Revival of Objectivity in Scientific Method

Doug Fraedrich

Introduction

One of Objectivism's many important contributions to philosophy is the distinction between the concepts of proof and validation. A proof is a deductive derivation from antecedent knowledge, while validation is establishing the relationship between the concept and reality (Peikoff 1991, 8). This distinction—that some concepts should be proven and others validated—has obviated many age-old problems in philosophy. However, this exploitation of concept validation puts a heavy burden on the philosophical system to use a reliable method for this validation.

Epistemology is the study of the nature of knowledge and its justification. This includes the sub-fields of perception, concept formation, and inference or scientific method. While perception and concept formation have been researched by Objectivist writers such as Rand (1990), Peikoff (1991) and Kelley (1986), the field of inference has received relatively little attention from them (see Kelley 1990a). Rand relied on others to provide a rigorous scientific method.

Inference is the domain of philosophers of science, who attempt to construct normative methodological frameworks for scientific inference. By and large, these philosophers are ontological realists and assume reliable, but not perfect, faculties of sense perception and concept formation. The question of the degree of objectivity normative inference possesses is of prime concern to Objectivist philosophers. Over the past half century, there has been a trend in this field toward more and more subjective methodologies. Objectiv-

The Journal of Ayn Rand Studies 3, no. 1 (Fall 2001): 29–46.

ism, which relies so heavily on the scientific method for concept validation, has no use for a method that is inherently subjective. A recent school of thought based on Error Statistics appears to restore objectivity to scientific method. It is the thesis of this article that Error Statistics-based inference is objective and that it meets the desiderata of a normative methodology for scientific inference, and thus meets a necessary condition for inclusion in Objectivism.

Scientific Method

Simply stated, the scientific method is a procedure for establishing the truth value of a hypothesis. In the simple case, there is a theory to be tested and data are used in the performance of this test. The *method* is a procedure used to perform this test and to decide whether the theory is verified or falsified. The theory under test is called the primary hypothesis.[1] There may or may not be an alternative that is also being tested. A third type of hypothesis is the auxiliary hypothesis, which serves to *connect* the primary hypothesis to observable data. These auxiliary hypotheses are often implicit and typically subsume common assumptions and previously well-verified theories that the scientist wants to build upon. Auxiliary hypotheses are frequently referred to as *ceteris paribus* conditions.

At this point, it is important to make a distinction between normative and descriptive models of scientific inference. A descriptive model attempts to describe how science *is* performed. This is predominately the domain of more historically and sociologically oriented philosophers of science. A normative model attempts to describe how science *ought to be* performed, using some optimal or ideal methodology. The study of normative methods is strictly an epistemological endeavor.[2]

The desiderata of a normative scientific method are summarized in Figure 1 below. Such a method must be both accurate and informative. With respect to accuracy, the test must have a high probability of being passed if the hypothesis is true and a low probability if the hypothesis is false.

Given the test is accurate, it must also be informative. If the

Figure 1. Summary of Requirements for a Normative Scientific Method

Outcome of Test		Truth Value of Primary + Auxiliary Hypothesis: True	Truth Value of Primary + Auxiliary Hypothesis: False
	Pass	High Probability (Must Solve MUD)	Low Probability (Type II Error)
	Fail	Low Probability (Type I Error)	High Probability (Must Solve Duhem-Quine Problem)

hypothesis is falsified, deductive logic tells us that either the primary hypothesis is false, or one of the auxiliary hypotheses is false (or both); this is known as the Duhem-Quine problem. A normative method must be able to determine which of these conditions is the case. A quick example is in order. When the Newtonian theory of gravity was being tested, the observed orbit of the planet Uranus was not what was predicted by the theory. Either the primary hypothesis was wrong (Newtonian Gravity) or one of the *ceteris paribus* conditions was wrong (e.g., there were no other unobserved planets perturbing Uranus's orbit). In fact, the theory was correct, and the previously undiscovered planet, Neptune, was responsible for Uranus's deviation from the predicted orbit.

If a primary hypothesis is verified, a normative methodology must be able to choose among other verified (but contrary) alternative hypotheses. If hypothesis tests are purely binary (pass-fail), then we can potentially be left with many conflicting theories that have passed, which is not terribly informative. This is known as the problem of Methodological Underdetermination.

A general and overarching desideratum is that the method must be applicable to both statistical and non-statistical (deterministic) hypotheses.[3] Deterministic and statistical hypotheses are best described by example. An example of a deterministic hypothesis is: "All ravens are black." This is a universal generalization. It can be stated in the logical form: If A is a raven, then A is black. An example of a statistical hypothesis is: "Drug A is more effective at treating headaches than a placebo." This is a statistical generalization about a population rather than a statement about particular subjects in a sample. Deterministic hypotheses are more common in the physical sciences such as physics. Statistical hypotheses are more common in clinical trials and in the socio-behavioral sciences. Fields such as biology routinely deal with both types. Finally, the Objectivist will add *at least* one more overarching requirement: the method must be objective.

One of Objectivism's important contributions to philosophy is the identification of "the differences separating an intrinsicist, a subjectivist and an objectivist approach to epistemology" (Peikoff 1991, 150). Intrinsicism holds that a concept is somehow inherent in reality, independent of human consciousness. Specifically regarding epistemology, "subjectivists hold that a man need not concern himself with the facts of reality; instead, to arrive at knowledge or truth, he need merely turn his attention inward, consulting the appropriate contents of consciousness, the ones with the power to make reality conform to their dictates" (Peikoff 1982, 58). Epistemological objectivity is defined by Rand as "the recognition of the fact that a perceiver's (man's) consciousness must acquire knowledge of reality by certain means (reason) in accordance with certain rules (logic)" (Rand 1965, 7). Another definition of objectivity that can be useful in assessing scientific method comes from Kelley: "a commitment to thinking in accordance with the facts and interpreting them logically, without bias or prejudice (Kelley 1990b, 110).

Before describing the Error Statistics methodology, several attempts at developing *normative* methodologies over the last several decades will be summarized. This will serve both to describe rival methodologies and to assess their ability to address the above desiderata.

Review of Previous Methods

Perhaps a natural starting point for such an overview is a school of epistemology known as Logical Positivism. Developed by a group of scientists, mathematicians, and philosophers (known as the Vienna Circle) in the 1920s and 1930s, this school of thought emphasized the definition and classification of epistemological terms rather than formulating a method *per se*.

Building upon the foundation of the Logical Positivists, the Logical Empiricists, particularly Rudolf Carnap (1962), attempted to develop a "Logic of Confirmation." This was a method to quantify a value C(h,e), which was a logical degree of implication C that evidence (e) affords to a hypothesis (h). While the Logical Empiricists never successfully developed their model of a normative method, this type of evidential-relationship approach was later revived by the Bayesians, as we shall see.

The first reasonably successful attempt at a normative methodology was the "critical rationalism" methodology of Karl Popper (1959), which was popularized in the late 1950s. In testing deterministic-universal hypotheses, Popper observed that hypotheses can be proven wrong with falsifying evidence by applying the logical rule of *modus tollens*. This states: If A then B; Not B, then Not A. In other words, if the hypothesis is true, then datum X will be observed; if datum X is not observed, the hypothesis is *not* true. However, when experimental evidence is consistent with the hypothesis, those who believe that this confirms the hypothesis fall prey to the "affirming the consequent" fallacy. Hence, according to Popper, scientific theories cannot be verified, but only falsified. Hypotheses that are not falsified are only tentatively corroborated. Thus, Popper seemed to have obviated Hume's "Problem of Induction" using this purely deductive and impeccably objective methodology.[4]

While this conceptual framework was well received by many physical scientists, philosophers quickly (and appropriately) judged this method against the desiderata above (Earman 1992). Its primary problem with respect to accuracy is its inability to handle experimental error; this can result in true hypotheses being falsified or Type I Errors. Another shortcoming is its treatment of the Duhem-Quine problem; the method does not adequately address how primary and auxiliary hypotheses might be tested independently. Finally, Popper's methodology is not very helpful in choosing between alternative hypotheses that are not falsified. Central to Popper's method is that hypotheses cannot be verified. All hypotheses that pass tests attain the equal status of *corroborated*, which is a particularly egregious example of Methodological Underdetermination. Although Popper's method was developed for deterministic hypotheses, its inability to handle statistical hypotheses limits its applicability as a general normative model.

In addition to the above problems, Popper's method (in addition to any other purely deductive method) suffers from what is known as the Raven Paradox. Consider the deterministic hypothesis: "All ravens are black." By applying the rules of deductive logic, this statement is equivalent to: "All non-black things are non-ravens." Thus, while observing a non-black thing that is a raven falsifies both hypotheses equally, seeing a white swan corroborates both hypotheses equally as well as the observation of a black raven.[5] Thus, while Popper's method was thoroughly objective, it failed several major "tests" as a comprehensive and normative method of inference.

The next major normative model to gain prominence was based on applying the principles of Bayesian Statistics to the problem of general scientific inference. The field of statistics saw a resurgence of Bayesian thought in the 1960s and 1970s; this movement was pioneered by statisticians such as Bruno DeFinetti, Denis Lindley and L. J. Savage. Ideas from these statisticians and a number of Bayesian epistemologists were synthesized into a general normative model by Howson and Urbach in their excellent book, *Scientific Reasoning: The Bayesian Approach* (Howson 1989). A brief summary of this method follows.

Let P(h) be the *a priori* measure of belief in the hypothesis h before the experiment. P(e) is the probability we would observe evidence, e, regardless of whether h is true or not; this term is often called the expectedness of the evidence. P(e|h) is the probability of observing the evidence given h is true (this conditional probability is known as a likelihood), and P(h|e) is the *a posteriori* probability of the hypothesis given the evidence. These terms are related by Bayes' theorem:

$$P(h|e) = P(e|h)\,P(h)/P(e)$$

Howson and Urbach adopt the following definitions:

e CONFIRMS or SUPPORTS h when $P(h|e) > P(h)$
e DISCONFIRMS or UNDERMINES h when $P(h|e) < P(h)$
e is NEUTRAL WITH RESPECT TO h when $P(h|e) = P(h)$

The relationship among the various terms as specified by Bayes' theorem readily appeals to the scientist. The posterior probability increases with the likelihood of the data given the hypothesis. When $P(e|h)=0$, disconfirmation is maximized. For deterministic hypotheses, h logically entails e resulting in maximum support.

The degree that the evidence supports h also increases with the prior probability. This also intuitively appeals to scientists, as it reflects all *circumstantial* evidence known ahead of time. Also, P(h|e) increases with decreasing P(e); this means that unexpected experimental results give added support to the theory under test.

While originally formulated for purely statistical hypotheses, the Bayesian model also easily applies to deterministic hypotheses. For illustration, we will again make use of the Raven example. To test the hypothesis "All Ravens are Black," the P(h) is assigned based on prior knowledge. For deterministic hypotheses such as this, $P(e|h)=1$; given the hypothesis is true, the probability of a raven being black is unity. Experimental uncertainty is easily included in this framework by allowing P(e | h) to be less than one, depending on the reliability of the experiment.[6] Popper's falsification method could not adequately account for experimental error. This model's applicability to statistical hypotheses is well known and will not be addressed further (Gelman 1995).

Accuracy issues are adequately addressed since Type I and Type II errors are readily computable. Bayesian methods easily handle the problem of Methodological Underdetermination. Once tested, the choice between the primary and potentially many alternative hypotheses is made by whichever one has the highest posterior probability. Thus, a primary hypothesis can be tested by itself (by comparing the posterior to the prior as specified by Howson and Urbach) or it can be tested against an arbitrary number of alternatives.

Bayesian methods *appear* to solve the Duhem-Quine problem by performing separate analyses on the primary (T) and set of auxiliary (A) hypotheses if the original tests fails. For example, if P(A | e) > P(A) and P (T | e) < P(T) then we conclude that the primary hypothesis is wrong and the auxiliary hypotheses are correct. While it has been demonstrated that it is possible, *post hoc*, to assign subjective probability values to fit simple historical examples (Dorling 1979), it can be quite difficult to execute this procedure in practice.

The Bayesian response to the Raven Paradox is a simple statistical one. Assume that there are some non-black ravens. A random sample of ravens is more likely to discover a non-black raven than is a random sample of non-black objects. Thus, using Bayes' Theorem, observing black ravens raises the posterior probability more than observing white swans due to the different resulting values in the expectedness of the data, P(e) (Suppes 1966).

While some philosophers have found faults with the Bayesian model as a normative method, it has become the dominant school of thought in this field.[7] One such fault is the *Problem of Old Evidence*. Since P(e) is a subjective degree of belief, if a scientist knows for a fact that e is the case, then he must assign the values P(e) = P(e | h) = 1 (Glymour 1980). However, using Bayes' theorem, this means that P(h | e) = P(h); and using Howson and Urbach's criteria, this corresponds to the hypothesis being neither confirmed nor falsified.

Another criticism came from Popper, which is referred to as the *Problem of Non-Zero Priors*. According to Popper (1959), in an "infinite universe," the prior probability of any (non-tautological) hypothesis is zero. In this case, P(h) = 0 and therefore P(h) = P(h | e), rendering any evidence neutral to all hypotheses.

To the Objectivist, the Bayesian model's major fault is its reliance on subjective prior probabilities.[8] Quite often, there is no objective method of determining the term P(h). In practice, this term ends up signifying a subjective *degree of belief* rather than an objective *frequentist* probability. So while this method appears much more useful than Popper's falsification, it is ultimately based on a subjective degree of belief as a starting point. Despite this drawback, the lack of any comprehensive, competing framework, coupled with a resignation to methodological subjectivity has resulted in Bayesian inference becoming the dominant school for normative inference.

This resignation to methodological subjectivity may, in part, be due to the work of Thomas Kuhn (1962) described in his book, *The Structure of Scientific Revolutions*. Kuhn developed a descriptive model for scientific change using his concept of paradigms. According to Kuhn's model, science is essentially a puzzle-solving endeavor. Progress is made during periods of "normal science" by using paradigms, which are networks of widely accepted theories, experimental procedures, and background knowledge. Occasionally, when paradigms lose their problem-solving ability, a scientific revolution simultaneously replaces the old paradigm by a new one. According to Kuhn, paradigm replacement does not happen gradually, but suddenly and completely. The new paradigm is not only a complete replacement, but the two are "incommensurable," meaning that scientific concepts from rival paradigms are not inter-translatable. Because of this incommensurability, the dynamics of paradigm shift are not and cannot be driven by objective scientific evidence. Rather, the dynamics of paradigm shift are mainly driven by social factors and, thus, are inherently subjective.

While Kuhn's analysis focused on actual or descriptive science, it is often misinterpreted as relating to normative methods of inference as well. This distinction between the descriptive and normative aspects of Kuhn's work has been the cause of great confusion and empty debate over the years, even though he was quite clear as to the scope of his own work: "many of my generalizations are about the sociology or social psychology of scientists; yet at least *a few* of my conclusions belong traditionally to logic or epistemology" (Kuhn 1962, 8; emphasis added). One of those few conclusions that some philosophers believe applies to *normative* methods is that for incommensurable (i.e., rival) theories, there are no objective, theory-neutral principles that can be used to compare them (Newton-Smith 1981, 109).

Error Statistical Approach to Scientific Method

This condition remained effectively unchallenged until recently, when a number of philosophers, such as Allen Franklin, Ian Hacking and Peter Galison, started to focus on the role of experiments in scientific inquiry. Looking at hypothesis testing from an experimental rather than a theory-laden perspective was seen to address many of the problems mentioned above. Galison (1987, 259) observed: "Experimental conclusions have a stubbornness not easily canceled by theory change." Despite this key insight, there lacked an integrated methodology that could replace that of the Bayesians.

In 1996, Deborah Mayo provided such a comprehensive and integrated methodology in her book, *Error and the Growth of Experimental Knowledge*, with what she calls an *Error-Statistical* approach to scientific inference. Mayo's framework is characterized as a *testing approach* to inference, which is fundamentally different from the *evidential-relationship* approaches of Popper and the Bayesians. It starts with a hierarchical Philosophy of Experiment. She analyses the connection between theory and experiment in order to better understand how and why valid theories can fail tests (when *ceteris paribus* conditions are not met) and how to distinguish among different valid theories.

Experimental inquiry is represented by a network of hierarchical models: Primary Models, Experimental Models and Data Models (Suppes 1969, 24–35). The Primary Model establishes the relationship between the theory under test and a testable hypothesis. This testable hypothesis is linked to observable data with the Experimental Model. Data Models address how to generate and model raw data such that it can be linked to experimental models. The last component of this model network is a "toolbox" of methods for testing the validity of these models and their assumptions (common examples are: normality, independence, variance homogeneity and random sampling). This network of models provides a philosophical basis and generalization of the statistical concept of *Experimental Design*. It enables the experimenter to make decisions such as: choice of test statistic, variable selection, sample size, and randomization technique.

A key ingredient to Mayo's methodology is the concept of severe tests, which is very powerful and widely applicable. Severity is defined as follows:

$$P(\text{Fail} \mid \sim h) \text{ is High}$$
$$P(\text{Pass} \mid \sim h) \text{ is Low}$$

Experiments are designed to ensure that there is a high probability of passing the test if the hypothesis is true and a low probability of accepting it if false, in other words, severe. This can be thought of as a generalization of the statistical concept of a test that simultaneously has a low probability of Type I and Type II errors.

Once the experiment has been designed and executed, it is analyzed using well-known Neyman-Pearson methods (Hogg 1995). These methods determine the optimal critical value of the test statistic and quantify error probabilities of committing Type I and Type II errors—hence, addressing the accuracy requirements.

With respect to the method's informational content, Methodological Underdetermination is handled by choosing whichever hypothesis passes the more severe test. The Duhem-Quine problem is solved by independently testing the primary and auxiliary hypotheses. While this may appear *prima facie* similar to the Bayesian solution, it is actually quite different. Under the Error Statistics methodology, the primary and auxiliary hypotheses are tested independently by designing separate severe tests into the experiment, rather than *post* *hoc* assignment of separate subjective probabilities. This is accomplished in practice by appropriate variable control and randomization. Also, the experimenter determines *a priori*, what errors would be expected if either the primary and auxiliary hypotheses are false, and designs the experiment to detect these errors if present. Mayo calls this creating an "error repertoire."

Deterministic hypotheses are brought into this framework by the observation that in testing such hypotheses, inevitable approximations, inaccuracies and uncertainties enable the application of standard statistical tests. Mayo also effectively uses the concept of severe tests to shed light on the other important epistemological problems such as *ad hoc* or use-constructed hypotheses and the utility of novel evidence. Severity also illuminates the varying information content of true experiments versus quasi-experiments versus observational studies; this is of particular use to the socio-behavioral sciences (Pedhazur 1991).

To further bolster the case for Error Statistics as a normative method (while not digressing too much), an outline is presented below on how different sub-categories of hypotheses can be tested within this framework.

A common statistical hypothesis is presented: Ho: "z% of X's have property β." An alternative hypothesis is formulated: "z'% of X's have property β." The steps to test this hypothesis are as follows:

1) Design a severe test
   a. Decide on acceptable levels of Type I and Type II errors
   b. Assign a Binomial Model
   c. Compute critical region and sample size n.
2) Execute and analyze experiment: We observe n X's, compute the percentage with observable β and see if it is in the critical region.
3) Check *ceteris paribus* assumptions (sampling, measurement bias, etc.).
4) Choose between primary and alternative hypothesis based on data from severe test.

Universal Hypothesis—Ho: "All X's have property β."

1) Design a severe test
   a. First we note that there is very little probability of a Type I error (Test failing when primary hypothesis is true) except when *ceteris paribus* conditions are violated.
   b. Decide on acceptable level of Type II error.
   c. Assign a Binomial Model and an appropriately small percentage of X's that do not have property β as an alternative hypothesis.
   d. Compute sample size n to achieve the desired level of Type II error.
2) Execute and analyze experiment: we observe n X's (unless we observe an X that does not have property β first).
3) Check *ceteris paribus* assumptions (sampling, measurement bias, etc.).
4) If all X's have property, then we accept Ho with specified levels of Type II error.

Existential Hypothesis—Ho: "There exists a Y that has property β."

1) Design a severe test
   a. First we note that there is very little probability of a Type II error (Test Passing when primary hypothesis is false) except when *ceteris paribus* conditions are violated.
   b. Decide on acceptable level of Type I error.
   c. Assign a Binomial Model and an appropriately small percentage z of Y's that have property β as an alternative hypothesis (Ha: Less then z% of Y's have property β).
   d. Compute sample size n to achieve the desired level of Type I error.
2) Execute and analyze experiment: We observe n Y's (unless we observe a Y with property β first).
3) Check *ceteris paribus* assumptions (sampling, measurement bias, etc.).
4) If no Y's have property then we reject Ho with specified levels of Type I error.

Metaphysical Hypothesis—Ho: "For every X, there exists a Y that has property β."

1) Design a severe test

a. First we note that, like statistical hypotheses, there is structural possibility of both Type I and Type II errors.

b. Decide on acceptable level of Type II error. To do this, we assume that for some small percentage of X's, there does not exist a Y that has observable property β. We assign a binomial model and compute how many X's 'm' we need to observe.

c. Then we decide on an acceptable level of Type I error. This is done by assuming an appropriately small percentage of Y's have observable property β. Again, we assign a Binomial Model and compute how many Y's 'n' we need to observe to achieve the desired level of Type I error.

2) Execute and analyze experiment: we observe 'm' X's and for each X, we observe 'n' Y's (unless we observe property β before we reach n).

3) Check *ceteris paribus* assumptions (sampling, measurement bias, etc.)

4) Choose between primary and alternative hypothesis based on data from severe test.

Thus, for Statistical, Universal and Existential hypotheses, we are only observing one type of entity (at least for the simple examples above). For Universal hypotheses, Type I error is inherently low and Type II error is controlled by observing more objects (plus assumption checking). For Existential hypotheses, Type II error is inherently low and Type I error is controlled by observing more objects (plus assumption checking). For Statistical hypotheses, both Type I and Type II errors are controlled by observing more objects.

For Metaphysical hypotheses, two types of entities are being observed, X and Y. Type II error is controlled by observing more X's and Type I error is reduced by observing more Y's. Thus, Error Statistics enables testing of Metaphysical hypotheses,[9] but significantly more measurements are required (n x m) compared to the other three

to achieve acceptable levels of severity.

This methodology achieves objectivity by using *a priori* knowledge in an objective way by designing experiments that result in severe tests. Specifically, this means making such decisions as: choice of variables, test statistic, sample size, randomization method and techniques for *post hoc* assumption checking. In response to Kuhn, Mayo shows that empirical data from these carefully designed and controlled experiments can bridge Kuhn's incommensurability gap. Mayo's Error-Statistical approach nicely meets the desiderata of a normative methodology of inference: accurate, informative, and applicable to all types of hypotheses.

Finally, a few comments on the descriptive aspects of the Error Statistics methodology. For statistical hypotheses, Neyman-Pearson type tests are currently the dominant school of thought, indicating that these methods are both useful and consistent with the needs of scientists in statistically oriented fields of research. As for non-statistical hypotheses, Mayo (1996) gives a descriptive account of how various historical examples are consistent with this methodology as well.

Summary

In summary, while Popper's method was thoroughly objective, it was not accurate or informative enough to be very useful for theory choice. The Bayesians responded with a method that addressed many of the shortcomings of Popper; however, this came at the price of methodological objectivity. This price was perceived to be minimal, since Kuhn claimed that no method could be objective due to the incommensurability of rival theories. Mayo's Error-Statistical approach successfully refutes this notion and provides an elegant theory of inference that is simultaneously objective, accurate and informative, and hence qualifies as a normative methodology. And because of this, it becomes a candidate for filling a long-standing void in Objectivist philosophy.

Additional research is required to answer the question of whether Error Statistics *can* be integrated into Objectivist philosophy. Specific questions are:

1) Are there other necessary conditions that must be met besides objectivity?
2) Are these conditions met?
3) What is the set of sufficient conditions?
4) Are these conditions met?
5) If Error Statistics does not meet these conditions, what will?

Acknowledgments

The author wishes to acknowledge valuable contributions to this essay by the referee.

Notes

1. In this essay, the terms "primary hypothesis" and "theory" will be used interchangeably.

2. It is worth examining an example of the demarcation between normative and descriptive. If a descriptive study of ethics determines that most people do not practice Rand's Objectivist (normative) ethics, this is not a valid argument against Rand's position. However, a careful study of descriptive ethics *can* be used to, at least partially, validate a normative theory. Thus, philosophical studies must precisely delineate between normative and descriptive aspects, and comparisons between the two must be made judiciously. The same applies to scientific method.

3. Both statistical and deterministic hypotheses are further classified into different types. A common type of statistical hypothesis is of the form "z% of X's have property β." Non-statistical hypotheses have been categorized by Popper (1959) into three types: Universal, Existential and Metaphysical. A Universal hypothesis has the form: "All X's have property β." An Existential hypothesis has the form: "There exists a Y that has property β." Finally, Metaphysical hypotheses can be thought of as a combination of these two: "For every X, there exists a Y that has property β." Clearly any normative model must be capable of addressing these various types of hypotheses.

4. While Popper's method was certainly not subjective, it could be considered intrinsicist under the Objectivist trinomial categorization scheme.

5. The main lesson from the Raven Paradox is that purely deductive models for scientific method cannot be normative.

6. For an illustrative example of how to assign values to the various terms in Bayes' Theorem, see Dorling 1979, 177–87.

7. It is important to differentiate between criticisms of the Bayesian Model as a: 1) purely statistical method; 2) descriptive model of scientific inference; and 3) normative model for scientific inference. The magnitude of criticism seems to scale with the viability of alternative methods. Thus, there has been the most criticism for applications 1 and 2 and somewhat less for application 3. This essay will focus exclusively on the Bayesian Model as a normative method.

8. An Objectivist might argue that the Bayesians' dogmatic insistence on using Bayes' Theorem here has intrinsicist elements, particularly in the Howson and Urbach formulation. However, their reliance and personal degrees of belief for the prior probability is clearly subjectivist.

9. Early normative methodologies like Popper's could not test Metaphysical hypotheses.

References

Carnap, Rudolph. 1962. *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Dorling, John. 1979. Bayesian personalism, the methodology of scientific research programmes and Duhem's problem. *Studies in the History and Philosophy of Science* 10, 177–87.

Earman, John and Wesley Salmon. 1992. The confirmation of scientific hypotheses. In *Philosophy of Science*, edited by Merrilee H. Salmon, John Earman, Clark Glymour, James G. Lennox, Peter Machamer, J. E. McGuire, John D. Norton, Wesley C. Salmon, Kenneth F. Schaffner. Indianapolis, Indiana: Hackett.

Galison, Peter. 1987. *How Experiments End*. Chicago: University of Chicago Press.

Gelman, Andrew, John Carlin, Hal Stern, and Donald Rubin. 1995. *Bayesian Data Analysis*. London: Chapman and Hall.

Glymour, Clark. 1980. *Theory and Evidence*. Princeton: Princeton University Press.

Hogg, Robert and Allen Craig. 1995. *Introduction to Mathematical Statistics*. Upper Saddle River, New Jersey: Prentice Hall.

Howson, Colin and Peter Urbach. 1989. *Scientific Reasoning: The Bayesian Approach*. La Salle, Illinois: Open Court.

Kelley, David. 1986. *Evidence of the Senses: A Realist Theory of Perception*. Baton Rouge: Louisiana State University Press.

___. 1990a. *Universals and Induction*. Audiotape lecture. Poughkeepsie: Institute for Objectivist Studies [now The Objectivist Center].

___. 1990b. *The Art of Reasoning*. New York: W. W. Norton.

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

Mayo, Deborah. 1996. *Error and the Growth of Experimental Knowledge*. Chicago: University of Chicago Press.

Newton-Smith, W. 1981. *The Rationality of Science*. London: Routledge.

Pedhazur, Elazar, and Liora Pedhazur. 1991. *Measurement, Design and Analysis*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Peikoff, Leonard. 1982. *The Ominous Parallels: The End of Freedom in America*. New York: Stein and Day.

___. 1991. *Objectivism: The Philosophy of Ayn Rand*. New York: Dutton.

Popper, Karl. 1959. *The Logic of Scientific Discovery*. New York: Basic Books.

Rand, Ayn. 1965. Who is the final authority on ethics? *The Objectivist Newsletter* 4, no. 2 (February): 7–8.

___. 1990. *Introduction to Objectivist Epistemology*. Expanded 2nd edition. New York:

Meridian.

Suppes, Patrick. 1966. A Bayesian approach to the paradoxes of confirmation. In *Aspects of Inductive Logic*, edited by Jaakko Hintikka and Patrick Suppes. Amsterdam, The Netherlands: North Holland.

___. 1969. Models of data. *Studies in the Methodology and Foundations of Science*, 24–35. Dordrecht, The Netherlands: D. Reidel.