutions that are thought to be in the data. All data are seen as mixtures of components—including, besides the signal, the background, the noise, and the underlying event—which are distorted and truncated through detector losses and smearings. "I cannot think of a case where (the data) is not mixed." This is where we come almost full circle, back to the starting point of the meaninglessness of measurement and the presence of antiforces in the experiment. Participants react to these convolutions by matching them with precise determinations of the single components, *and by converting the principle of mixtures, which they find in nature, into an epistemic resource in their own endeavors.* They draw upon the different possibilities of calculating mathematical functions theoretically, of simulating events that are neither easily calculable nor measurable, and of making measurements experimentally; and they bring them together continuously by specifying the results of experiments. They do not shy away from extending data points through theoretical or Monte Carloed predictions, from superimposing one element upon another, or from correcting and revising data through Monte Carloed theories and predictions. The crucial difference this approach has from some other fields is that the *convoluted tangles* thereby created are expedient in creating *experimental outcomes.* They are *not* tests of theories through data, simulations used for purposes other than data processing (as in the planning and design of an experiment), or similar conjunctions of elements before and after the specification of data. Convolutions of experimental data with Monte Carlos and theory precede the framing strategy outlined earlier. But they are also an extension of it, as when they occur in the interplays created between the respective elements. They are a stratagem in generating experimental outcomes in a world that refers back upon itself and seeks recourse in manipulating its own components.

try again some combination—I try to exhibit the articulation, in each science, of a basic form of order, but also the infiltration, in each field, of these forms of order by additional orderings that supplement, contradict, or in some sense dissipate the other. You see I am not very consistent in my strategies.

R: We all can learn! Though it might be harder for the type of unashamed empiricist you seem to be. Perhaps I shall do some further deconstructions in these footnotes.

KK: Right on! Like the bathroom, Anthony Grafton says in a historical paper on the footnote—given at the Princeton Davis Center colloquium in 1993, a paper now part of a book (Grafton 1997)—like the bathroom, the footnote enables us to deal with ugly tasks in private. Like the bathroom, it is tucked genteelly away. I shall accept your obstinate meddling, especially if we agree that the text persuades, while the footnote has only secondary significance!

R: Actually, Grafton says, the footnote is like an engineer's diagram of a splendid building. It reveals the occasionally primitive structures, the unavoidable weak points, and the hidden stresses that an elevation of the facade conceals . . .

3 PARTICLE PHYSICS AND NEGATIVE KNOWLEDGE

1. PHYSICIST: In the spirit of our discussion of the footnotes (which are actually endnotes!) at the beginning of Chapter 2, let me get back into the picture. I shall be the physicist who read your chapter and suggested improvements, which, as I can see, you mostly included. However, some sloppiness in your language remains. A hadron collider, for example, does not produce large inelastic cross-sections because . . .

KARIN KNORR CETINA: Wait one moment. If we are going to bring in some context to the statements I make in the text we might as well be precise. Several physicists read this particular chapter, and many more listened at various places to talks I gave about HEP. Will you be the "generalized other" or any particular one of them? I am sure that those of my readers seasoned in recent discussions of ethnography would like to know.

P: We are not just bringing in context, we are going to reveal some of the occasionally primitive structures, the unavoidable weak points and the hidden stresses that your text conceals, just as Grafton says, if only by way of a few examples. As to who I am, you know me very well. I am a member of UA2 and now of ATLAS, you transcribed my comments, I taught you a few things about physics analysis, and now I shall simply stand for others. I am specific, but my name is of no relevance.

KK: Your evasiveness gives me the opportunity to take the literary freedom of editing and summarizing your comments, and of joining them with a few other comments. Go ahead!

P: You are also selecting them, as the reader should know. I began by saying that there is some sloppiness in your language. A hadron collider does not "produce" large inelastic cross-sections because the large inelastic cross-section is just the probability for a given type of event, which is nothing produced by an accelerator. It is just that as a *consequence* of this large inelastic cross-section the accelerator produces a large number of events of this type. The cross-section is a physical fact.

KK: But I am quite sure I took this wording from a physicist.

P: The cross-section is a physical quantity like mass or charge. The low energy events are more common because the cross-section is many orders of magnitude larger than the cross-section of high energy events, but the accelerator produces events, not cross-sections. Sometimes physicists speak a bit carelessly about these things in conversations.

KK: I think this sort of formulation can also be found in technical writings, though I would have to go back to the literature to show you where.

P: It's like saying the collider produces the big Z mass which of course it doesn't. The Z mass is what it is.

KK: Right. Mhm.

2. P: There is an intermediary step missed here. The fact that the transverse momentum of the W is important is because we are not really measuring the W mass itself. The smearing here comes in because we measure a correlated quantity, not the transverse momentum on the W.

KK: But is this possibility not implied, since I talk here only about the uncertainty which surrounds the measurement of certain physics processes and quantities?

P: It's almost there. I admit this is a very detailed comment. You say the transverse momentum distribution is seen to smear the mass measurement of the W because it is not well known, but if we really measured the mass it would not matter at all. It is not only the fact that the W mass is not well known that makes this smearing. It is also the fact that we measure the mass in an indirect way. The relation between what we measure and the quantity we are trying to establish gets blurred because of that. It's a very detailed comment.

KK: I can at least put a footnote there to point this out.

3. P: This is a point I am coming back to from our private conversation, because I think it is important, and because you haven't done anything about it in the text! You say in several places, and this is one, that we "choose" an approach or some strategy, as a cultural preference, part of our culture. Here you say we choose to be obsessed with understanding ourselves, and later you say we are obsessed with (systematic) errors. I, however, don't think we have a choice.

KK: You think you haven't the choice?

P: I think we have to act the way we act. The reason we must be interested in self-understanding is because there is only one experiment like ours. If there was only one thermometer in the world, we would have to spend a lot of time trying to find out how it works. Some people once spent a lot of time trying to find out how these thermometers work and then they put a scale on them. If there were thousands of experiments like ours, someone would have found out how we work and would give us just a book saying "if you are getting this measurement in the detector, the following thing has happened inside." But we have to learn from scratch how everything works. So I don't really think these strategies are part of our culture. I think they are necessary.

KK: I find it interesting that you compare your experiment with a thermometer, with an instrument. Isn't it much more than an instrument? Also, you seem to think that because something is cultural it is irrational. Whereas I believe cultural preferences usually have good reasons, but the reasons are local. I argue from a comparative point of view. What I have in mind, for example, are disciplines like the social sciences which, like you, deal with aggregate phenomena, yet their concern with self-understanding is rudimentary compared to yours. Our error potential is at least as great. There are the uncertainties of measurement (Cicourel 1964), the smearing of distributions, there are underlying events—we often measure more than one event simultaneously with our data—we have plenty of fake signals, and great uncertainties surrounding our theories. We have lots of reasons to deal with our measurement process. I was not implying that you should be doing something different, but rather that I know a number of disciplines which, given the problems they face, should have a more developed concern with their "self" than they do. Epistemic strategies do not seem to simply flow from objective constellations of problems. The necessity you feel becomes a contingency when you see that other disciplines, confronted with similar problems, develop different strategies.

P: But you say here it is us who choose a different road, whereas I say we do not have a choice. If you use apparatus which is unique, you have to do things in the way we do. You can measure the temperature of the water in your bathtub or you can measure the temperature outside, once you have a thermometer you don't need to recalibrate it every time you measure something.

KK: Our measurement instruments, for example, our questionnaires, are usually unique too. So are the designs molecular biologists use in their experimental activities—there are very few attempts at exact replications of experiments, as we know (e.g., Collins 1975; Mulkay 1985). Moreover, the positive use and mobilization of errors and uncertainties in your case is not just relevant with respect to calibration; you also deploy the strategy when you deal with different theories, it is just one component of the shift toward

the study of liminal phenomena which you do not deny. When one talks about cultures one usually has in mind a whole system of procedures which somehow hang together and sustain each other. Given the framework of the system, every single strategy appears necessary. But the whole is a specific development and articulation of strategies.

THEORETICAL PHYSICIST: I would like to jump in on this and support my fellow physicist, albeit from the point of view of a theorist. The calibration example does strike me as illuminating. What came to my mind when I heard you talk about the need of these experiments to understand their detector are scales. Suppose I want to know my weight. To know my weight I need an instrument—some way to measure what I am after, plus the exact knowledge how this instrument reacts. What, if I put five pounds on it, it will indicate, its errors and efficiency, and the systematic error I make by using this instrument rather than another. What more is an experiment in HEP? Is it not just the testing of scales for measuring something that comes from theory? The weight, the thing we are really interested in, are the theoretical predictions. Experiments are, in the last analysis, preparing and calibrating the equipment.

KK (to TP): Frankly, I am not surprised by your comment. It corresponds to an understanding I have come across over and over again by theorists—the understanding that the intellectually important ingredient is physics theory, while the experiment is just what we imagine happens on the scales (or thermometer) level, writ larger. What interests me about this understanding is how it renders HEP experiments invisible despite their size and complexity, and despite authors, such as Galison, Traweek, and Sutton (and perhaps even Pickering from the theory side), who bring them to our attention.

READER: You forget, if you allow me to come back just briefly from the notes in Chapter 2, you forget the role our disciplines, philosophy and history (but also sociology), have played in promoting this understanding, by paying attention almost exclusively to physics theory. You also forget that physics is a segmented universe in which fundamental theorists and experimentalists live, work and think detached from each other, without knowing much about each other. Theoretical physicists often simply do not know what goes on in experiments. You also forget that studies such as the above are only appearing in the last ten years.

PWA (second theoretical physicist): The question by the theorist should be answered, nonetheless. Perhaps I may point out what I, as another theorist (however, a theorist working at Bell Labs!), wrote more than twenty years ago in an article in *Science* (Anderson 1972: 393), namely that the main fallacy in the above kind of thinking is that the reductionist hypothesis on which it is based does not by any means imply a "constructionist" one.

KK: The reductionist hypothesis supposes that once theoretical predictions have been made there is nothing left but "device engineering"?

PWA: Sometimes. What I said was that "the ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe. And the more the elementary particle physicists tell us about the nature of the fundamental laws, the less relevance they seem to have to the very real problems of the rest of science." I was not talking about experimental HEP then. But it seems to me now that it is a case in point within particle physics. Experimental HEP is a constructionist project. It needs to reconstruct conditions that existed at the beginning of the universe, and design and develop the equipment to do so. It needs to create an operation in which fundamental particles, like the top quark, and qualities of particles *not* predicted by theory, like the Higgs mass, may reveal themselves—and this is an operation no less complex, time-consuming and energy-expending (and only somewhat less costly) than perhaps the military invasion of a foreign state, or a space exploration. It needs to exactly predict the environment in which this operation takes place and its own future states in order to conduct the research. In short, it operates at an entirely new level of complexity at which new properties appear, understanding of new behaviors is required, and research that is as fundamental in its nature as any other is needed. None of this can be done by a simple extrapolation of the standard model.

KK: In my terminology, we are in a different but equally fundamental epistemic culture. Whose existence is concealed by the emphatic articulation of physics theory—and the celebration of individual physicists—in the more popular press and the academic literature. For example, as Galison said in the preface to his book on *How* (physics) *Experiments End* (1987: ix), "despite the slogan that science advances through experiments, virtually the entire literature of the history of science concerns theory."

P: You are, of course, right about the annoying neglect and misconstruction of our project. But I reiterate, you are wrong in calling what we do cultural. To go back to the text, look at Section 3.5.

4. P: Here you go again. And also later on when you talk about variation. The point is that there is only one experiment like ours. The more you know about how your detector reacts, the better you can tell what happens, and that is always true. There is nothing magical about our use of errors or variation. You have to subject a tool to variations such that you understand it. It is not a cultural preference. These other sciences you were thinking of use tools that exist in large numbers and have been calibrated . . .

KK: Actually, one attempts to standardize them, as I said earlier (3.3)—think of the standardized questionnaires in sociology—but standardization is not calibration. Standardization just means trying to use the same stimulus or the same procedure. In molecular biology, one also attempts to use standard procedures, the procedures fixed in laboratory protocols. The point of the laboratory work is to get the procedure to work—to use it in a way which brings the desired results. The point is not to understand exactly why or how it works, in particular what systematic and statistical errors, uncertainties, efficiencies, and response functions exist in all composing steps. A protocol is like a recipe, you use it without necessarily caring about the chemical, biological, thermodynamical and other processes which make it work.

P: I have been thinking a lot about biology, trying to understand how they work and why we are different. One of the reasons is that when we look at matter at the quantum-mechanical level, we get probabilities. For example, you don't know if you produce an electron pair, and because of the width of the Z particle, the mass of that electron pair is going to be 91 GeV plus minus 2.5 GeV. There is no unique number you can get on this. But if you take two hydrogen atoms and one oxygen atom they always make one water molecule. Chemistry at that level is not probabilistic, it is completely predictable. I believe it's the same in molecular biology. If they have a certain enzyme that cuts at a given sequence, it always cuts at that sequence and nowhere else. And if they run their inverse chain reaction, they always produce a particular molecule.

KK: Molecular biology may not be probabilistic in the same describable sense, but many reactions seem to be contingent. It is not feasible or conducive to the goals to isolate each factor and deal with it separately. You have to think in processes and systems. If you deal with a vector, for example, however simple it may be, you deal with a changing biological system and not with an elementary, undifferentiated point-like entity. The results of biological reactions are usually unclear and need interpretation.

P: I think problems arise from the messy initial state. They don't come from their apparatus. It's just that at this time molecular biologists may not know what they have in their bottle before say some enzyme. But the enzyme does not have to be investigated, it always does the same thing. In science, we always do the same thing . . .

KK: Each science always thinks the other has more control. Molecular biologists often make comments about physicists in this spirit. They will also say that any attempt to investigate their procedures in your way is far too time-consuming and probably doomed to failure in a situation in which they know little about their systems, and all investigations are subject to the same uncertainty as the one that gives rise to the investigation (see Section 4.4). They optimize the system in some holistic way.

P: We cannot do that. We couldn't optimize without knowing. I don't know enough about their work to explain why there is a difference. But it is clear to me that *we* cannot do our work in any other way. If we did not use *these* methods, we would not generate any results. They can obviously do without this obsession with errors, for example. While we could not possibly. There is no cultural choice for us in this.

KK: In a cultural analysis, one explains, in Churchland's sense of explanation (1992: 198), by recognizing a pattern. You may well have set up your problem in a way which demands certain kinds of treatment. But your problem definition is itself part of the pattern, isn't it?

P: Well, the terms *acceptance* and *efficiency* flow a bit into each other. If you want to differentiate them I would rather say that *acceptance* usually refers to geometrical terms, for example, to pseudo-rapidity cuts or transverse momentum cuts. Whereas *efficiency* refers to the less clear cuts we make, like the electron quality cuts. The pseudo-rapidity is a very clear cut. Either the particle is inside or outside this cut. This defines your acceptance in a very clear way. The same holds for the transverse momentum cut. Either it is above 20 GeV or below 20 GeV. Whereas for efficiencies you almost always have a sliding scale where at some point 50 percent of your events pass the cut and at some point 90 percent. So you get the types of efficiency *curves* you know. The definition you give here is probably a better definition of the efficiency than of the acceptance. But you could in some instances define acceptance that way as well.

KK: Aren't there two notions of acceptance? The overall acceptance and the geometrical acceptance?

P: The geometrical acceptance is an efficiency, if you like. Here I think it is rather common that we talk about it in terms of acceptance when it is geometrical—or kinematical. If you look at the very basics, you would probably define both notions (efficiency and acceptance) in the same way which is this way. But . . .

KK: It did pose a lot of problems trying to work out from documents, comments and interviews what you mean, and how to separate the two notions. It did not seem very clear!

P: Perhaps you clarified things too much. They do slide into each other.

6. P: Here you actually quoted me in the text, thank you very much!
7. P: There is actually an important difference between errors and uncertainties. We have uncertainties when we need some input from theory and that input is not well defined. For example, you may have a range of possible (theoretical) inputs and this range of inputs propagates through your analysis and gives a range of possible answers—and this is not really an error, it is an uncertainty. But if you make systematic errors in your measurement, it is just you—you, yourself have an error in your measurement which you do not quite control. *That* is an error. However, in your example it is true if you say here that these systematic errors are also called uncertainties, because we are sloppy when we talk about these things.

KK: So uncertainties refer to an additional distinction you make?

P: We usually lump uncertainties together with systematic errors because traditionally the statistical and systematic errors separate the ones which would be improved by better statistics from the ones which would not. That is why the uncertainties end up with the systematic error, but in reality there are more categories than just the two.

8. P: If you come from the theoretical side and generate top quarks (through Monte Carlo generators) then the mass of the top quark reflects the number of events you would produce. But that is not what we measure. We measure that there are zero top quarks and then we lead that back to a mass limit. The limit we set is the limit on the number of events.
9. P: The limits we produce also affect theory quite a lot.

KK: In what sense?

P: Take LEP, for example. LEP set the number of neutrinos to three. That immediately rules out everything which has more than three neutrinos. Every theorist who has been working on more than three neutrinos can stop doing so and do something else.

KK: You are confident that such limits are not contested by theorists?

P: They need our input. Experimental results overrule theoretical assumptions.

KK: In fact, some theorists regularly attend experimental meetings. Someone (an experimentalist!) once told me they flock to them like vultures looking for prey. The prey was supposedly experimental results which indicate to theorists ways to go and channels to avoid, they provide empirical parameters which enter calculations and above all food—food for new theoretical work. Not surpisingly, with the cancellation of the SSC in Texas it has been theorists who cried out because of the scarcity of input, the "lack of food" if you want to stay in the metaphor, they anticipate if only one collider, CERN, will produce "prey" in the future (Horgan 1994).

R: If you allow me to jump in once more, are you suggesting a reversal of the theory-experiment relation we usually assume, where the theory makes the predictions and the experiment tests them?

KK: Not just a reversal, but, a variety of different relations. Among them the one where the theorist needs experiments to continue in interesting and intelligent ways. Or the one, mentioned in Chapter 1, where the theorist (a phenomenologist) is thought of as a kind of calculation machine for experimental needs. But this is not a matter for this chapter . . .

4 MOLECULAR BIOLOGY AND BLIND VARIATION

1. KLAUS AMANN: Since we collaborated on this chapter, I want to come in as your co-worker and add a few things. When you talk about the laboratory