The New Empiricism: Systematic Musicology in a Postmodern Age[1]

DAVID HURON
Ohio State University

ABSTRACT: A survey of intellectual currents in the philosophy of knowledge and research methodology is given. This survey provides the backdrop for taking stock of the methodological differences that have arisen between disciplines, such as the methods commonly used in science, history or literary theory. Postmodernism and scientific empiricism are described and portrayed as two sides of the same coin we call skepticism. It is proposed that the choice of methodological approach for any given research program is guided by moral and esthetic considerations. Careful assessment of these risks may suggest choosing an unorthodox method, such as quantitative methods in history, or deconstruction in science. It is argued that methodological tools (such as Ockham's razor) should not be mistaken for philosophical world-views. The article advocates a broadening of methodological education in both arts and sciences disciplines. In particular, it advocates and defends the use of quantitative empirical methodology in various areas of music scholarship.

KEYWORDS: *methodology, empiricism, postmodernism, musicology*

INTRODUCTION

SCHOLARLY disciplines distinguish themselves from one another, principally by their subject matter. Musicology differs from chemistry, and chemistry differs from political science because each of these disciplines investigates different phenomena. Apart from the subject of study, scholarly disciplines also frequently differ in how they approach research. The methods of the historian, the scientist, and the literary scholar often differ dramatically. Moreover, even within scholarly disciplines, significant methodological differences are common.

Over the past two decades, music scholarship has been influenced by at least two notable methodological movements. One of these is the so-called "new musicology." The new musicology is loosely guided by a recognition of the limits of human understanding, an awareness of the social milieu in which scholarship is pursued, and a realization of the political arena in which the fruits of scholarship are used and abused. The influence of the new musicology is evident primarily in recent historical musicology and ethnomusicology, but it has proved broadly influential in all areas of music scholarship, including music education.

Simultaneously, the past two decades have witnessed a rise in scientifically inspired music research. This increase in empirical scholarship is apparent in the founding of several journals, including *Psychomusicology* (founded 1981), *Empirical Studies in the Arts* (1982), *Music Perception* (1983), *Musicae Scientiae* (1997), and *Systematic Musicology* (1998). This new empirical enthusiasm is especially evident in the psychology of music and in the resurrection of systematic musicology. But empiricism is also influential in certain areas of music education and in performance research. Music researchers engaged in empirical work appear to be motivated by an interest in certain forms of rigor, and a belief in the possibility of establishing positive, useful musical knowledge.

The contrast between the new musicology and the new empiricism could hardly be more stark. While the new musicology is not merely a branch of Postmodernism, the influence of Postmodern thinking is clearly evident. Similarly, while recent music empiricism is not merely the offspring of Positivism, the family resemblance is unmistakable. Yet the preeminent intellectual quarrel of our time is precisely that between Positivism and Postmodernism – two scholarly approaches that are widely regarded as mortal enemies.[2] How have these diametrically opposed methodologies arisen, and what is a thoughtful scholar to learn from the contrast? How indeed, ought one to conduct music research?

By *methodology*, I mean any formal or semi-formal approach to acquiring insight or knowledge. A methodology may consist of a set of fixed rules or injunctions, or it may consist of casual guidelines, suggestions or heuristics. From time to time, a particular methodology emerges that is shared in common by several disciplines. One example is the so-called Neyman-Pearson paradigm for inductive empirical research commonly used in the physical sciences (Neyman and Pearson, 1928, 1967). But not all disciplines adopt the same methodologies, nor should they.

Different research goals, different fears, different opportunities, and different dispositions can influence the adoption and development of research methods. For any given scholarly pursuit, some research methods will prove to be better suited than others. Part of the scholar's responsibility then, is to identify and refine methods that are appropriate to her or his field of study. This responsibility includes recognizing when a popular research method ceases to be appropriate, and adapting one's research to take advantage of new insights concerning the conduct of research as these insights become known.

Two Cultures

Historically, the most pronounced methodological differences can be observed in the broad contrast between the sciences and the humanities. (For convenience, in this article I will use the term "humanities" to refer to both the humanities and the arts.) In humanities scholarship, research methods include historiographic, semiotic, deconstructive, feminist, hermeneutical, and many other methods. In the sciences, the principal scholarly approaches include modeling and simulation, analysis-by-synthesis, correlational and experimental approaches.

Many scholars presume that methodological differences reflect basic philosophical disagreements concerning the nature of scholarly research. I think this view masks the more fundamental causes of methodological divergence. As I will argue in this article, in most cases, the main methodological differences between disciplines can be traced to the materials and circumstances of the particular field of study. That is, differences in research methods typically reflect concrete differences between fields (or sub-fields) rather than reflecting some underlying difference in philosophical outlook. This is the reason, I will contend, why Muslims and Christians, atheists and anarchists, liberals and libertarians, have little difficulty working with each other in most disciplines. Although deep personal beliefs may motivate an individual to work on particular problems, one's core philosophical beliefs often have little to do with one's scholarly approach.

Philosophy of Knowledge and Research Methodology

In addressing issues pertaining to scholarly methodology, there is merit in dividing the discussion into two related topics. One topic relates to broad epistemological issues, while the second topic relates to the concrete issues of how one goes about doing practical scholarship. In short, we might usefully distinguish philosophy of knowledge (on the one hand) from research methodology (on the other). One rightly expects that the positions we hold regarding the philosophy of knowledge would inform and shape the concrete procedures we use in our day-to-day research methods. However, the information flows in both directions. Practical research experiences also provide important lessons that shape our philosophies of knowledge.

In the training of new scholars, it appears that academic disciplines often differ in the relative weight given to philosophy of knowledge compared with research methodology. My experience with psychologists, for example, is that they typically receive an excellent training in the practical nuts and bolts of research methodology. In conducting research, there are innumerable pitfalls to be avoided, such as *confirmation bias, demand characteristics*, and *multiple tests*. These are the sorts of things experimental psychologists learn to recognize, and devise strategies to avoid or minimize. However, most psychologists I have encountered have received comparatively less training in the philosophy of knowledge. Most have only heard of Hume and Popper, van Quine and Lakatos, Gellner, Laudan, and others. The contrast with the training of literary scholars is striking. There is hardly an English scholar, trained in recent decades, who has not read a number of books pertaining to the philosophy of knowledge. The list of authors differs, however – emphasizing the anti-foundationalist writers: Kuhn and Feyerabend, Derrida and Foucault, Lacan, Leotard, and others.[3] On the other hand, most English scholars receive relatively little training in research methodology, and this is often evident in the confusion experienced by young scholars when they embark on their own research: they often don't know how to begin or what to do.

The philosophical and methodological differences between the sciences and the humanities can be the cause of considerable discomfort for those of us working in the gap between them. As a cognitive musicologist, I must constantly ask whether I should study the musical mind as a humanities scholar, or as a scientist? Having given some thought to methodological questions, my purpose in this article is to share some observations about these convoluted yet essential issues.

OVERVIEW

My goal in this article is to take stock of the methodological differences that arise between disciplines and to attempt to understand their origins and circumstantial merits. As I've already noted, I think the concrete circumstances of research are especially formative. However, before I argue this case, it behooves me to address the noisy (and certainly interesting) debates in the philosophy of knowledge. In particular, it is appropriate to address the often acrimonious debate between empiricism and postmodernism.

Of course not all sciences are empirical and not all humanities scholarship is postmodern. The field of mathematics (which is popularly often considered "scientific") relies almost exclusively on deductive methods rather than empirical methods. Similarly, although postmodernism has been a dominant paradigm in many humanities disciplines over the past two decades, there exist other methodological traditions in humanities scholarship. The reason why I propose to focus on the empirical and postmodernist traditions is that they are seemingly the most irreconcilable. I believe we have the most to learn by examining this debate.

This paper is divided into two parts. In Part I, I outline some of the intellectual history that forms the background for contemporary empiricism and postmodernism. Part II focuses more specifically on methodology. In particular, I identify what I think are the principal causes that lead to the adoption of different methodologies in different fields and sub-fields. Part II also provides historical examples where disciplines have dramatically changed their methodological preferences in response to new circumstances. My claim is that the resources available for music scholarship are rapidly evolving, and that musicology has much to gain by adapting empirical methods to many musical problems. I conclude by outlining some of the basic ideas underlying what might be called the "new empiricism."

PART ONE: PHILOSOPHY OF KNOWLEDGE

Empiricism and Science

The dictionary definition of "empirical" is surprisingly innocuous for those of us arts students who were taught to use it as a term of derision. Empirical knowledge simply means knowledge gained through observation. Science is only one example of an empirical approach to knowledge. In fact, many of the things traditional historical musicologists do are empirical: deciphering manuscripts, studying scores, and listening to performances.

The philosophical complexity begins when one asks how it is that we learn from observation. The classic response is that we learn through a process dubbed *induction*. Induction entails making a set of specific observations, and then forming a general principal from these observations. For example, having stubbed my toe on many occasions over the course of my life, I have formed a general conviction that rapid movement of my toe into heavy objects is likely to evoke pain. We might say that I have learned from experience (although my continued toe-stubbings make me question how well I've learned this lesson).

The 18th-century Scottish philosopher, David Hume, recognized that there are serious difficulties with the concept of induction. Hume noted that no amount of observation could ever resolve the truth of some general statement. For example, no matter how many white swans one observes, an observer would never be justified in concluding that *all swans are white*. Using postmodernist language, we would say that one cannot legitimately raise local observations to the status of global truths.

Several serious attempts have been made by philosophers to resolve the problem of induction. Three of these attempts have been influential in scientific circles: *falsificationism*, *conventionalism* and *instrumentalism*. However these attempts suffer from serious problems of their own. In all three philosophies, the validity of empirical knowledge is preserved by forfeiting any strong claim to absolute truth.

One of the most influential epistemologies in twentieth-century empiricism was the philosophy of conventionalism. The classic statement is found in Pierre Duhem's *The Aim and Structure of Physical Theory* originally published in 1905, but reprinted innumerable times throughout the past century. In his book, Duhem notes that science never provides theories or explanations of some ultimate reality. Theoretical entities and mathematical laws are merely conventions that summarize certain types of relationships. It can never be determined whether scientific theories are "true" in the sense of explaining or capturing some underlying reality. Scientific theories are merely conventions that help scientists organize the observable patterns of the world.

A variation of conventionalism, known as *instrumentalism* similarly posits that empiricism does not provide ultimate explanations: the engineer has no deep understanding of why a bridge does not fall down. Rather, the engineer relies on theories as tools that are reasonably predictive of practical outcomes. For the instrumentalist, theories are judged, not by their "truthfulness," but by their predictive utility.

The most well-known attempt to resolve the problem of induction was formulated by Karl Popper in 1934. Popper accepted that no amount of observation could ever verify that a particular proposition is true. That is, an observer cannot prove that *all swans are white*. However, Popper argued that one could be certain of falsity. For example, observing a single black swan would allow one to conclude that the claim – *all swans are white* – is false. Accordingly, Popper endeavored to explain the growth of knowledge as arising by trimming the tree of possible hypotheses using the pruning shears of falsification. Truth is what remains after the falsehoods have been trimmed away.

Popper's approach was criticized by van Quine, Lakatos, Agassi, Feyerabend and others. One problem is that it is not exactly clear what is falsified by a falsifying observation. It may be that the observation itself is incorrect, or the manner by which the phenomenon of interest is defined, or the overall theoretical framework within which a specific hypothesis is posited. (For example, the observer of a purported black swan might have been drunk, or the swan might have been painted, or the animal might be claimed to be a different species.) A related problem is fairly technical, and so difficult to describe succinctly. In order to avoid prematurely jettisoning a theory, Popper abandoned the notion of a *falsifying observation* and replaced it with the concept of a *falsifying phenomenon*. Yet to establish a falsifying phenomenon, researchers must engage in an activity of verification – an activity which Popper himself argued was impossible. In Popper's methodology, the nasty problem of inductive truth returns through the rear door.

Despite such difficulties, Popper's falsificationism has remained highly influential in the day-to-day practice of empirical research. In the professional journals of science, editors regular remove claims that *such-and-such is true*, or that *such-and-such a theory is verified,* or even that *the data "support" such-and-such a hypothesis*. On the contrary, the boiler-plate language for scientific claims is: *the null hypothesis was rejected* or *the data are consistent with such-and-such a hypothesis*. Of course this circumspect language is abandoned in secondary and popular scientific writings, as well as in the informal conversations of scientists. This gap between official skepticism and colloquial certainty is a proper subject of study for sociologists of science.

Another, less influential scientific epistemology in the twentieth century was positivism. Positivism never provided a proposal for resolving the problem of induction. Nevertheless, it is worth brief mention here for two reasons. First logical positivism drew attention to the issue of language and meaning in scientific discourse, and secondly, "positivism" has been the preeminent target of postmodernist critiques.

Positivism began as a social philosophy in France, initiated by Saint-Simon and Comte, and spread to influence the sciences in the early twentieth century. The tenants of positivism were articulated by the so-called Vienna Circle (including Schlick and Carnap) and culminated in the classic statement of 1936 by A.J. Ayer. In science, logical positivism held sway from roughly 1930 to 1965. However, this influence was almost exclusively restricted to American psychology; only a small minority of empiricists ever considered themselves positivists.

For most of the twentieth century, the preeminent philosophical position of practicing scientists (at least those scientists who have cared to comment on such matters) has been conventionalism or instrumentalism. Popper's emphasis on falsifying hypotheses (which is consistent with both conventionalism and instrumentalism) has proved highly influential in the day-to-day practice of science, largely because of the Pearson/Neyman/Popper statistically-based method of inductive falsification. (Many epistemologists consider Popper's most important and influential writings to be his appendices on probability and statistics.)

This is by no means a complete story of the philosophy of science in the twentieth century, but before we continue our story, it is appropriate to turn our attention to postmodernism.

Postmodernism

Postmodernism is many things, and any attempt to summarize it is in danger of oversimplification. (Indeed, one of the principal tenants of postmodernism is that one should not attempt to represent the world-views of others.) In the same way that philosophers of science disagree with one another, those who call themselves postmodernists also are not of one mind. Nevertheless, there are a number of common themes that tend to recur in postmodernist writings. Postmodernism is a philosophical movement that focuses on how meanings get constructed, and how power is commandeered and exercised through language, representation and discourse.[4]

Postmodernism is interested in scholarship, because scholarly endeavors are among the preeminent meaning-conferring activities in our society. Postmodernism is especially interested in science, principally because, at least in Western societies, science holds a power of persuasion second to no other institution. It is a power, of which the most powerful politicians can only express envy.

Postmodernism begins from a position surprisingly similar to Popper's anti-verification stance and Duhem's conventionalism. Where Duhem and Popper thought that the truth is unknowable, postmodernism assumes that there is no absolute truth to be known. More precisely, "truth" ought to be understood as a social construction that relates to a local or partial perspective on the world. Our mistake is to assume that as observers, we can climb out of the box which is our world. There is no such objective perspective.

There are, rather, a vast number of interpretations about the world. In this, the world is akin to a series of texts. As illustrated in the writings of Jacques Derrida, any text can be deconstructed to reveal multiple interpretations, no one of which can be construed as complete, definitive, or privileged. From this, postmodernists conclude that there is no objective truth, and similarly that there is no rational basis for moral, esthetic or epistemological judgment.

If there is no absolute basis for these judgments, how do people in the world go about making the decisions they do? The most successful achievements of postmodernism have been in drawing attention to the power relations that exist in any situation where an individual makes some claim. As Nancy Hartsock has suggested, "the will to power [is] inherent in the effort to create theory" (1990; p.164). Like the politician or the business person, scholars are consciously or unconsciously motivated by the desire to commandeer resources and establish influence. Unlike the politician or the business person, we scholars purport to have no hidden agenda – a self-deception that makes us the most dangerous of all story-tellers.

It is the most powerful members of society who are able to establish and project their own stories as so-called "master narratives." These narratives relate not only to claims of truth, but also to moral and artistic claims. The "canons" of art and knowledge are those works exalted by, and serving, the social elites. Insofar as works of art give legitimacy to those who produce them, "A work of art is an act of power." (Rahn, 1993)

This admittedly pessimistic view of the world could well lead one to despair. Since there is no legitimate power, how does the conscientious person act so as to construct a better world? Postmodernism offers various strategies that might be regarded as serving the goal of *exposé*. That is, the postmodernist helps the cause through a sort of investigative journalism that exposes how behaviors are self-serving. At its best, postmodernism is a democratizing ladle that stirs up the political soup and resists the entrenchment of a single power. By creating a sort of chaos of meaning, it calls existing canons into question, subverts master narratives, and so gives flower to what has been called "the politics of difference".

FEYERABEND AND THE GALILEO-SCHOLASTICS DEBATE

In the world of the sciences, a concrete demonstration of such power relations is examined in the work of Paul Feyerabend. In his book, *Against Method*. Feyerabend used scientific method itself to show the failures of scientific discourse, and the role of power in presumed rational debate.

It is worth discussing Feyerabend's work at some length because his work has led to widespread misconceptions, many of which were promoted by Feyerabend himself.

Contemporary scientific method embraces certain standards for evidence in scientific debates. For example, when two competing theories (*X* and *Y*) exist, scientists attempt to construct a "critical experiment" where the two theories are pitted against each other. If the results turn out one way, theory *X* is

rejected; if the results turn out another way, theory *Y* is rejected. In addition, contemporary scientific method frowns upon so-called ad hoc hypotheses. Suppose that the results of a critical experiment go against my pet theory. I might try to save my theory by proposing that the experiment was flawed in various ways. I might say that the reason the experiment failed to be consistent with my theory is that the planet Mercury was in retrograde on the day that the experiment was carried out, or that my theory is true except on the third Wednesday of each month. Of course *ad hoc* hypotheses need not be so fanciful. More credible *ad hoc* hypotheses might claim that the observer was poorly trained, the equipment not properly calibrated, or the control group improperly constructed, etc. Although an *ad hoc* hypothesis might be true, such appeals are considered very bad form in scientific circles whenever the motivation for such claims is patently to "explain away" a theoretical failure.

Feyerabend uses the case study of the famous debate between Galileo and the Scholastics. In the popular understanding of this history, Galileo argued that the sun was positioned in the center of the solar system and the Scholastics, motivated by religious dogma, maintained that the earth was in the center of the universe.

Historically, this popular view is not quite right – as Feyerabend points out. The Scholastics argued that motion is relative, and that there is, in principle, no way that one could determine whether the earth was rotating about the sun or the sun was rotating about the earth. Since observation alone cannot resolve this question, the Scholastics argued that the Bible implies that the earth would be expected to hold a central position.

However, Galileo and the Scholastics agreed on a possible critical experiment. Suppose that your head represents the earth. If you rotate your head in a fixed position, the angles between various objects in the room will remain fixed. However, if you walk in a circle around the room, the visual angles between various objects will change. As you approach two objects, the angle separating them will increase. Conversely, as you move away from two objects, the angle separating them will decrease.

According to this logic, if the earth is in motion, then one ought to be able to see slight angular shifts between the stars over the course of the year. Using his new-fangled invention, the telescope, Galileo did indeed make careful measurements of the angular relationships between the stars over the course of a year. He found, however, that there was no change whatsoever. In effect, Galileo carried out a critical experiment – one whose results were not consistent with the idea that the earth is in motion. How did Galileo respond to this result? Galileo suggested that the reason why no parallax shifts could be observed was because the stars are extremely far away.

Feyerabend pointed out that this is an *ad hoc* hypothesis. A critical experiment was carried out to determine whether the earth or the sun was in motion, and Galileo's theory lost. Moreover, Galileo had the audacity to defend his theory by offering an *ad hoc* hypothesis. By modern scientific standards, one would have to conclude that the Scholastics' theory was superior, and that, as a scientist, Galileo himself should have recognized that the evidence was more consistent with the earth-centered theory.

Of course, from our modern perspective, Galileo was right to persevere with his sun-centered theory of the solar system. As it turns out, his *ad hoc* hypothesis regarding the extreme distance to the stars is considered by astronomers to be correct.

From this history, Feyerabend draws the following conclusions. First, the progress of science may depend on bad argument and ignoring data. Second, Galileo should be recognized, not as a great scientist, but as a successful propagandist. Third, had Galileo followed modern standards of scientific method the result would have been scientifically wrong. Fourth, the injunction against *ad hoc* hypotheses in science can produce scientifically incorrect results. Fifth, the use of critical experiments in science can produce scientifically incorrect results. Sixth, no methodological rule will ensure a correct result. Seventh, there is no scientific method. And eighth, in matters of methodology, concludes Feyerabend, *anything goes*. Like Popper and Lakatos, Feyerabend argued that there is no set of rules that guarantees the progress of knowledge.

In assessing Feyerabend's work, we need to look at both his successes and failures. Let's begin with some problems. Recall that the problem of induction is the problem of how general conclusions can be drawn from a finite set of observations. Consider, the fourth and fifth of Feyerabend's conclusions. He notes that two rules in scientific methodology (namely, the rule forbidding *ad hoc* hypotheses, and the instruction to devise critical experiments) failed to produce a valid result in Galileo's case. From these two historical observations, Feyerabend formulates the general conclusion: no methodological rule will ensure a correct result. By now you should recognize that this is an inductive argument, and as Hume pointed out, we can't ever be sure that generalizing from specific observations produces a valid generalization.

Showing that *some* methodological rules don't work in a single case, doesn't allow us to claim that *all* methodological rules are wrong. Even if one were to show that all known methodological rules were inadequate, one can't logically conclude than there are *no* true methodological rules.

A further problem with Feyerabend's argument is that he exaggerates Galileo's importance in the promotion of the sun-centered theory. The beliefs and arguments of a single person are typically limited. Knowledge is socially distributed, and ideas catch on, only when the wider population is prepared to be convinced. In fact, the heliocentric theory of the solar system was not immediately adopted by scientists because of Galileo's arguments. The heliocentric theory didn't gain many converts until after Kepler showed that the planets move in elliptical orbits. Kepler's laws made the sun-centered theory a much simpler system for describing planetary motions. In short, Galileo's fame and importance as a scientific champion is primarily retrospective and ahistorical.

Feyerabend's historical and analytic work is insufficient to support his general conclusion: namely that in methodology, the only correct rule is "anything goes." Moreover, Feyerabend's own dictum is not born out by observation. Anyone observing any meeting of any academic group will understand that, in their debates, it is not true that `anything goes.' All disciplines have more or less loose standards of evidence, of sound argument, and so on. Although a handful of scholars might wish that debates could be settled through physical combat, for the majority of scholars such "methods" are no longer admissible. There may be no methodological recipe that guarantees the advance of knowledge, but similarly, it is not the case that anything goes.

On the positive side, Feyerabend has drawn attention to the social and political environment in which science takes place. Feyerabend stated that his main reason for writing *Against Method* was "humanitarian, not intellectual". Feyerabend wanted to provide rhetorical support for the marginalized and dispossessed (p.4). In drawing attention to the sociology of science, Feyerabend and his followers have met strong resistance from scientists themselves. Until recently, most scientists rejected the notion that science is shaped by a socio-political context. The failings of science notwithstanding, this does not mean that scholars working in the sociology of science have been doing a good job.

KUHN AND PARADIGMATIC

The most influential study of science is probably Thomas Kuhn's *The Structure of Scientific Revolutions*. As a historian of science, Kuhn set out to describe how new ideas gain acceptance in a scientific community.

From his studies in the history of science Kuhn distinguished two types of science: *normal* science and revolutionary science. The majority of scientific research can be described as normal science. Normal science is a sort of puzzle-solving activity, where the prevailing scientific theory is applied in various tasks, and small anomalies in the prevailing theory are investigated. Many anomalies are resolved by practicing such "normal" science. However, over time, certain anomalies fail to be resolved and a minority of scientists begin to believe that the prevailing scientific theory (or "paradigm") is fundamentally flawed.

Revolutionary science breaks with the established paradigm. It posits an alternative interpretation that meets with stiff resistance. Although the new theory might explain anomalies in the prevailing theory, inevitably, there are may things that are not (yet) accounted for by the new theory. Opponents of the new paradigm contrast these failures with the known successes of the existing paradigm. (In part, the problems with the new paradigm can be attributed to the fact that the new theory has not yet benefitted from years of normal science that resolve apparent problems that can be explained using the old paradigm.)

An important claim made by Kuhn is that debates between supporters of the old and new paradigms are not rational debates. Changing paradigms is akin to a religious conversion: one either sees the world according to the old paradigm or according to the new paradigm. Supporters of the competing paradigms are incapable of engaging each other in reasoned discussion. Scientists from competing paradigms "talk past each other." Technical terms, such as "electron" begin to have different meanings for scientists supporting different paradigms.

Kuhn argued that there is no neutral or objective position from which one can judge the relative merits of the two different paradigms. Consequently, Kuhn characterized the paradigms as *incommensurable* – not measurable using a single yard-stick. Paradigm shifts occur, not because supporters of the old paradigm become convinced by the new paradigm. Instead, argues Kuhn, new paradigms replace old paradigms because old scientists die, and new paradigm supporters are able to place their colleagues and students in important positions of power (professorships, journal editors, granting agencies, etc.) Once

advocates of the new paradigm have seized power, the textbooks in the discipline are re-written so that the revolutionary change is re-cast as a natural and inevitable step in the continuing smooth progress of the discipline.

While Kuhn's work had an enormous impact in the social sciences, it had comparatively little impact in the sciences themselves. *The Structure of Scientific Revolutions* portrayed science as akin to fashion: changes do not arise from some sort of rational debate. Change is simply determined by who holds power. Although Thomas Kuhn denied that he was arguing that science does not progress, his study of the history of science strongly implies that "scientific progress" is an illusion perpetrated by scientists who re-construct history to place themselves (and their paradigms) at the pinnacle of a long lineage of achievement.

Many social sciences and humanities scholars applauded Kuhn because his portrayal removed science from the epistemological high ground. The presumed authority of science is unwarranted. Like different cultures around the world, there is no valid yardstick by which one can claim that one scientific culture is better than another.

Kuhn's writings also appealed to those scientists (and other scholars) whose views place them outside the mainstream. For those scientists whose unorthodox views are routinely ignored by their colleagues, Kuhn's message is highly reassuring. The reason why other people don't understand us and don't care about what we say, is that they are enmeshed in the old paradigm: no amount of reasoned debate can be expected to convince the existing powers. In short, Kuhn's characterization of science provides a measure of comfort to the marginalized and dispossessed.

Shortly after the publication of Kuhn's book, a young Bengali philosopher named Jagdish Hattiangadi wrote a detailed critique of the work. Although Kuhn regarded himself as a historian of science with great sympathies for science, Hattiangadi noted that Kuhn's work removed any possibility that science could be viewed as a rational enterprise. Although Kuhn never said as much, his theory had significant repercussions: for example, a chemist who believes that modern chemistry is *better* than ancient chemistry must simply be deluded. Hattiangadi noted that, either there is no progress whatsoever in science, or Kuhn's portrayal of science is wrong. Hattiangadi concluded that Kuhn's work failed to account for the widespread belief that scientific progress is a fact. Moreover, as early as 1963, Hattiangadi predicted that Kuhn's book would become wildly successful among social and humanities scholars – a prediction that proved correct.

POSTMODERNISM: AN ASSESSMENT

With this background in place, let's return to our discussion of postmodernism. In general, postmodernism takes issue with the Enlightenment project of deriving absolute or universal truths from particular knowledge. That is, postmodernism posits a radical opposition to induction. We cannot generalize from the particular; the global does not follow from the local.

At first glance, it would appear that postmodernism would be as critical of Feyerabend and Kuhn as of the positivists. For the arguments of Feyerabend and Kuhn also rest on the assumption that we can learn general lessons from specific historical examples. However, postmodernism is less concerned with such convoluted issues than it is with the general goal of causing intellectual havoc for those who want to make strong knowledge claims. Accordingly, the works of Feyerabend and Kuhn are regarded as allies in the task of unraveling science's presumed authority.

Of course postmodernism also has its critics. Much of the recent unhappiness with postmodernism is that it appears to deny the possibility for meaningful human change. For example, many feminist thinkers have dismissed a postmodernist approach because it removes the high moral ground. In lobbying for political change, most feminists have been motivated by a sense of injustice. However, if there are no absolute precepts of justice, then the message postmodernism gives to feminists is that they are simply engaged in Machiavellian maneuvers to wrest power. In the words of Joseph Natoli, "postmodernist politics here has nothing to do with substance but only with the tactics." (1997, p. 101) On the one hand, postmodernism encourages feminists to wrest power away from the male establishment; but at the same time, postmodernism tells feminists not to believe that their actions are at all justified. Understandably, many feminists are uncomfortable with this contradiction.

The nub of the issue, I think, is evident in the following two propositions associated with postmodernism:

(1) There is no privileged interpretation.
(2) All interpretations are equally valid.

As the postmodernist writer Catherine Belsey has noted, postmodernism has been badly received by the public primarily because postmodernists have failed to distinguish between sense and nonsense. This is the logical outcome for those who believe that (2) is simply a restatement of (1).

If we accept the proposition that there is no privileged interpretation, it does not necessarily follow that all interpretations are equally valid. For those who accept (1) but not (2), it follows that some interpretations must be "better" than others – hence raising the question of what is meant by "better."

Postmodernism has served an important role by encouraging scholars to think carefully, laterally, and self-reflectively. Unfortunately, postmodernism encourages slovenly research and a disinterest in pursuing rigor. Postmodernism draws welcome attention to the social and political context of knowledge and knowledge claims. But postmodernism goes too far when it concludes that reality is socially *constructed* rather than socially *mediated*. Postmodernism serves an important role when it encourages us to think about power relations, and in particular how certain groups are politically disenfranchised because they have little control over how meanings get established. But at the same time, postmodernism subverts all values, and transforms justice into mere tactical maneuvers to gain power. In reducing all relationships to power, postmodernism leaves no room for other human motivations. Scholarship may have political dimensions, but that doesn't mean that all scholars are plotting power-mongers. Postmodernism is important insofar as it draws attention to the symbolic and cultural milieu of human existence. But, while we should recognize that human beings are cultural entities, we must also recognize that humans are also biological entities with *a priori* instinctive and dispositional knowledge about the world that originates in an inductive process of evolutionary adaptation (Plotkin, 1994). Foucault regrettably denied any status for humans as biological entities whose mental hardware exists for the very purpose of gaining knowledge about the world.

When pushed on the issue of relativism, postmodernists will temporarily disown their philosophy and accept the need for some notion of logic and rigor. Belsey, for example, claims that as postmodernists, "we should not abandon the notion of rigor; the project of substantiating our readings" (Belsey, 1993, p. 561) Similarly, Natoli recognizes that "logic" (1997, p.162) and "precision" (p.120) make for compelling narratives. However, postmodernists are oddly uninterested in how these approaches gain their rhetorical power. What is "logic"? What is "rigor"? What is it about rationality that makes some narratives so mentally seductive or compelling? It is exactly this task that has preoccupied philosophers of knowledge over the past 2,500 years and was the focus of Enlightenment efforts in epistemology. The Enlightenment project of attempting to characterize the value of various knowledge claims is not subverted by postmodernism. On the contrary, postmodernism simply raises anew the question of what it means to do good scholarship.

PART TWO: PHILOSOPHY OF METHODOLOGY

How then, should scholars conduct research? What does the philosophy of knowledge tell us about the practicalities of scholarship? As we have seen, the philosophy of knowledge suggests that we abandon the view that methodology is an infallible recipe or algorithm for establishing the truth. The epistemological role of methodology is much more modest. At the same time, what the new empiricism shares in common with postmodernism is the conviction that scholarship occurs in a moral realm, and so methodology ought be guided by moral considerations.

Methodological Differences

As noted in the introduction, one of the principal goals of this paper is to better account for why methodologies differ for different disciplines. In pursuing this goal I will outline a taxonomy of research methodologies based on four distinctions. In brief, these are:

False-positive skepticism versus false-negative skepticism. False-positive skepticism holds that theories or hypotheses ought to be rejected given the slightest contradicting evidence. False-negative skepticism holds that theories or hypotheses ought to be conserved unless there is overwhelming contradicting evidence.

High risk versus low risk theories. Theories, hypotheses, interpretations and intuitions carry moral and esthetic repercussions. In testing some knowledge claim, the burden of evidence can shift depending on the consequences of the theory. Many theories carry negligible risks, however.

Retrospective versus prospective data. Some areas of research (such as manuscript studies) have only pre-existing evidence or data. Other areas of research (such as behavioral studies) have opportunities to collect newly generated evidence. Prospective data allows researchers to more rigorously test knowledge claims by attempting to forecast properties of yet-to-be-collected data.

Data-rich versus data-poor fields. Fields of study can also be characterized according to the volume of pertinent evidence. When the evidence is minimal, researchers in data-rich fields have the luxury of suspending judgment until more evidence is assembled. By contrast, researchers in data poor fields often must interpret a set of data that is both very small and final – with no hope of additional forthcoming evidence.

Below, I will describe more fully these four distinctions. My claim is that fields of study can be usefully characterized by these taxonomic categories. Each of these four distinctions has repercussions for formulating field-appropriate methodologies. I will suggest that these taxonomic distinctions not only help us to better understand why methodologies diverge for various fields, but also help us to better recognize when an existing methodology is inappropriate for some area of study.

Additionally, I will note that fields of research sometimes experience major changes in their basic working conditions – changes that precipitate shifts in methodology. A formerly uncontentious field of research (such as education) may abruptly find that its latest theories[5] carry high moral risk. A previously data-poor field (such as theology) may become inundated by new sources of information. And a formerly retrospective discipline (such as history) may unexpectedly find a class of events for which it can offer testable predictions. Later in this article I will briefly discuss two case examples of such shifts in resources and methods. My first example is the transformation of sub-atomic physics so that its methods increasingly resemble those in philosophy and literary theory. My second example will be the increasing influence of empirical methods in music scholarship.

Two Forms of Skepticism

From at least the time of the ancient Greeks, the essence of scholarship has been closely associated with skepticism. Most scholars evince a sort of love/hate relationship with skepticism. On the one hand, we have all experienced annoyance at the credulity of those who accept uncritically what we feel ought to evoke wariness. On the other hand, we have all experienced exasperation when someone offers belligerent resistance to the seemingly obvious. What one person regards as prudent reserve, another considers bloody-mindedness.

Science is often portrayed as an institutionalized form of skepticism. Unfortunately, this portrayal can leave the false impression that the arts and humanities are *not* motivated by skepticism – that the humanities are somehow credulous, doctrinaire, or gullible. Contrary to the views of some, most humanities disciplines also cultivate institutionalized forms of skepticism; however, the type of skepticism embraced is often diametrically opposed to what is common in the sciences.

These differences are illustrated in Table 1. The table identifies four epistemological states related to any knowledge claim (including the claim that something is unknowable). Whenever a claim, assertion, or mere insinuation is made, two types of errors are possible. A *false positive error* occurs when we claim something to be true or useful or knowable when it is, in fact, false, useless or unknowable. A *false negative error* occurs when we claim something to be false/useless/unknowable when it is, in fact, true/useful/knowable. Methodologists refer to these errors as Type I and Type II respectively.

Table 1

	Thought to be True, Useful or Knowable	***Thought to be False, Useless or Unknowable***
Actually True, Useful or Knowable	***Correct Inference***	***False Negative Error (Type II Error)***
Actually False, Useless or Unknowable	***False Positive Error (Type I Error)***	***Correct Inference***

The *false-positive skeptic* tends to make statements such as the following:

> "You don't know that for sure."
> "I really doubt that that's useful."
> "There's no way you could ever know that."

By contrast, *false-negative skepticism* is evident in statements such as the following:

> "It might well be true."
> "It could yet prove to be useful."
> "We might know more than we think."

In short, the two forms of skepticism might be summarized by the following contrasting assertions:

> *False-Positive Skeptic*: "There is insufficient evidence to support that."
> *False-Negative Skeptic*: "There is insufficient evidence to reject that."

Speaking of false-negative and false-positive skepticism can be a bit confusing. For the remainder of this article, I'll occasionally refer to false-positive skepticism as *theory-discarding skepticism* since these skeptics look for reasons to discard claims, theories or interpretations. By contrast, I'll occasionally refer to false-negative skepticism as *theory-conserving skepticism* since these skeptics are wary of evidence purporting to disprove a theory or dismiss some claim, view, interpretation or intuition.

In the case of the physical and social sciences, most researchers are theory-discarding skeptics. They endeavor to minimize or reduce the likelihood of making false-positive errors. That is, traditional scientists are loath to make the mistake of claiming something to be true that is, in reality, false. Hundreds of thousands of scientific publications begin from the premise of theory-discarding skepticism.[6] This practice has arisen in response to researchers' observations that we are frequently wrong in our intuitions and all too eager to embrace suspect evidence in support of our pet theories.

In the past two decades or so, medical researchers have raised serious challenges to this orthodox scientific position. The U.S. Food and Drug Administration formerly approved only those drugs that had been proved to be effective (i.e., "useful") according to criteria minimizing false-positive errors. (That is, drugs that might be useful were rejected.) The AIDS lobby drew attention to the illogic of denying seemingly promising drugs that had not yet been shown to be useless. For the patient facing imminent

death, it is the enlightened physician who will recommend that her patient seek out the most promising of recent "quacks."[7] In other words, the medical community has drawn attention to the possible detrimental effects of committing false-negative errors. Theory-discarding skeptics are prone to the error of claiming something to be useless that is, in fact, useful.

This shift in attitude has moved contemporary medical research more closely towards dispositions more commonly associated with traditional arts/humanities scholars. Broadly speaking, traditional humanities scholars (including scholars in the arts) have tended to be more fearful of committing false-negative errors. For many arts and humanities scholars, a common fear is dismissing prematurely an interpretation or theory that might have merit – however tentative, tenuous or incomplete the supporting evidence. Arts scholars (in particular) have placed a premium on what is regarded as sensitive observation and intuition: no detail is too small or too insignificant when describing or discussing a work of art.

Another way that traditional humanities scholars exhibit theory-conserving tendencies is evident in attitudes toward the notion of *coincidence*. For traditional scientists, the principal methodological goal is to demonstrate that the recorded observations are unlikely to have arisen by chance. In the common Neyman-Pearson research paradigm, this is accomplished by disconfirming the *null hypothesis.* That is, the researcher makes a statistical calculation showing that the observed data[8] are inconsistent with the hypothesis that the data would be expected to arise by chance. For many traditional humanities scholars, however, dismissing an observation as a "mere coincidence" is problematic. If the goal is to minimize false negative claims, then a single "coincidental" observation should not be dismissed lightly. For many arts and humanities scholars, apparent coincidences are more commonly viewed as "smoking guns."

In summary, both traditional scientists and traditional humanities scholars are motivated by skepticism, but they often appear to be motivated by two different forms of skepticism. One community appears to be wary of accepting theories prematurely; the other community appears to be wary of dismissing theories prematurely.

A concrete repercussion of these two forms of skepticism can be found in divergent attitudes towards the language of scholarly reporting.

Open Accounts versus Closed Explanations

Scientists are apt to take issue with the idea that traditional humanities scholars are more likely to give interesting hypotheses or interpretations the benefit of the doubt. A scientist might well point out that many traditional humanities scholars are often skeptical of scientific hypotheses for which a considerable volume of supporting evidence exists. How, it might be asked, can a humanities scholar give credence to Freud's notion of the Oedipal complex while entertaining doubts about the veracity of Darwin's theory of evolution? I think there are two answers to this question – one answer is substantial, while the second answer arises from an understandable misconception.

The substantial answer has to do with whether a given hypothesis tends to preclude other possible hypotheses. The Oedipal complex might be true without significantly precluding other ideas or theories concerning human nature and human interaction. However, if the theory of evolution is true, then a large number of alternative hypotheses must be discarded. It is not necessarily the case that the humanities scholar holds a double standard when evaluating scientific hypotheses. If a scholar is motivated by theory-conserving skepticism (that is, avoiding false-negative claims), then a distinction must be made between those theories that claim to usurp all others, and those theories that can co-exist with other theories. The theory-conserving skeptic may cogently choose to hold a given hypothesis to a higher standard of evidence precisely because it precludes such a wealth of alternative interpretations.

In the humanities, young scholars are constantly advised to draw conclusions that "open outwards" and to "avoid closure." This advice contrasts starkly with the advice given to young scientists who are taught that "good research distinguishes between competing hypotheses." From the point of view of the false-negative skeptic, a "closed" explanation greatly increases the likelihood of false-negative errors for the myriad of alternative hypotheses.

This fear is particularly warranted whenever the volume of available data is small, as is often the case in humanities disciplines. A low volume of evidence means that no single hypothesis can be expected to triumph over the alternatives, and so claims of explanatory closure in data-poor fields are likely to be unfounded. For this reason, many humanities scholars regard explanatory "closure" as a provocation – a political act intended to usurp all other views.

Of course many scientific theories do indeed achieve a level of evidence that warrants broad acceptance and rejection of the alternative theories. Still, not all humanities scholars will be convinced that the alternative accounts must be rejected. I suspect that all researchers (both humanities scholars and scientists) tend to generalize from their own discipline-specific experiences when responding to work reported from other fields. Since humanities scholars often work in fields where evidence is scanty, the humanities scholar's experience shouts out that no knowledge claim warrants the kind of confidence commonly expressed by scientists. Objecting to scientific theories on this basis is clearly a fallacy, but it is understandable why scholars from data-poor disciplines would tend to respond skeptically to the cocky assurance of others. We will return to consider the issue of explanatory closure later, when we discuss Ockham's razor and the issue of reductionism.

Having proposed this association between theory-discarding skepticism and science (on the one hand) and theory-conserving skepticism and the humanities (on the other hand), let me now retract and refine it. I do not think that there is any *necessary* association. The origin of this tendency, I propose, has nothing to do with the nature of scientific as opposed to humanities scholarship. I should also hasten to add that I do not believe that individual scholars are solely theory-discarding or theory-conserving skeptics. People have pretty good intuitions when to approach a phenomenon as a false-positive skeptic and when to approach a phenomenon as a false-negative skeptic.

If there is no necessary connection between theory-discarding skepticism and science, and theory-conserving skepticism and the humanities, where does this apparent association come from? I think there are two factors that have contributed to these differing methodological dispositions. As already suggested, one factor relates to the quantity of available evidence or data for investigating hypotheses or theories. A second factor pertains to the moral and esthetic repercussions of the hypotheses. These two factors are interrelated so it is difficult to discuss each factor in isolation. Nevertheless, in the ensuing discussion, I will attempt to discuss each issue independently.

HIGH RISK VERSUS LOW RISK THEORIES

For the casual reader, one of the most distinctive features of published scientific research are those strings of funny Greek letters and numbers that often pepper the prose. Some statement is made, such as "X is bigger than Y," and this is followed in parentheses by something like the following:

$$X^2=8.32;\ df=4;\ p<0.02$$

There is some skill involved in understanding these numbers, but the essential message is conveyed by the value of p.

In statistical inference, the value p is a calculated value that estimates the probability of making a false-positive error. If the researcher is endeavoring to avoid making a false positive claim, then the value of p should be as small as possible. As we have seen, depending on the circumstances, the researcher may wish to minimize the possibility of making a false negative error (i.e. theory-conserving skeptic). How does a researcher know what type of error to minimize? Should the researcher be skeptical of negative claims or skeptical of positive claims? Should the researcher aim to conserve theories or discard them?

The answer to this question is that it depends upon the moral (and esthetic) consequences of making one kind of error versus another kind of error. Consider, for example, the difference between civil and criminal cases in jurisprudence. Civil cases (such as trespassing) require comparatively modest evidence in order to secure a conviction ("preponderance of evidence"). Criminal cases (such as murder) require much more convincing evidence ("beyond a reasonable doubt"). These different standards of evidence are warranted due to the different moral repercussions of making a false-positive error. Securing the conviction of an innocent person in a murder trial is a grave blunder compared to convicting an innocent person of trespassing.

Fields of inquiry that carry significant risks (such as medicine, jurisprudence and public safety) ought to have high standards of confidence. If the field is data rich, it is especially important to collect a sufficient volume of evident so the researcher can assemble a convincing case. If the field is data poor (such as often happens in jurisprudence), then one must expect to make a lot of errors; the moral repercussions of a false-positive versus a false-negative error will determine whether the researcher should adopt a theory-conserving or theory-discarding skepticism. In criminal law, one can expect many failures to convict guilty people in order to minimize the number of wrongful convictions.

In contrast with legal proceedings, most scholarly hypotheses have marginal moral or esthetic risk. For example, whether a theory of the origins of Romanesque architecture is true or false has little moral impact. However, risk is never entirely absent. Suppose that a musicologist found evidence suggesting that one composer had plagiarized a melody from another composer. If the claim of plagiarism was in fact false, then the first composer's reputation would be unjustly tarnished. If that composer were still living, then a false claim of plagiarism would be morally reprehensible.

To the knowledgeable statistician there is nothing new in this discussion. Modern statisticians have always understood the reciprocal relationship between false positive and false negative errors, and have long recognized that whether a researcher endeavors to reduce one or the other depends entirely on the attendant risks of making either error. In most traditional arts and humanities scholarship, making a false positive claim rarely has onerous moral or esthetic repercussions. Conversely, false-negative claims have often been seen as reckless.

Perhaps the best-known theory-conserving argument is *Pascal's Wager*. Unconvinced by the many proofs offered for the existence of God, Pascal asked what would be lost if the proposition were true but our evidence scant? Pascal argued that the repercussions of making a false-negative error were simply too onerous. He chose to believe in God, not because the positive evidence was compelling, but because he thought that the moral risk associated with wrongly dismissing the hypothesis would require an extraordinary volume of contradicting evidence (Pascal, 1669).

Historically, statistical tests have been used almost exclusively to minimize false-positive errors. It is the community of theory-discarding skeptics who have made the greatest use of statistics. I suspect that this historical association between the use of statistical inference and false-positive skepticism may account for much of the widespread suspicion of statistical arguments among arts and humanities scholars. Yet there is nothing in statistical inference *per se* that is contrary to the traditional arts/humanities scholar's penchant for false negative skepticism. As statisticians well know, common statistical procedures are equally adept at serving the theory-conserving skeptic.

As noted earlier, the science/false-positive and humanities/false-negative association is changing. Contemporary medicine has become more cognizant of the dangers of prematurely discarding theories. Concurrently, many arts and humanities researchers are becoming more aware of the problems of theory-conserving skepticism. In the case of music, several hundred years of speculative theorizing has led to the promulgation of innumerable ideas – many of which surely lack substance[9]. Until recently, there was little one could do about this. The scarcity of pertinent data in many humanities fields simply made it impossible to satisfy statistical criteria for minimizing false positive errors. The opportunities to address these problems have been immensely expanded due to the growing availability of computer databases, comprehensive reference tools, and the growing use of experiment-based data collection. We will return to these issues shortly.

HISTORICAL FIELDS

Fields can be characterized according to whether the principal evidence or data arise from the past or from the future. Historical fields are fields whose fundamental data already exist. Archeology, paleontology and art history are examples of historical fields. In each of these fields, the principal phenomena of study are ones that occurred in the past. These phenomena are accessible for study only through the tenuous traces of currently existing data. Historical data might include paper documents, physical objects, oral histories, or unspoken memories. Normally, the existing evidence constitutes a proper subset of all of the pertinent evidence, most of which has been destroyed by the passage of time.

It would be wrong to think of historical fields as principally belonging to the humanities. The sciences of astronomy, geology, and paleoanthropology are predominantly historical fields. Each of these sciences is concerned primarily with evidence of past events. Indeed, the preeminent historical discipline, it might be argued, is *astronomy*: the light that reaches astronomers' telescopes is typically hundreds or millions of years old. It is rare that astronomers get to study "current events."

RETROSPECTIVE VERSUS PROSPECTIVE DATA

Historical data should not be confused with what may be called *retrospective* evidence or data. *Retrospective* data is evidence that is already in-hand – evidence that is known to the researcher. Prospective data, by contrast, is data that is not yet available to the researcher. Prospective data includes

evidence that will be collected in the future, but prospective data also includes existing evidence that a researcher has not yet seen – such as data published in a forgotten article, or manuscripts in an overlooked archive.

Note that prospective data can be entirely historical. Consider, by way of example, weather forecasting. We normally think of meteorologists testing their models by forecasting future weather, such as predicting tomorrow's weather, the weather next week, or the weather next year. However, most meteorological theories are tested using historical data. Given the antecedent data, a theory might be used to predict the weather on, say, March 2nd, 1972.

Similarly, suppose that an ethnomusicologist formulates a theory based on a study of three hunter-gatherer societies. For example, the ethnomusicologist might theorize that matrilinear hunter-gatherers employ predominantly ascending melodic contours whereas patrilinear hunter-gatherers exhibit predominantly descending melodic contours. This theory might be tested by predicting specific cultural patterns in other hunter-gatherer groups. We might test the ethnomusicologist's predictions by carrying out new field research in as-yet-unstudied cultures. However, we could also test the ethnomusicologist's predictions against already existing data about other societies, provided the data is prospective rather than retrospective. Similarly, historians might test specific theories by predicting the contents of newly discovered (yet unopened) documents pertaining to a particular historical event.

Of course in some areas of research, all of the pertinent data is already available. No amount of money will necessarily increase the volume of documents relating directly to Petrarch's life. In other words, all of the data is retrospective and researchers hold little hope of future prospective data. The loss of opportunities for prospective data removes the possibility of evaluating a theory by testing predictions. This situation has onerous repercussions for the affected area of research.

PRE-DATA THEORY AND POST-DATA THEORY

One of the most pernicious problems plaguing historical disciplines is the tendency to use a single data set both to generate the theory and to support the theory. Formally, if observation O is used to formulate theory T, then O cannot be construed as a predicted outcome of T. That is, observation O in no way supports T.

The origin of the Theory of Continental Drift arose from observing the suspicious visual fit between the east coasts of the American continents and the west coasts of Europe and Africa. The bulge of north-west Africa appears to fit like a piece of a jig-saw puzzle into the Caribbean gulf. This observation was ridiculed as childish nonsense by geologists in the first part of the twentieth century. Geologists were right to dismiss the similarity of the coast-lines as *evidence* in support of the theory of continental drift, since this similarity was the origin of the theory in the first place. Plate tectonics gained credence only when independent evidence was gathered consistent with the spreading of the Atlantic sea-bed.

Such *post hoc* theorizing has particularly plagued evolutionary theorizing (see Gould, 1978; Gould & Lewontin, 1979; Lewontin, 1991; Rosen, 1982). Nevertheless, in some cases, evolutionary theories can arise that make predictions about yet-to-be-gathered data (such as the Trivers-Willard hypothesis). Good theories are *a priori*; that is, the theory suggests or predicts certain facts or phenomena before those facts are ascertained or observed.

Fields that rely exclusively on retrospective data are susceptible to *post hoc* theorizing where hypotheses are easy to form and difficult to test. This is a problem that is endemic to many fields, especially historical fields (including astronomy). Nevertheless, careful attention to the underlying logic of a theory may permit testing of unexpected predictions of pre-existing prospective data. The fields of astronomy and evolutionary biology have demonstrated that there are many more opportunities for testing historical theories than is recognized by historians working in humanities disciplines.

EXPERIMENTAL VERSUS CORRELATIONAL DATA

A further distinction can be made between two types of prospective data. When making predictions about prospective data, a distinction can be made between phenomena that can be influenced by the researcher and phenomena that are beyond the researcher's influence. In some cases (such as weather forecasting), researchers have little or no opportunity to manipulate the initial conditions and observe the consequences. In other cases, researchers can initiate phenomena themselves or contrive or influence the initial conditions or context for some phenomenon, and then observe the ensuing consequences.

Disciplines that can or cannot influence the phenomena under study are methodologically distinct. When significant interaction with the phenomenon is possible, scholars can carry out formal experiments. For example, a psychomusicologist can directly manipulate the timbre of a sound and determine whether listeners from different cultures perceive the sound as "more cute" or "less cute." By manipulating single variables, an experiment allows the researcher to infer causality. A properly designed experiment allows the researcher to demonstrate that A has affected B rather than B affecting A. By contrast, researchers in historical disciplines cannot carry out controlled experiments. There is no way to go back into the past to change a single variable, nor is there any way to construct an independent world and observe the effects of specific manipulations. In the language of empirical methodology, historical disciplines necessarily rely on *correlational* rather than *experimental* methods.

In correlational studies, the researcher can demonstrate that there is a relationship or association between two variables or events. But there is no way to determine whether *A* causes *B* or *B* causes *A*. Moreover, the researcher cannot dismiss the possibility that *A* and *B* are not causally connected. It may be the case that both *A* and *B* are caused by an independent third variable. By way of illustration we might note that there is a strong correlation between consumption of ice cream and death by drowning. Whenever ice cream consumption increases there is a concomitant increase in drowning deaths (and vice versa). Of course the likely reason for this correlation is that warm summer days lead people to go swimming and also leads to greater ice cream consumption. In historical disciplines, one can never know whether the association of two events is causal, accidental, or the effect of a third (unidentified) event or factor.

DATA RICH AND DATA POOR

Of all the taxonomic distinctions made in this article, probably the most seminal is the distinction between *data*-rich and *data*-poor areas of research. Although the term "data" unfortunately implies something scientific, I intend the term to be construed in the broadest possible sense, meaning any information, observation, artifact, or evidence that may be pertinent to some theory, hypothesis, interpretation, or intuition. (In Latin, *datum*: a thing known, or passed around.)

Data-rich disciplines are in principal able to uncover or assemble as much information, evidence, observations, etc. as they wish, limited only by financial resources. Data-poor disciplines have little control over the volume of pertinent data. As noted earlier, no amount of money will necessarily increase the volume of documents relating directly to a historical figure's life.

There are four ways a field can be data-poor. One way is that the phenomenon itself is comparatively rare. It is difficult to study phenomena such as ball lightning, monosyllabic vowel-consonant verbs, white Bengali tigers, or multiple personality disorder. Few historical musicologists will experience the thrill of discovering a manuscript for an unknown work by a major composer.

A second way by which a field may be data-poor, is that the data may be volatile or is quickly destroyed. For the paleontologist, soft body tissues disappear in a matter of years and so are difficult to study from fossilized rock samples. Some sub-atomic particles exist for less than a millionth of a second. For the psychomusicologist, the moment-by-moment expectations of a music listener are ephemeral and evanescent.

A field may also be data-poor because the data is inaccessible. Archeological data is smothered by dirt. Neutrinos are thought to be everywhere in large quantities, but they have no electrical charge and no mass, so they resist interacting with any detection device. Although hundreds of thousands of amateur sound recordings are made each year, musicologists find them difficult to study: how does one assemble the recordings of Bach keyboard works performed by amateurs in 1999?

Finally, data can simply be lost. The destruction of the famed ancient library at Alexandria transformed pre-Socratic philosophy into a notoriously data-poor field. A modern translation of *all* of the surviving pre-Socratic Greek texts runs to just 162 pages (Fitt, 1959). This includes the complete extant texts from the writings of Pythagoras, Thales, Anaximander and dozens of other classical thinkers. Musical examples abound: for example, not a trace remains of Dufay's *Requiem*.

POSITIVIST FALLACY

Data poor fields raise some special methodological concerns. One of these is the problem known as the *positivist fallacy*. If a phenomenon leaves no trail of evidence, then there is nothing to study. We may even

be tempted to conclude that nothing has happened. In other words, the positivist fallacy is the misconception that *absence of evidence* may be interpreted as *evidence of absence*.

Positivism had a marked impact on mid-twentieth century American psychology. In particular, the influence of logical positivism was notable in the behaviorists such as J.B. Watson and B.F. Skinner. The classic example of the positivist fallacy was the penchant of behaviorists to dismiss unobservable mental states as non-existent. For example, because "consciousness" could not be observed, for the positivist it must be regarded as an occult or fictional quality with no truth status (Ayer, 1936).

If it is true that the positivist fallacy tends to arise from data-poor conditions, then it should be possible to observe this same misconception in humanities scholarship – whenever data is limited. Consider, by way of example, the following argument from the distinguished historical musicologist, Albert Seay. At the beginning of his otherwise fine book on medieval music, Seay provides the following rationale for focusing predominantly on sacred music in preference to secular music:

"Although much music did exist for secular purposes and many musicians satisfied the needs of secular audiences, the Church and its musical opportunities remained the central preoccupation. No better evidence of this emphasis on the religious can be seen than in the relative scarcity of both information and primary source materials for secular music as compared to those for the sacred." (Seay, 1975, p.2)

In other words, Seay is arguing that, with regard to secular medieval music-making, absence of evidence is evidence of absence. Since secular activities generated little documentation, we have almost no idea of the extent and day-to-day pertinence of medieval secular music-making. For illiterate peasants, "do-it-yourself" folk music may have shaped daily musical experience far more than has been supposed. Of course Seay may be entirely right about the relative unimportance of secular music-making, but in basing his argument on the absence of data, he is in the company of the most rabid logical positivist. The positivist fallacy is commonly regarded as a symptom of scientific excess. However, it knows no disciplinary boundaries; it tends to appear whenever pertinent data are scarce.

PARSIMONY VERSUS PLURALISM

An important intellectual precursor of logical positivism can be found *Ockham's Razor*. William of Ockham promoted the idea that the number of factors entailed by an explanation should not be multiplied beyond those necessary. Modern philosophers more commonly refer to this as the *principle of parsimony* – namely, that one should prefer the simplest hypothesis that can account for the observed evidence. Unessential concepts, factors, or causes should be excised.

Of course the simplest explanation may not be the correct explanation. Biologists in particular have discovered that physiological processes are typically much more convoluted than would seem to be necessary. Nevertheless, there is methodological merit in eschewing unnecessary complexity. Every time an additional parameter or factor is introduced, the capacity for false-positive errors is increased considerably.

By way of illustration, consider a hypothetical music theory that purports to explain every possible 8-note melodic phrase constructed using pitches within the range of an octave. (There are over 800 million possible phrases of this sort.) Mathematically, every conceivable 8-note pitch sequence can be perfectly modeled using just 7 parameters. Any music theorist can easily posit 7 plausible factors that influence the shape of a phrase. For example, a phrase might be influenced by scale type, contour shape, degree of chromaticism, Schenkerian line, pitch proximity, gap-fill tendency, stylistic period, etc. However, if a researcher claims to have a melodic model that accounts for all possible 8-note pitch sequences using just 7 factors, then the researcher has done no better than chance. Limiting the number of parameters or factors dramatically decreases the likelihood of constructing a spurious model or explanation.

For the false-positive skeptic, the principal of parsimony holds merit, not because it reduces complex phenomena to simple phenomena, but because decreasing the number of variables reduces the chances of making a false-positive error. While increasing the number of contributing factors can make a model more realistic, regrettably, it also greatly increases the capacity for self-deception.

Three Faces of Reductionism

There are at least three ways of interpreting the term *reductionism*. One is the methodological injunction to use the least number of variables possible when formulating a theory. This view of reductionism is synonymous with the principal of parsimony, which we have just discussed. A second way of

understanding reductionism is the "divide and conquer" method of research. A third interpretation of reductionism is the "nothing but" mode of explanation. These latter two notions of reductionism are described below.

"Divide and conquer" reductionism endeavors to elucidate complex phenomena by isolating constituent relationships. Classically, the principal research tool for this form of reductionism is the concept of "control." It is commonly thought that control entails holding one or more factors constant while the "independent variable" is manipulated and the "dependent variable" is observed. However, control more commonly entails randomizing the potentially confounding variables. In taking a political poll, for example, pollsters expect that the number of variables influencing a particular opinion is very large. It is hopeless to assume that one can hold constant such a large number of factors. Consequently, researchers seek a random sample with the hope that unknown influences will tend to cancel each other out. The formal statistical argument in support of random sampling is quite compelling, so there is considerable merit to this method of control.

Using such methods of control, it becomes possible for a researcher to investigate the effect of a given factor on some complex phenomenon. By investigating one factor at a time, it is often possible to build a sophisticated model or theory of the phenomenon in question. When the number of factors is more than five or six, the divide and conquer strategy often becomes intractable due to the explosion of possible interactions between purported factors. Nevertheless, the approach can still help identify important relationships in real-world phenomena.

The more contentious form of reductionism may be called the "nothing but" mode of explanation. A reductionist attempts to explain complex phenomena as merely the interaction of simpler underlying phenomena; explanation proceeds by accounting for complex wholes in terms of simpler components. In this form of reductionism, the researcher aims to make statements of the form "*X is nothing but Y*."

Used in this sense, reductionism can be contrasted with what is sometimes called *holism*. A `holist' expects to explain phenomena as being greater than the sum of its parts (a process dubbed *synergism* by Buckminster Fuller). Frequently, synergism leads to "emergent properties" where complex phenomena cannot be predicted even when a thorough understanding exists of the underlying constituent phenomena.

In contrast to the holist, the "nothing but" reductionist seeks to explain all complex phenomena as convoluted manifestations of a handful of fundamental causes or interactions. Culture is just sociology, sociology is just psychology, psychology is just biology, biology is just chemistry, and chemistry is just physics.

One cannot help but be impressed by the breathless grandiosity of this program. If such a scientific reductive synthesis is true, it will represent one of the pinnacle achievements of human inquiry. If it is false, it will represent one of the preeminent intellectual blunders in human history.

Humanities scholars of many stripes have derided the reductionist project. Much of the objection originates in the unsavory esthetic repercussions of `nothing but' reductionism. It is argued that such reductionistic accounts "explain" only in the sense of *making flat* (*ex planum*). The world as an enchanting place is transformed into a prosaic, colorless, and seemingly senseless enterprise. Among humanities scholars, musicians and musicologists have been among the most vocal critics of `nothing but' reductionism. Music theorists explicitly embrace complexity and scorn simplicity.[10] John Cage cautioned strongly against such "logical minimizations." Moreover, Cage was prescient in recognizing that this reductive tendency is not limited to the sciences. It is surprising where one can find such "nothing but" forms of reductionism.

Consider, once again, postmodernism. The postmodernist/deconstructionist philosophy advocates the unpacking of concepts and utterances in terms of socially constructed roles and power relations (e.g. Hacking, 1995). Postmodernism has helped to expose innumerable subtle and not-so-subtle ways in which ostensibly rational discourse manifests convoluted forms of dominance and control. But postmodernism goes much further. The most abstract principles of law, philosophy, and even science are best understood from the point-of-view of politics: *everything reduces to politics*. Notice that in this formulation, postmodernism and deconstruction bear all the hallmarks of *nothing-but* reductionism. Any thought you care to express can be reduced to a political motive. A sociobiologist may believe a social phenomenon to be ultimately reducible to underlying chemical interactions. But the postmodernist trumps this reductionism by viewing all scientific discourses as ultimately reducible to power ploys. As in the case of the scientific reductive synthesis, one cannot help but be impressed by the breathless grandiosity of such postmodernist patterns of explanations.

There is, I would suggest a more helpful way of understanding the value of reductionism while avoiding some of the more unsavory excesses (in both the sciences and the humanities). A helpful distinction is to treat "reductionism" as *a potentially useful strategy for discovery* rather than *a belief about how the world is*. Concretely, the postmodernist might *use* the assumption of hegemony as a technique to help unravel a complex behavior. Similarly, the sociobiologist might use the assumption of a recessive gene as a technique to help analyze a personality trait. In both cases, there are dangers in assuming that the tool is the reality. But in both cases, there remains the possibility that the reductive explanatory principle proves useful in understanding the phenomenon in question.

Humanistic and Mechanistic Beliefs

Our understanding of reductionism can be aided by contrasting the terms *reductionism* and *holism* with the philosophical distinction between *humanistic* and *mechanistic* views. The latter concepts might be defined as follows:

Humanistic: A belief in spirit and consciousness as fundamental, and not reducible to mechanical descriptions.

Mechanistic: A belief in a mechanical conception of life and consciousness. A belief that there is no essential mystery or enigma – there is only our ignorance of how things work.

Humanism and mechanism (as defined above) are beliefs, whereas reductionism and holism (as I've defined them) are methodological approaches. It is true that researchers who hold a mechanistic view of the world also tend to prefer reductionistic methods. It is also true that researchers who hold a humanistic view of the world tend to prefer or advocate holistic methods. However, there is no necessary link between humanism and holism, nor between mechanism and reductionism. There are many scientists (especially those working in the areas of complexity and chaos) who hold a mechanistic view of the world but who presume that complex interactions can lead to emergent properties that cannot be predicted (e.g., Anderson, 1972; Gell-Mann, 1994; Gleick, 1987; Pagels, 1988). In addition, a researcher can cogently hold a humanistic view of the origins of human behavior, yet rely on reductionism as a useful method for investigation. That is, one need not *believe* that human behavior is mechanistic in order to use reductionism as a way of probing the complexities of the world. Using reductionism as a research strategy does not commit a researcher to a mechanistic world-view. Similarly, analyzing a phenomenon as a holistic emergent property does not thereby transform the researcher into a spiritualist.

A Quantitative Role

Earlier we noted that "empiricism" simply means knowledge gained through observation. For many critics of empiricism, it not the idea of observational knowledge *per se* that raises concerns, but empiricism's widespread reliance on quantitative methods.

Perhaps the preeminent concern is that quantitative methods force phenomena into numerical categories that may or may not be appropriate. A researcher, for example, might ask listeners to rate musical excerpts on a scale from 1 to 10, where 1 represents "maximum sadness" and 10 represents "maximum happiness." This practice is open to innumerable objections: happiness and sadness may be independent phenomena that do not exist on some unified continuum; the musical excerpt may not retain a consistent character throughout the passage; a "poignant" passage might be both "happy" and "sad" simultaneously; a passage might be recognizable as *intending* to portray happiness, but a listener may find the portrayal unconvincing, and so "sadly" a failure; the numerical judgments may be uninterpretable (is the value 2 intended to be half as sad as the value 1?), etc.

Concerns such as these actually form much of the fundamental curriculum for training in quantitative methodology. For example, empiricists are taught that any judgment scale should use a single adjective (ranging from "least *X*" to "most *X*") rather than using mixed adjectives ("most *X*" to "most *Y*"). Similarly, empiricists learn that measurements are never to be construed as direct indices of actual phenomena, and operational definitions should not be reified. Statisticians have devised completely independent analytic procedures, depending on the properties of various measurement scales.

For many humanistically-inclined scholars, however, there remains something inherently wrong about quantifying human experiences – especially those experiences related to human attachment, esthetic experience, and spiritual life. Many scholars would agree with Renato Poggioli's view that the technical and quantitative have their place, but not in the arts:

> "'Technicism' means that the technical genius invades spiritual realms where technique has no raison d'être. ... It is not against the technical or the machine that the spirit justly revolts; it is against this reduction of nonmaterial values to the brute categories of the mechanical and technical." [p.138]

Once again, let me respond to this view by distinguishing *methodologies of scholarly inquiry* from *philosophical beliefs about the nature of the world*. Lest this distinction seem too abstract, consider the following extended illustration, which draws a parallel to scholarly attitudes regarding the use of writing and musical notation.

Socrates famously criticized the new fangled invention of writing. He rightly pointed to a number of predictable, yet questionable, consequences of relying on written texts. Specifically, Socrates predicted a decline in the importance of rote memory, and the waning of oratory skills.

Socrates' predictions have been amply proved correct. Few modern children can recite more than a single poem, politicians rely on teleprompters, and humanities scholars make public presentations with their heads buried in dense texts that leave listeners confused. Socrates' legitimate criticisms notwithstanding, writing caught on. In fact, writing was soon recognized as providing an invaluable window on previously unknown phenomena. With writing, for example, the Greeks discovered grammar. By removing speech from the ephemeral moment, the ancients discovered "parts of speech" (nouns, adjectives, particles, etc.) as well as tenses, conjugations, sentences, plots, and other structures. In short, the invention of writing provided an unprecedented opportunity to better understand language, and (paradoxically) speech.

An almost identical history attended the advent of musical notation. Music theorizing was common long before music was written down. But music notation unquestionably inspired and facilitated the growth of music theory in the West. As in the case of written language, musical notation allowed those who study music to identify patterns of organization that would otherwise be difficult or impossible to discern.

Of course, like Socrates, musical notation has drawn its critics. Jazz musicians are likely to resonate with the observations of a nineteenth century Arab traveler to Europe, Faris al-Shidyaq:

> "The Franks [Europeans] have no `free' music unbound by those graphic signs of theirs ... so that if you suggest to one of them that he should sing a couple of lines extempore ... he cannot do so. This is strange considering their excellence in this art, for singing in this fashion is natural and was in use among them before these graphic signs and symbols came into being." [As quoted in Nettl, 1985, p.123]

A perhaps unfortunate repercussion of musical notation has been the reification of notation as music. The very noun "music" has today acquired meanings that would have confounded ancient musicians. In modern times it is possible for "music" to fall off a stand or to be eaten by one's dog. Consider philosopher Nelson Goodman's well-known conception of the identity of the musical work:

> "A score, whether or not ever used as a guide for a performance, has as a primary function the authoritative identification of a work from performance to performance. Often scores and notations – and pseudo-scores and pseudo-notations – have such other more exciting functions as facilitating transposition, comprehension, or even composition; but every score, as a score, has the logically prior office of identifying a work." (Goodman, 1976/1981; p.128).

For Goodman, the notion of the existence of a musical work devoid of any score is a highly complex and thorny philosophical issue. In Goodman's view, the very identity of "music" is intimately linked and equated with material notational artifacts of a certain sort. This is what is meant by "reification."

As in the case of written language and musical notation, quantitative methods provide (1) important opportunities for glimpsing otherwise invisible patterns of organizations, and (2) similar opportunities for reification and fetishism. Scholarly attitudes toward musical notation are rightly mixed: notation has provided extraordinary opportunities for scholarly inquiry, but it has also spawned some moot and questionable beliefs regarding the nature of the musical world.

In the case of applying quantitative methods in music scholarship, we are a long way away from such excesses. On the contrary, music scholarship has barely begun to take advantage of the genuine opportunities provided for better understanding musical organization. Of the many examples that can be used to illustrate the promise of quantitative empirical methods, two examples must suffice. My first example relates to the concept of the "melodic arch" whereas the second example relates to the concept of "gap fill".

THE MELODIC ARCH

For centuries, music theorists have drawn attention to the so-called "melodic arch" – a presumed general tendency for melodic phrases to ascend and then descend. An example of an arch-shaped phrase might be the opening phrase *of My Bonnie Lies Over the Ocean*. Unfortunately, there are also lots of counter-examples*: Joy to the World* and the *Star Spangled Banner* are just two of many melodies that exhibit "convex" initial phrases.

What is one to make of the concept of the "melodic arch"? Is it true that there is a general arch tendency in musical phrases? Or have textbook writers simply been selective in their examples?

Huron (1996) carried out a study involving more than 36,000 melodic phrases sampled from European folksongs. The first question to resolve is one of definition: what is an "arch"? One way to define an arch is that all of the notes in the first half of the phrase rise upward in pitch, while all of the notes in the second half of the phrase move downward. A less restrictive definition might simply require that the average pitch-height of the initial and final notes of a phrase are lower than the average pitch-heights of the mid-phrase notes. Alternatively, one might determine phrase contours only after non-structural tones have been discarded. Without resolving the issue of what we mean by an "arch", Huron's study used several different operational definitions and found that the results were the same no matter how one defines an arch. By way of illustration, Figure 1 (below) shows the results of just one way of addressing the matter. The figure shows what happens when the pitch heights of 6,364 seven-note phrases are averaged together.

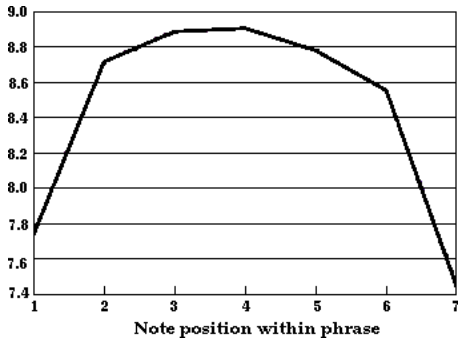

Fig. 1. Average melodic contour calculated for 6,364 seven-note folksong phrases displaying a melodic arch (Huron, 1996).

In Huron's study, each of the alternative notions of a "melodic arch" converged on the same answer. Although there are many individual phrases that do not exhibit an arch-shape, the great majority of phrases do indeed have a roughly ascending-descending contour. That is, the results are consistent with a general theoretical notion of a melodic arch (at least in Western folksong melodies). One might suppose that averaging together thousands of melodic phrases constitutes the epitome of quantitative lunacy. Yet, such simple quantitative procedures can prove remarkably useful in addressing certain kinds of musical questions.

GAP FILL

A common criticism of empirical studies in music is that they merely confirm our intuitions. A good counter-example is provided by the phenomenon of "gap fill." For 500 years, music scholars have observed that large melodic leaps tend to be followed by changes of melodic direction. This phenomena goes under a number of names, but let us use Leonard Meyer's terminology: "gap fill."

In a series of empirical studies, Paul von Hippel (2000a, 2000b, von Hippel & Huron, 2000) carried out extensive empirical investigations of the gap fill concept. The results are not at all consistent with music theorists' intuitions about gap fill. The story has two parts:

> (1) It is indeed the case that the majority of large intervals tend to be followed by a change in melodic direction. This pattern occurs in melodies from cultures spanning five continents and 500 years. This pattern is evident both for immediate pitch continuations, as well as delayed pitch continuations.
>
> However ...
>
> (2) If you completely scramble the order of notes within a melody, you end up with "random" melodies that tend to have exactly the same amount of gap fill as the original melodies themselves. This pattern occurs in melodies from cultures spanning five continents and 500 years.

The fact that scrambled (randomly reordered) versions of the same melodies produce the same gap fill tendency suggests that gap fill is not a consequence of compositional intent.

There is a straightforward explanation for why this happens – a phenomenon that statisticians call "regression toward the mean". A large leap will have a tendency to take the melody towards the upper or lower extremes of a melody's range. Having landed (say) near the top of the range, the melody has little choice but to continue with one of the lower notes. In real music, the closer the leap is to the extremes of the range, the more likely the contour changes direction. When a leap lands in the middle of the tessitura, reversing direction is no more common than continuing in the same direction.

Quantitatively, this account is very strong. After accounting for regression-toward-the-mean, there is no residual melodic behavior that can be attributed to a hypothetical principle of gap fill. While research on peripheral aspects of this issue continues, at this point it appears that "gap fill" is a musical concept without any reality in a large and diverse sample of actual notated music[11].

The Promise of Quantitative Methods

As I have argued, quantitative methods are important for the same reason that musical notation can be important: like musical notation, quantitative methods allow us to observe patterns of organization that might otherwise be difficult or impossible to decipher. For the new empiricist, an interest in quantitative methods has nothing to do with science. It has everything to do with becoming a more observant music scholar.

Consider, finally, the value of quantitative methodology in resolving how assertions are made in humanities scholarship. To the outsider, it often appears that the essence of scholarly debate is that one scholar believes that *X* is true, whereas another scholar believes that *X* is false. Most scholarly disagreements, however, relate to subtle shades of certainty. Consider, for example, the following assertions:

1. Tchaikovsky most certainly did not commit suicide.
2. Tchaikovsky very likely did not commit suicide.
3. Tchaikovsky probably did not commit suicide.
4. Tchaikovsky perhaps did not commit suicide.
5. Tchaikovsky may or may not have committed suicide.
6. Tchaikovsky perhaps committed suicide.
7. Tchaikovsky probably committed suicide.
8. Tchaikovsky very likely committed suicide.
9. Tchaikovsky most certainly committed suicide.

Most Tchaikovsky scholars suspect that Tchaikovsy did not commit suicide, but they disagree about the strength of the evidence, and hence they disagree about how a scholar should express this idea. Different scholars will accept (2), (3), or (4) above, but (1) will be considered excessive. While a given scholar may write (2) in a peer-reviewed journal, the ire of his or her colleagues may be provoked if his/her ensuing book prints (1) instead. Such are the proper nuances of scholarship.

Scholars familiar with quantitative methodology will immediately recognize that the disagreement amounts to uncertainty about the value of *p* (described earlier) – namely, the probability of making a false positive claim. In empirical research, the potential for mischief in reporting this idea would be circumvented by simply reporting the statistical confidence level.

Quantitative methods provide little benefit when the amount of data is as miniscule as that pertaining to Tchaikovsky's death. But there are innumerable musical issues where quantitative methods are indispensable and powerful. Conductors may pride themselves on their unprejudiced golden ears, but economists Claudia Goldin and Cecilia Rouse have assembled the concrete numbers comparing blind and non-blind auditions: the results are consistent with a rampant and systematic discrimination against female orchestral musicians (Goldin & Rouse, 2000).

In assessing the writings of another scholar, how are we to know whether the writer is guilty of using exaggerated rhetoric? Like the Tchaikovsky researcher commenting on Tchaikovsky's death, scholars may rightly wonder whether the assertion of, say, a feminist scholar, is being overstated or understated. But for those who understand quantitative methods, the numbers can be far more compelling – and far more damning – than any rhetorical flourish.

USING THE RIGHT METHODOLOGY

By now it should be clear that I regard methodologies as tools for conducting research, not as philosophical belief systems. Like all tools, a given methodology is suitable for certain kinds of research but not other kinds. In pursuing a research program, thoughtful scholars will take stock of the conditions of their area of research and choose a methodological approach that best caters to the research goals – in light of the opportunities and dangers inherent in specific tasks. The most appropriate methodology may change depending on the specific task or hypothesis being addressed.

The principal impediment to careful selection of field-appropriate methods is the methodological inertia to be found in most disciplines. Researchers typically are taught only a single methodology and so expect this method to be applicable (to a greater or lesser degree) in virtually all research tasks. In the words of Abraham Maslow, to the person who holds a hammer, all the world looks forever like a nail.

Even when researchers tend to use field-appropriate methods, we are often insensitive to the subtle changes in a field that ought to cause us to revisit and revise our methodological strategies and commitments. In the remaining sections, we consider some of the misconceptions and failures that attend either (1) failing to recognize field-specific differences, or (2) failing to recognize changing conditions within a field of research.

Understanding Humanities Methods

Scientists sometimes express dismay at the low levels of evidence that appear to be common in humanities disciplines. These views are often misplaced for two reasons. First, many humanities activities address `low risk' hypotheses in the sense that committing a false-positive error has modest repercussions. Second, data-poor disciplines simply cannot be expected to satisfy high standards of evidence.

Faced with often paltry volumes of data, most scientists would never consider pursuing the sort of research projects found in the humanities. The scientist might be tempted to conclude that no knowledge claims ought to be made. *However, this presumes that false-negative errors have no moral repercussions*. It may be the case that 'the lessons of history' are poorly supported and unreliable, but what are the consequences of concluding that it is impossible to learn from history? Historians are right, I believe, to try to make sense of incomplete and eclectic historical evidence – since our failure to learn from this material may doom us to repeat past mistakes.

In general, it should not be surprising that researchers in data-poor fields are typically oriented to theory-conserving skepticism rather than theory-discarding skepticism. When data is scarce, pursuing a theory-discarding skepticism means that one must always conclude that no conclusion is possible: no hypothesis can be supported, no theory is tenable. Any scholar having this disposition will naturally

abandon the discipline.

There are circumstances, however, where the dismay expressed by scientists concerning evidence in humanities disciplines is proper and appropriate. Specifically, these criticisms are warranted (1) when the risks of committing a false-positive error have significant moral (or esthetic) repercussions, and (2) when the field is not, or need not be, data poor. Both of these circumstances arise with some regularity in traditional humanities disciplines. Moreover, either one of these circumstances necessitates significant changes in methodology. An instructive historical example may be found in the splitting of the social sciences from the humanities.

The Social Sciences Split: Risky Hypotheses and Data Riches

Humanities disciplines deal with human behavior, civil society, and culture. Humanities scholars regularly make claims about human nature, about moral and immoral conduct, and render advice about political, educational and cultural institutions. Scholars' views concerning these areas of human affairs can, and often do, have significant impact. At the end of the nineteenth century, the social sciences began to drift away from traditional humanities approaches precisely because thoughtful scholars recognized the need for higher standards of evidence in support of knowledge claims, especially those claims that might influence public attitudes and public policy.

In recognizing the risks of committing false-positive errors, social scientists were right to initiate changes in their research methods. Contributing to this revolution in methodology was the realization that the social sciences could conduct research that would significantly increase the volume of evidence that could inform researchers' theorizing.

Over the decades, a number of humanities scholars have criticized contemporary psychology and sociology for adopting methods more commonly associated with the physical sciences. However, these criticisms are based on the false assumption that disciplines are defined, not only by their subject matter, but also by their methods. As we have seen, methods arise not from the subject of research, but by the riskiness of the hypotheses, by the availability of pertinent data, by the ability of researchers to observe the effects of *a priori* manipulations, and by the opportunity to collect evidence independent from the original evidence used to formulate some theory or interpretation.

It is wrong, I believe, to portray methodologies as competing philosophical allegiances. It is not a question of whether "scientific" methods prevail over interpretive, hermeneutic, phenomenological, or other traditional humanities methods, or vice versa. The question is whether researchers use the best methodology (or `basket' of methods) for the task at hand.

To many scholars, it appears that over the course of the twentieth century, the humanities "lost" a number of disciplines – including linguistics, archeology, psychology, and (to a lesser extent) anthropology and sociology. I disagree. The subject matter of these disciplines has changed little over the past century. Linguists are still interested in the origins, structures and acquisition of human languages. Archaeologists are still interested in how artifacts inform us about past human civilizations. Psychologists are still interested in human thoughts and motivations. Sociologists and anthropologists are still interested in the nature of human interaction and the nature of culture. In each discipline, human beings and human lives remain central. What has changed for these disciplines is primarily the volume of available evidence – and consequently the opportunities to address more refined questions using methods that better exploit the expanded data resources.

The prospect of gaining access to increased data is not merely an opportunity to be taken or ignored, as one pleases. Where pertinent data is readily available, it is morally reprehensible not to use it since failing to use the data increases the likelihood of making *both* false-positive and false-negative errors. In short, empirical data deserves our attention for precisely the same reason that small amounts of historical data warrant the historian's best interpretive efforts: failing to attempt to learn from the information at hand is to encourage and condone ignorance.[12]

PARTICLE PHYSICS: THE REPERCUSSIONS OF DECREASING DATA

Although circumstances can open the flood-gates of data, circumstances can also close them. Admittedly, it is less common for a discipline to experience a reduction in the volume of data, but it does happen. The field of particle physics is arguably such a field. The very success of sub-atomic physics has pushed the frontier of study to more and more esoteric corners of reality. Particle physicists cannot carry out

experiments without access to enormously costly machinery. After spending roughly $2 billion preparing to build the super-conducting super-collider (SSC), in 1993 the U.S. government decided to abandon the venture as too costly. Although particle physicists can continue to collect data, physicists have few opportunities to collect data that is pertinent to the latest theoretical models and issues.

Even if the SSC had been built, its utility would have been limited. The most developed theories of physical reality exceed our abilities to test them. For example, in order to test hypotheses arising from *superstring theory*, it has been estimated that a suitable particle accelerator would need to be 1,000 light-years in circumference (Horgan, 1996; p.62). With the increasing scarcity of pertinent data, sub-atomic physics is slowly being transformed into a purely theoretical enterprise. Already, quantum physics has attracted innumerable competing interpretations with little hope that tests will ever be done that might prune away the incorrect interpretations. Nobel laureate, Sheldon Glashow expresses the malaise in his field as follows: "contemplation of superstrings may evolve into an activity ... conducted at schools of divinity by future equivalents of medieval theologians." (Glashow & Ginsparg, 1986; p.7).

Glashow's allusion to theology is derisive. But particle physicists may need to get used to the apparently inevitable methodological transformation that awaits their discipline. Humanities scholars can be forgiven for shedding crocodile tears: for centuries, historians have had to struggle to make sense of manuscript fragments that they knew would never be made whole. When data is finite, interpretation is the only scholarly activity that remains. Moreover, the interpretive, hermeneutic enterprise is an activity that remains of value.

MUSICOLOGY: THE REPERCUSSIONS OF INCREASING DATA

While sub-atomic physics is moving into a period of data scarcity, the reverse situation appears to be happening for music. As noted earlier, technical and organizational innovations can transform data-poor fields into data-rich fields. Over the past 25 years, such innovations have arisen in many areas of musical study – following the trends of such disciplines as linguistics, education and anthropology. Contemporary music scholars have access to computational and database resources, comprehensive reference tools, high quality data acquisition methods, sophisticated modeling techniques, and other innovations that make it far easier to collect, analyze and interpret musically-pertinent evidence and artifacts. There is hardly any area of music that cannot benefit from the increased resources, and from the ensuing opportunity to adopt more rigorous standards of evidence. This includes areas such as manuscript studies, poietics, history, iconography, analysis, performance, pedagogy, reception, esthetics and criticism, phenomenology, social and critical theory, cultural studies, cultural policy, media, and ethnology. Not all areas of music scholarship have, or will be touched by the expanding resources. Nor will speculative and creative music philosophy entirely lose its value.

The changing landscape in musicology towards more empirical approaches is not a displacing of the humanities spirit by an antithetical scientific ethos. It is fundamentally a response to a clearer epistemological understanding of the role of methodology. Changing conditions simply allow us to be better music scholars, to embrace higher standards of evidence, and to be more acutely aware of the moral and esthetic repercussions of our knowledge claims, including claims that something is unknowable or that some phenomena ought not to be investigated. Our strongest criticisms should be levied at those who insist on speculative discourse when the resources are readily available to test such knowledge claims.

IMPACT ASSESSMENTS IN HUMANITIES DISCOURSE

The above discussion has only cursorily addressed the issue of evaluating the moral and esthetic repercussions of various knowledge claims. Few aspects of humanities discourse are in greater need of discussion. I believe it is imperative that humanities scholars not be cavalier about the impact and importance of ideas. It is dangerous to suppose that, in comparison to *technologies* (with their considerable potential for mischief), *ideas* are somehow fragile and innocent. Karl Marx never failed to denigrate what he called "mere ideas." Philosophers, he said, have been content simply to talk about the world, with little interest in changing it. It is unfortunate that Marx never lived to see the cruel irony of his words. No other individual had so marked a *moral* effect on twentieth century lives as Karl Marx. Yet Marx himself was the quintessential closeted philosopher. Before letting an idea loose on the world, ideas ought to be subject to the same environmental impact assessments we apply to roadways and chemicals. Half-baked ideas have been just as disruptive and damaging as any technological innovation – probably more so. It is important

that humanities scholars stop underestimating our power to change the world. At the same time, it is important not to underestimate our culpability when we get things wrong.

METHODOLOGY AS POT-HOLE GUIDES

Possibly the most pervasive misconception about methodology is that scholarly methods provide algorithms for carrying out research. According to this view, a methodology is a sort of recipe that scholars follow in the course of their studies. In this view, the function of epistemologists is presumed to be to concoct increasingly refined and more detailed methodological algorithms. The origin of this view may be linked to similar misconceptions about procedures in mathematical proofs. While the deductive procedures used by mathematicians are indeed rule-bound, mathematical research itself is a much more woolly-headed enterprise.

As noted in the Part I, in the twentieth century, the idea of "methodology as algorithm" has come under sustained and devastating attack (Agassi, 1975; Feyerabend, 1975; Gellner, 1974; Kuhn, 1962; Laudan, 1977; Popper, 1934; Polanyi, 1962; Quine, 1953; and others). Many of these attacks have come from authors whose motivation was a defense of the rationality of science. The overwhelming conclusion from these critiques is that no known set of rules can guarantee the advance of knowledge. Moreover, as we have seen, even the most flexible known methodology `rule' yet proposed, Feyerabend's *anything goes*, fails to be born out by observation.

Of the various efforts to reformulate our understanding of scholarly methodology, one of the best informed and most nuanced has been the view offered by the epistemologist Jagdish Hattiangadi. In his *Methodology without Methodological Rules*, Hattiangadi (1983) argues that, like scientific theories, methodological theories are activities of discovery, for which there are not fixed rules. The scholar who slavishly follows a fixed methodology will ultimately make an onerous mistake.

Hattiangadi regards fields of scholarship as debating traditions that develop problems and criteria as they go. Although rationality is tradition-bound, rationality is not constrained solely by what we believe. What methodologists discover is a series of guidelines or heuristics.

In our long history of making mistakes, scholars have come to identify common `pot-holes' on the road to understanding. Humanities scholars have learned to recognize and avoid a multitude of logical and rhetorical fallacies, including *ad hominem* arguments, appeals to authority (*ipse dixit*), the naturalist fallacy, the positivist fallacy, reification or hypostatization, and a host of pitfalls in forming historical explanations (Elster, 1989; Fischer, 1970; Roberts, 1996). Similarly, contemporary scientists have identified innumerable additional dangers. Among these dangers are the problem of hindsight reasoning, experimenter bias, ceiling effects, demand characteristics, the multiple tests problem, the third variable problem, cohort effects (Schaie, 1986), and the reactivity problem (Webb, Campbell, Schwartz, Sechrest, & Grove, 1981). These (and many other problems) are all well documented, and in many cases effective guidelines have been devised to recognize, avoid or minimize their detrimental effects on scholarship.

Researchers are free to choose or develop their own methodology – whether deductive, empirical, phenomenological, or whatever. But the pursuit of knowledge is best served when scholars learn from the various existing debating traditions. Although there is no detailed road-map for pursuing research, there exist sketches of well-documented pot-holes that others scholars have already encountered. It is important for scholars to be aware of these known hazards and for disciplines to keep abreast of methodological discoveries. Methodology is not simply some abstract specialty of philosophy. It is a utilitarian cross-disciplinary consultancy that offers pragmatic day-to-day assistance for all researchers.

Here, regrettably, postmodernism has done humanities scholarship a grave disservice. Many otherwise thoughtful people are convinced there is no possibility of rigor, and that methodology is a dangerous illusion. As a result, an entire generation of students in the arts and humanities has been deprived of adequate practical education relating to methodology. To the postmodernist skeptic, one must respond with the reverse skepticism: What if there are truths? What if some truths are knowable? What if some interpretations are better than others? What if we fail to learn from the evidence that is available to us?

CONCLUSION

By way of review, the basic arguments I have presented can be reconstructed and summarized as follows:

1. Postmodernists are right to note that knowledge claims do not take place in a moral vacuum. Theories, hypotheses, interpretations and opinions carry moral (and esthetic) repercussions. Moreover, choosing to avoid making knowledge claims is similarly an act with moral consequences.
2. Anyone wishing to make any knowledge claim about the world, has no choice but to navigate the treacherous path between false positive and false negative errors. This includes claims that say `I don't know' and `We cannot know.' There is nothing epistemologically safer about these negative claims compared with the corresponding positive claims `I know' or `In principal, we can know.'
3. The "Problem of Induction" is intractable and omnipresent: no amount of observation can establish the truth of some proposition. This problem applies not only to empiricism, but also to the critiques of empiricism offered by anti-foundationalist writers like Feyerabend. No amount of observation about the history of science can establish the general claim that the enterprise of science is irrational or arational.
4. Despite the problem of induction, observation remains indispensable to knowledge in ways we do not understand. Our very biological machinery has evolved to facilitate acquiring knowledge about the world. We can show that observations are consistent with some theories and not other theories – even though we cannot *prove* that one theory is better than another.
5. Fields of study differ according to the volume and quality of available evidence ("data") used to support or assess different claims, views, interpretations, or theories.
6. When data are inaccessible or non-existent, the field is susceptible to the *positivist fallacy* – that absence of evidence can be interpreted as evidence of absence.
7. Data-poor fields are unable to support research whose goal is to minimize false-positive claims. Theory-discarding skeptics therefore avoid pursuing research in data-poor fields; they conclude that no conclusions can be drawn from the available data.
8. Other scholars will recognize the possibly onerous moral repercussions from failing to attempt to learn from small amounts of data/evidence. Data-poor fields will attract only theory-conserving skeptics, that is, scholars whose goal is to minimize false-negative claims.
9. When the volume of data is small, false-negative skeptics are logically consistent when they support multiple alternative hypotheses or interpretations. Pluralism is therefore preferred over parsimony. Conclusions are open rather than closed.
10. Unfortunately, scholars working in data-poor fields will typically make innumerable false-positive errors. That is, many ideas will be promulgated that lack merit.
11. Data-rich fields provide greater power for hypothesis testing. More stringent criteria allow testing that minimizes false-positive claims. As a result, competing hypotheses can be rejected with some assurance. Parsimony is therefore preferred to pluralism. Researchers aim for closed explanations.
12. Data can also be characterized as retrospective or prospective. Retrospective data invites two methodological problems. First, retrospective data is susceptible to unfettered "story-telling:" scholars are adept at formulating theories that account for any existing set of data. That is, it is tempting to use retrospective data both to formulate an explanatory theory and to provide evidence in support of the theory. A second problem with retrospective data is that possible causal relationships cannot be inferred.
13. In contrast to retrospective data, prospective data makes it possible to challenge theories or stories by comparing predictions to new data. Few demonstrations of the possibility of knowledge are more compelling than predicting otherwise improbable observations.
14. A distinction can be made between two types of prospective data: data that can be influenced by the researcher, and data that cannot be influenced. Influenced future data allows the manipulation of initial conditions, and so in principle allows the researcher to infer possible causality. If the researcher cannot manipulate experimental variables, then possible causal relationships cannot be inferred.
15. Whether one holds a theory-conserving or theory-discarding skeptical attitude should depend on the moral repercussions of making a false-positive or false-negative error. This risk will change from one claim/hypothesis/interpretation to the next.

16. Scholars in all fields of study ought to maintain flexibility in choosing a methodology that is suited to the task at hand. That choice should be informed by both the ethical repercussions of making various types of errors, as well as by the particular circumstances of the field itself.
17. In nearly every case, scholarship is enhanced by the availability of additional evidence. Like prosecuting attorneys, scholars have a moral obligation to seek out additional sources of evidence/data whenever these can be obtained. The magnitude of this obligation is proportional to the moral repercussions of the hypothesis.
18. Inferential statistical tests can be used equally effectively by both theory-conserving and theory-discarding skeptics. Theory-conserving skeptics have under-utilized statistical tests.
19. The material and structural conditions of any field of research are susceptible to change. A common source of change is either an increase or decrease in available pertinent data. Changing conditions often demand changes in research methodologies in order to minimize moral risks.
20. The selection of an appropriate methodology is a moral decision. When a scholar is unaware of the methodological choices, the selection of a methodology will be morally uninformed.
21. Research methodologies should be regarded as scholarly tools; researchers should resist the tendency to hold methodologies as comprehensive belief systems about the world.
22. There is no known methodological algorithm that ensures the advance of knowledge. Methodology consists primarily of a set of pointers that warn scholars of previously encountered pitfalls. Methodologies are extended and refined in the same manner as other theories.

In this paper, I have endeavored to rekindle the view that the humanities are distinguished from the sciences primarily by their subject matter, and secondarily by a philosophical tendency towards humanistic rather than mechanistic conceptions of the world. More importantly, I have argued *against* the idea that the sciences and humanities are necessarily distinguished by their methodological habits. It is true that humanities disciplines currently tend to embrace false-negative skepticism, tend to be historical in orientation, tend to prefer pluralism to parsimony, and tend to prefer open accounts rather than closed explanations. However, I have noted that these methodological tendencies primarily arise from the structures and material circumstances attending the particular fields of study involved. Specifically, many humanities disciplines (though not all) are comparatively data-poor, deal with lower risk hypotheses, and are unable to carry out formal experiments. Data-poor disciplines repel false-positive skeptics because such disciplines provide an environment where false-positive skepticism is not productive.

My claim that methodological differences arise primarily from the concrete research conditions of individual disciplines should evoke no surprise. Philosophers of knowledge all presume that what might loosely be called “rationality” is not discipline-specific. What is good for the epistemological goose ought to be good for the epistemological gander as well.

Fields of study do have discipline-specific methodological needs. For example, manuscript studies have developed analytic methods based on water marks, chain lines, binding patterns, and so on. But there are also underlying patterns to how different disciplines approach their goals, and there are some unifying principles in research. In summary, while the humanities and sciences may rightly diverge in their philosophical conceptions about the nature of the world, they nevertheless share deep methodological commonalities. All fields of study can greatly benefit from an awareness of both the wide variety of available research methods and the innumerable pointers to methodological potholes.

The New Empiricism

Research begins when we ask questions about the world. In the case of music, there is a multitude of worthwhile questions that can be posed. In many cases, there are negative moral repercussions if we choose not to investigate some question. Offering the excuse that “we could never be certain about the answer to that question” is hollow rather than noble, since it applies to all empirical questions. Good questions rightly challenge scholars to do our best to assemble evidence that might help produce informed (albeit limited and provisional) answers.

Over the past decade, increasing numbers of music scholars have become attracted by the opportunities offered through empirical methods. The new empiricism recognizes that formal observation can indeed potentially lead to genuine insights about music and musicality. As I have noted, what the new empiricism shares in common with postmodernism is the conviction that scholarship occurs in a moral realm, and so methodology ought to be guided by moral considerations.

Of course some research questions are hampered by a dearth of pertinent evidence. Nevertheless, there are reasonable ways of trying to decipher *likelihood* – even if we can never divine the *Truth*. Many questions allow us to collect lots of pertinent data, and to use inferential statistical methods that allow us to minimize both false-positive and false-negative errors.

The new empiricism has three bones to pick with the sciences. Scientists are wrong to denigrate or ignore fields that are data-poor and areas of research where experimentation is impossible. Scientists are wrong to treat the 0.05 confidence level as some sort of immutable inferential standard. For severely data-limited fields, 0.10 and 0.20 confidence levels ought to be entertained when the risks associated with making a false positive error are low. Scientists are also wrong to assume that the goal of research must always be to minimize false-positive errors.

Similarly, the new empiricism also has some bones to pick with our colleagues in the humanities. Empiricism is not a dirty word. There are many musical questions, from history, esthetics, culture, analysis, theory, performance, poeitics, reception, listening, etc. which can be usefully addressed using inferential statistical methods. Contrary to a popular belief, statistics cannot be used to prove any point of view.

To the traditional music scholar, it must look for all the world like science is muscling-in on musicology. But the rise of empiricism has nothing to do with "science". It arises from within music scholarship, and is motivated by the desire to learn as much as possible from the information available to us – including the additional information that might be assembled with a little effort. The pursuit of evidence is a moral obligation. Once again, the analogy to jurisprudence is compelling: if a prosecuting attorney has the opportunity to gain access to a wealth of new evidence, it would be morally reprehensible not to examine the material in order to better establish the guilt or innocence of someone.

The pursuit of rigor is not some sort of methodological fetish. It is simply an attempt to avoid well-documented pitfalls in research. We ought not to be cynical of those scholars who aspire to do their best.

In light of the above observations concerning methodology, it should be obvious that I think both humanities scholars and scientists should be educated with an aim to providing a broader repertoire of research methodologies. In particular, humanities scholars ought to learn the basics of statistical inference, and scientists ought to be exposed to phenomenological and deconstructionist approaches.

Finally, moral and ethical philosophers should take a greater interest in epistemological ethics. Knowledge claims have consequences, and it is important for scholars to be cognizant of the moral and esthetic repercussions of their views – including the view that something is unknowable. Better research on risk is needed in order to help researchers recognize when to adopt a theory-conserving or theory-discarding stance.

NOTES

[1] This paper was presented on October 8, 1999 as part of the Ernest Bloch Lectures in the Department of Music, University California, Berkeley. A printed version of this lecture was posted on the world-wide web on April 2, 2001 (http://www.musiccog.ohio-state.edu/Music220/Bloch.lectures/3.Methodology.html).

[2] It should be noted that the term "Positivism" is rarely used by modern empiricists; however, it is a designation commonly used in humanities scholarship, hence our use of it here. For a discussion of the so-called "culture wars" see: Alan Sokal and Jean Bricmont, *Fashionable Nonsense: Postmodern Intellectuals' Abuse of Science,* New York: Picador, 1998; and Joseph Natoli's *A Primer to Postmodernity*, Oxford: Blackwell Publishers, 1997 – notably Chapter 8: Postmodernity's War with Science.

[3] See Belsey (1993), Feyerabend (1975), Foucault (1970, 1977), Hartsock (1990), Kuhn (1962/1970), Natoli (1997).

[4] In the pithy words of Foucault, "There is no power relation without the correlative constitution of a field of knowledge, nor any knowledge that does not presuppose and constitute at the same time power relations." (p. 27).

[5] Throughout this article, the word "theory" should be interpreted broadly to mean any claim, hypothesis, theory, interpretation or view.

[6] A standard textbook on scientific method notes the following: "In contrast to the consequences of publishing false results, the consequences of a Type II error are not seen as being very serious." (Cozby, 1989; p. 147).

[7] It is essential to recommend *new* rather than established quacks. Established quackery has usually been

the subject of research that has failed to establish its efficacy. Untested quackery has a better chance of being helpful.
[8] Once again, the reader is reminded that throughout this article, the word "data" should be interpreted broadly to mean any information or evidence.
[9] An example will be given later in this article.
[10] "In musical interpretations, complexity is cherished ... In the social sciences, complexity seems to be avoided: the details of phenomena are levelled so that the findings can be expressed in the simplest possible way." (Rahn, 1983; p. 197).
[11] Statisticians have written extensively about the phenomenon of regression-toward-the-mean. Unfortunately, it appears to be a concept that is difficult for humans to grasp. Even Nobel laureate, W.F. Sharpe, incorrectly mistook regression-toward-the-mean for a new economic phenomenon (see, for example, Gary Smith, "Do Statistics Test Scores Regress Toward the Mean?"). As often happens with significant discoveries, a careful literature search sometimes finds that the same discovery was made decades earlier by another scholar. In a 1924 study, Henry Watt suggested that gap-fill in music can be attributed to regression toward the mean. Given the poor level of statistical numeracy among music scholars, I predict that it will take another 70 years before the preponderance of music theorists understand what has been demonstrated regarding gap fill.
[12] There may be statistical reasons for excluding some data from an analysis.

REFERENCES

Agassi, Joseph. 1975. *Science in Flux.* Dordrecht, Holland: D. Reidel Publishing Co.

Anderson, Philip. 1972. More is different. *Science,* Aug. 4, 1972, pp.393-

Ayer, Alfred J. 1936. *Language, Truth and Logic.* London: Victor Gollantz.

Belsey, Catherine. 1993. "Towards Cultural History," in: *A Postmodern Reader,* J. Natoli & L. Hutcheon (eds.) New York: SUNY Press, pp. 561, 555-556 [sic]. Quote is from page 561.

Cozby, Paul C. 1989. *Method in Behavioral Research.* Mountain View, CA: Mayfield Publishing Co., 4th edition.

Cook, Nicholas. 1990. *Music, Imagination, and Culture.* Oxford: Clarendon Press.

Duhem, Pierre. 1905. *La Théorie Physique: Son Objet, Sa Structure.* Paris: Marcel Rivière; 2nd edition, 1914. Translated by P.P. Wiener as: *The Aim and Structure of Physical Theory* Princeton: Princeton University Press, 1954; reprinted New York: Atheneum, 1962, 1977.

Elster, Jon. 1989. *Nuts and Bolts for the Social Sciences.* Cambridge: Cambridge University Press.

Feyerabend, Paul. 1975. *Against Method: Outline of an Anarchistic Theory of Knowledge.* London: Verso Editions 7.

Fischer, David H. 1970. *Historians' Fallacies: Towards a Logic of Historical Thought*. New York: Harper Collins.

Fitt, Mary. 1959. *Diels' Fragmente der Vorsokratiker.* Translated by Kathleen Freeman as *Ancilla to The pre-Socratic philosophers: A complete translation*. Cambridge: Harvard University Press, 1983; 5th edition.

Foucault, Michel. 1970. *The Order of Things: An Archaeology of the Human Sciences.* New York: Pantheon, 1970.

Foucault, Michel. 1977. *Discipline and Punish: The Birth of the Prison.* New York: Vintage/Random

House.

Gell-Mann, Murray. 1994. *The Quark and the Jaguar.* New York: W.H. Freeman.

Gellner, Ernest. 1974. *Legitimation of Belief.* Cambridge: Cambridge University Press.

Glashow, Sheldon & Ginsparg, Paul. 1986. Desperately seeking superstrings. *Physics Today*, May 1986, p.7.

Gleick, James. 1987. *Chaos: Making a New Science.* New York: Penguin Books.

Goldin, Claudia & Rouse, Cecilia. 2000. Orchestrating impartiality: The impact of `blind' auditions on female musicians. *American Economic Review,* September 2000.

Goodman, Nelson. 1976/1981. *Languages of Art; An Approach to a Theory of Symbols,* Brighton, Sussex: Harvester Press.

Gould, S.J. 1978. Sociobiology: The art of storytelling. *New Scientist,* Vol. 80, pp. 530-533.

Gould, S.J., & Lewontin, R.C. 1979. The spandrels of San Marco and the Panglossian program: A critique of the adaptationist programme. *Proceedings of the Royal Society of London,* Vol. 250, pp. 281-288.

Hacking, Ian. 1995. *Rewriting the Soul: Multiple Personality and the Sciences of Memory.* Princeton, N.J.: Princeton University Press.

Hartsock, Nancy. 1990. Foucault on power: A theory for women? In: Linda Nicholson (ed.) *Feminism/Postmodernism.* New York: Routledge.

Hattiangadi, Jagdish N. 1983. A methodology without methodological rules. In: R. Cohen & M. Wartofsky (eds.), *Language, Logic, and Method.* Dordrecht, Holland: D. Reidel Publishing Co., pp. 103-151.

von Hippel, Paul. 2000a. Redefining pitch proximity: Tessitura and mobility as constraints on melodic intervals. *Music Perception,* Vol. 17, No. 3, pp. 315-327.

von Hippel, Paul. 2000b. Questioning a melodic archetype: Do listeners use gap-fill to classify melodies? *Music Perception,* Vol. 18, No. 2, pp. 139-153.

von Hippel, Paul, & Huron, David. 2000. Why do skips precede reversals? The effect of tessitura on melodic structure. *Music Perception,* Vol 18, No. 1, pp. 59-85.

Horgan, John. 1996. *The End of Science: Facing the limits of knowledge in the twilight of the scientific age.* Reading, MA: Addison-Wesley Publisher.

Huron, David. 1996. The melodic arch in Western folksongs. *Computing in Musicology,* Vol. 10, pp. 3-23.

Kuhn, Thomas S. 1962/1970. *The Structure of Scientific Revolutions.* Chicago: University of Chicago Press.

Laudan, Larry. 1977. *Progress and Its Problems: Towards a Theory of Scientific Growth.* Berkeley: University of California Press.

Lewontin, R.C. 1991. *Biology as Ideology.* Concord, Ontario: Anansi Press and the Canadian Broadcasting Corporation.

Natoli, J. 1997. *A Primer to Postmodernity.* Oxford: Blackwell Publishers.

Nettl, Bruno. 1985. *The Western Impact on World Music,* New York: Schirmer Books.

Neyman, J. and Pearson, E.S. 1928 On the use and interpretation of certain test criteria for purposes of statistical inference. *Biometrika, A,* Vol. 20, pp. 175-240,263-294.

Neyman, J. and Pearson, E.S. 1967. *Joint Statistical Papers.* Cambridge: Cambridge University Press.

Pagels, Heinz. 1988. *The Dreams of Reason.* New York: Simon and Schuster.

Pascal, Blaise. 1669. *Pensées.*

Plotkin, Henry. 1994. *Darwin Machines and the Nature of Knowledge.* Cambridge, MA: Harvard University Press.

Poggioli, Renato. 1962/1968. *Teoria dell'arte d'avanguardia. Società editrice il Mulino.* Translated by G. Fitzgerald as *The Theory of the Avant-Garde.* Cambridge, MA: Harvard University Press.

Polanyi, Michael. 1962. *Personal Knowledge: Towards a Post-Critical Philosophy.* Chicago: University of Chicago Press.

Popper, Karl. 1934/1959. *Logik der Forschung.* Vienna, 1934. Trans. as The Logic of Scientific Discovery. New York: Basic Books, 1959.

Rahn, John. 1983. *A Theory for All Music: Problems and solutions in the analysis of non-Western forms.* Toronto: Toronto University Press.

Rahn, John. 1993. Differences. *Perspectives of New Music,* Vol. 31, No. 2, pp. 58-71.

Quine, Willard Van. 1953. *From a Logical Point of View.* Cambridge, MA: Harvard University Press.

Roberts, Clayton. 1996. *The Logic of Historical Explanation.* University Park, PA: Pennsylvania State University Press.

Rosen, D.E. 1982. Telostean interrelationships, morphological function and evolutionary inference. *American Zoologist,* Vol. 22, pp. 261-273.

Seay, Albert. 1975. *Music in the Medieval World.* Englewood Cliffs, NJ: Prentice-Hall Inc.

Watt, Henry J. 1924. Functions of the size of interval in the songs of Schubert and of the Chippewa and Teton Sioux Indians. *British Journal of Psychology*, Vol. 14, No. 4, pp.370-386.

Webb, E.J., Campbell, D.T., Schwartz, R.D., Sechrest, R., and Grove, J.B. 1981. *Nonreactive Measures in the Social Sciences.* Boston: Houghton Mifflin, 2nd edition.