# LOGIC AND SCIENCE IN WILDLIFE BIOLOGY

BYRON K. WILLIAMS, North American Waterfowl and Wetlands Office, U.S. Fish and Wildlife Service, 4401 North Fairfax Drive, Arlington, VA 22203, USA

*Abstract:* Descriptions of wildlife science often inadequately treat the interrelation of theory, formal logic, and data analysis that define scientific methods in the investigation of causal hypotheses. I provide 2 definitions for causation, and describe scientific methods in terms of theory, causal hypotheses, and the comparison of testable predictions against experimental and/or field data. Hypothesis assessment is developed with principles of logical inference, emphasizing differences in the burden of proof for hypothesis confirmation and rejection. Inductive logic and statistical inference are emphasized in the identification and testing of hypotheses. This framework is illustrated by means of the comparison of additive and compensatory hypotheses relating harvest and annual mortality.

***J. WILDL. MANAGE. 61(4):1007–1015***

**Key words:** cause, hypothesis testing, inference, logic, scientific method, statistical inference, wildlife science.

However varied the practice of wildlife science, a common feature is the comparison of predictions, deduced from biological hypotheses, with data collected pursuant to the comparison. Much has been written about the testing of biological/ecological hypotheses (Romesburg 1981, Hurlbert 1984, Peters 1991), and specifically about sampling designs and statistical inferences for hypothesis testing (Green 1979, Hairston 1989, Skalski and Robson 1992). However, much of this documentation has focused on characterizing biological hypotheses in terms of statistical distributions, and on investigating distribution attributes with sample data Brownie et al. 1985, Burnham et al. 1987, Lebreton et al. 1992). Left undeveloped is an exposition of how these activities fit into a broader context of theory, logic, and data analysis that is definitive of scientific methodology. Thus, the objective of this paper is to provide a biological context for scientific methodology, and in so doing to clarify the respective roles of theory development, statistical inference, and the structures of formal logic in wildlife science.

## Causation in Wildlife Science

Whatever else it is, science is about the identification and confirmation of causes for observed phenomena, where by "cause" is meant a generic explanation of patterns observed for a class of phenomena. The explanatory power of a cause results from the ability to entail many, often apparently disparate, phenomena under its rubric. Causes are recognized as "explanatory" in the context of a scientific theory of which they are components, the theory itself consisting of relatively few causal factors entailing a wide range of phenomena.

More formally, causation can be described in terms of antecedent conditions, consequent effects, and a rule of correspondence for their conjoint occurrence. In wildlife biology the "effect" of a cause typically is a biological event (e.g., mortality, growth, population change) that occurs subsequent to the occurrence of some prior condition. Provided the joint occurrence of the prior condition and the subsequent event meet certain theoretical and logical requirements, the prior condition is held to be the cause of the event.

The causal linkage between a prior condition and a subsequent effect can be described in terms of the logic of material implication (Copi 1982). The expression $A \rightarrow B$ describing material implication is taken to mean that affirmation of the premiss $A$ implies affirmation of the conclusion $B$. However, material implication is silent about the affirmation of $A$ given that $B$ is affirmed. More formally, material implication establishes the equivalence of $A \rightarrow B$ with the assertion that either $A$ is false or $B$ is true. Thus, one can look to the premiss of $A \rightarrow B$ to confirm its conclusion, but one cannot look to the conclusion of $A \rightarrow B$ to confirm its premiss.

The concept of causation in scientific enquiry is informed by the logic of material implication, by identifying cause ($C$) and effect ($E$) as either premiss or conclusion. Two distinct definitions of causation can be identified.

*Necessary Causation:* $E \rightarrow C$.—In this case an effect $E$ points to a presumptive cause $C$, in that the occurrence of the effect guarantees the occurrence of condition *C*. A logically equivalent argument is $\sim C \rightarrow \sim E$, i.e., the non-occurrence of *C* guarantees the non-occurrence of the effect (the symbol $\sim$ in this expression is used to indicate logical negation, so that $\sim C$, which is read "not *C*," means that the truth of *C* is negated). Thus, necessary causation asserts that the absence of an effect follows from the absence of the cause. However, it is silent about effect *E* in the presence of *C*. Examples of necessary causation might include: light as a cause of photosynthesis; smoking as a cause of certain forms of lung cancer; ingestion of lead shot as a cause of lead toxicosis in waterfowl; grazing by domestic herbivores as a cause of rangeland degradation; and salmonella bacteria as a cause of typhoid fever. In each example the effect may or may not be present when the presumptive cause occurs; however, the effect is held to be absent when the cause is missing.

*Sufficient Causation:* $C \rightarrow E$.—In this case the presumptive cause *C* points to the effect *E*, in that the occurrence of condition *C* guarantees the occurrence of the effect. Thus, sufficient causation asserts that the occurrence of an effect follows from the presence of condition *C*. However, it is silent about the effect in the absence of *C*. Sufficient causation might underlie an argument that heat causes fluid dynamics; that a low level of ambient oxygen during respiration causes the production of lactic acid; that oxygenation of pig iron under high pressure causes the production of steel; that drought causes physiological stress in non-succulent plants; that competition causes population size to be reduced to levels below an environmental carrying capacity. In these examples the presence of the cause is held to ensure the presence of the effect; however, the effect may or may not be present in the absence of the cause.

Sufficient causation is a logically stronger definition than necessary causation, in that necessary causation specifies *C* as one condition (possibly among many) that must be present to ensure the occurrence of an effect, whereas sufficient causation specifies that *C* alone ensures its occurrence. An otherwise necessary cause can be recognized as sufficient by restricting the range of conditions in which it is operative. Thus, a concentrated source of heat (e.g., a lighted match) is a necessary cause of combustion, but a heat source in the presence of combustible material in a cool, dry oxygenated environment becomes a sufficient cause (under these conditions).

The importance of maintaining a clear distinction between necessary and sufficient causation can be illustrated by the controversy about smoking as a potential cause of lung cancer. Advocates for restricting the advertisement and sale of tobacco products base their arguments on the strong statistical association between tobacco use and the occurrence of lung cancer, wherein the great majority of lung cancer victims in the United States also have a history of smoking. On the other hand, opponents of tobacco restrictions have argued repeatedly that the association between smoking and lung cancer is not causal, and cite as evidence the fact that a majority of smokers in the United States do not have lung cancer. Clearly, these conflicting positions (and different assessments of evidence) point to inconsistent uses of the concept of causation. Apparently advocates of tobacco restrictions assume necessary causation, such that a history of tobacco use is inferred from the occurrence of lung cancer. Evidence for smoking as a necessary cause of lung cancer focuses on the fact that lung cancer victims overwhelmingly have a history of smoking, and a key implication is that the avoidance of smoking implies the near absence of lung cancer. On the other hand, opponents of tobacco restrictions appear to use sufficient causation, wherein smoking should lead to the occurrence of cancer. By implication, the absence of cancer therefore should imply the absence of smoking, which is inconsistent with the fact that the overwhelming proportion of smokers have no record of lung cancer. Hence, tobacco is held not to be a cause of lung cancer by opponents of tobacco restrictions. Given the inconsistent uses of causation, it is not surprising that the controversy between advocates and opponents of tobacco restrictions has not been amenable to data-based resolution. Indeed, the evidence likely will continue to indicate that tobacco use is simultaneously a cause of lung cancer (in the necessary sense) and not a cause of lung cancer (in the sufficient sense).

## Approaches to the Investigation of Causes in Wildlife Science

A definition of cause as necessary often applies to the control of unwanted effects, whereby the elimination of an effect (e.g., typhoid fever) is assured by the elimination of the cause (e.g., destruction of salmonella bacilli through sterilization). Scientific investigation thus involves a search for conditions that are predictive of the non-occurrence of an effect of concern. Necessary causation often is implied in wildlife biology when biological effects in the presence of a particular condition are attenuated by the restriction or removal of the condition. A particular example is duck nest predation as a presumptive (necessary) cause of reproductive failure in cultivated prairie lands under non-drought conditions. The implication is that reducing predation will reduce reproductive failure.

On the other hand, a definition of cause as sufficient applies to causes (e.g., drought) that guarantee an effect (e.g., physiological stress in plants). Scientific investigation in this case involves the search for conditions that are predictive of the occurrence of an effect. Sufficient causation is implied in wildlife biology when the influence of a prior condition is both direct and adequate to produce an effect of concern. A relevant example is the investigation of sport hunting as a potential cause of declining waterfowl population trends. Thus, heavy hunting pressure is hypothesized to reduce survival and depress population levels, recognizing that population declines can occur even in the absence of hunting.

## Scientific Methods

A useful context for scientific method involves scientific investigation both before and during a period when it is guided by a recognized theoretical framework. Thus, in its early stages, scientific activity consists of observation guided primarily by intuition, tradition, guesswork, and perceived pattern. Its function initially is to organize observations into coherent categories, to explore these observations for patterns, and to describe the patterns clearly. The process of recognizing the underlying causes of patterns comes as the scientific discipline matures, and a set of relations is formulated that are accepted as "explanatory." These relations are sometimes called a theoretical paradigm or, more briefly, a theory (Kuhn 1970).

A standard for the operation of science, including wildlife science, involves a comparison of theoretically based predictions against data, recognizing that a match between data and prediction provides evidence of hypothesis confirmation, and the lack of such a match disconfirms a hypothesis (Hempel 1965). A somewhat more detailed treatment includes the following 5 elements:

*Theory*.—First, explicit statement of a relevant theory is necessary, or at least the reference to it is necessary. The theory is expressed in terms of the axioms, postulates, theoretical constructs, and causal relations among constructs that constitute the corpus of the theory. This corpus, involving such biological elements as genetics, taxonomy, evolutionary principles, and ecological relations (Hull 1974), is operationally accepted as verified and true. A theory is noted in what follows by $\{T\}$.

Every scientific discipline is founded on an operational theory, which provides a conceptual framework through which the world is observed and facts about the world are discerned. Broadly recognized examples might include the theory of relativity, electromagnetic field theory, the theory of plate techtonics in geology, thermodynamic theory, and the theory of evolution by natural selection. An operational theory allows one to recognize patterns among apparently disparate phenomena, and to explain relations among them. It also is the foundation for hypothesis formulation, prediction, and testing. In short, a theory is essential to the conduct of scientific investigation.

*Hypotheses*.—Second, a hypothesis that is relevant to the theory is identified, often through field or laboratory observations that appear to be anomalies to the theory, i.e., that appear not to be explained adequately by the theory as it currently is understood. A hypothesis, denoted here by $H$, asserts a claim about relations among components of the theory, or about relations of these components to observed reality, or about relations among entities in the observed world that are presumed to follow from the hypothesis. An example of the first kind of claim might be the recognition that one component of the theory entails another; an example of the second kind is the predicted existence of heretofore unrecognized sociobiological patterns; and an example of the third kind are the dynamics of dispersal following certain kinds of environmental disruption. I emphasize in what follows the investigation of causal hypotheses, involving antecedent conditions and consequent effects that are identified in a theoretical context.

A hypothesis is recognized as potentially true or false. When added to a theory, it renders the theory potentially inconsistent, or potentially false. In what follows an amended theory is designated by $\{T\}+H$, to indicate that $H$ is included as one of the elements defining the amended theory. This notation suggests an attendant increase in theory complexity. Alternatively, $H$ can replace a particular hypothesis $H_0$ within the body of the theory. This is designated by $\{T_0\}+H$, where $\{T_0\}+H_0$ represents the theory before amendment. Scientific investigation then becomes a comparison of the relative explanatory power of the 2 theoretical constructs $\{T_0\}+H_0$ and $\{T_0\}+H$. To simplify notation I will use $\{T\}+H$ to represent both the appending of $H$ to $\{T\}$ and the replacement of $H_0$ in $\{T\}$ by $H$.

*Predictions.*—Third, potentially observable conclusions are deduced from the amended theory. These follow from logical relations inherent in the amended theory, or they are derived from relations between the amended theory and observed reality. The derivation of predictions is designated by $\{T\}+H \rightarrow P$, where $P$ represents a prediction and the arrow indicates logical inference. The notion here is that the addition of $H$ to $\{T\}$ allows for inferences that otherwise would not follow from $\{T\}$ in the absence of $H$. At least some of these inferences are testable, in that they predict observable phenomena that potentially are verifiable with field or experimental data. The key here is that $P$ consists of potentially observable predictions.

*Observations.*—Fourth, field or experimental data are collected that are pertinent to the predictions. The investigator's attention is directed to these data by the amended theory, which is used as above to derive predictions for which the data are relevant. Field and/or experimental data, designated by $O$ for observation, are essential components by which the amended theory is to be evaluated. Key to successful data collection are statistically sound surveys, experiments, and other data collection instruments.

*Comparison of Predictions against Data.*—Fifth, predictions from the amended theory are compared to observations $O$ from the field or laboratory. This comparison is used to determine the acceptability of the amended theory and hence the acceptability of the hypothesis $H$. If $O$ conforms to $P$, i.e, if the predicted results of $\{T\}+H$ are in fact observed, then the investigation provides evidence to confirm $H$. If $O$ does not conform to $P$, then the evidence disconfirms $H$. Statistical testing procedures play a crucial role in the process of hypothesis confirmation.

An ideal approach to scientific investigation consists of repeated applications of this sequence across all levels of investigation. Thus. alternative hypotheses often are part of a study design, wherein 2 or more hypotheses may be considered as alternatives for theory amendment. For a given hypothesis $H$, numerous predictions may be identified, each worthy of field investigation. For each prediction $P$, data from several different field and laboratory studies may be appropriate. In addition, studies involving the same hypothesis, the same prediction, and the same kind of data collection often are repeated numerous times, to add to the strength of evidence for confirmation or disconfirmation.

## Additive and Compensatory Mortality

To illustrate the process of scientific investigation, consider a wildlife species that is exposed annually to sport hunting. A traditional concern in game management is the effect of harvest on future population status, and in particular the effect of harvest on annual survival. Two competing hypotheses have been identified: (1) The hypothesis of additive mortality asserts that harvest is additive to other forms of mortality such as disease and predation. Under this hypothesis the annual mortality rate increases approximately linearly in response to increases in harvest rate. (2) The hypothesis of compensatory mortality asserts that harvest mortality may be compensated by corresponding changes in other sources of mortality. Thus, increases in harvest rate have no effect (up to some critical level $c$ of harvest) on the annual mortality rate. In the standard formulation of the compensatory hypothesis, harvest rates beyond $c$ result in an approximately linear increase in annual mortality. Refer to Anderson and Burnham (1976) and U.S. Fish and Wildlife Service (1988) for a more complete development.

The compensatory and additive hypotheses provide a convenient point of reference for the process of scientific investigation. Research on the effect of hunting is conducted in the context of a theory of population dynamics recognizing structural, functional, and dynamic characteristics of wildlife populations in an ecosystem of interrelated organisms and abiotic processes. Elements of the theory involve reproduction, survival, and migration as influenced by such factors as interspecific interactions, physiological condition, behavioral adaptations, and seasonal habitat conditions. The edifice of concepts, relations, axioms, and terms relating to wildlife populations constitutes the scientific paradigm of population biology (see Baldassarre and Bolen 1994 for rev. of waterfowl population biology). It is in the context of this paradigm that the relation between mortality and harvest rate can be investigated.

The investigation proceeds with deduction of testable predictions, following from inclusion in the paradigm of the compensatory and additive hypotheses. Three general predictions can be recognized for waterfowl populations (Nichols et al. 1984):

1. (1) The compensatory mortality hypothesis leads to a prediction that there is no relation between survival rate and hunting mortality, so long as harvest rate is less than the critical value defined in the hypothesis. On the other hand, the additive mortality hypothesis suggests that there is negative relation between survival rate and hunting mortality over the whole range of potential harvest rates.
2. (2) Under reasonable conditions, the compensatory mortality hypothesis leads to a prediction that there is a negative relation between hunting mortality (during the hunting season) and nonhunting mortality rates (during and after the hunting season). The additive mortality hypothesis leads to a prediction that there is no such relation.
3. (3) The compensatory mortality hypothesis leads to a prediction that there is a positive relation between nonhunting mortality rate and population size or density at some time in the year. In many circumstances nonhunting mortality after the hunting season should be positively related to population size at the end of the hunting season. The additive mortality hypothesis leads to a prediction that there is no relation between nonhunting mortality and population size.

These predictions differ considerably in the degree to which they represent explanatory causes of population dynamics, and the difficulty with which data can be collected and used informatively for testing (Conroy and Krementz 1990). Indeed, it always is an outstanding challenge in scientific investigation to devise ways of collecting data that are pertinent to testable predictions. In this particular case population surveys (Thompson 1992), radiotelemetry (White and Garrott 1990), mark-recapture procedures (Nichols 1992), banding studies (Brownie et al. 1985) and other field procedures can provide valuable data by which to test the predictions. Such studies can be replicated at different times and different locations, under a variety of different field conditions and different harvest strategies, with a focus on one or any combination of the predictions listed above. Each study adds evidence by which investigators can confirm or disconfirm the hypotheses. Replication and redundancy of this kind play an important role in preventing unwarranted generalizations of study results.

## Hypothesis Confirmation

The logic of hypothesis confirmation can be expressed in general terms by means of material implication. The process is denoted by

$$\begin{array}{lll@{\qquad}c@{\qquad}lll}
 & \{T\} + H \rightarrow \mathrm{P} & & & & \{T\} + H \rightarrow \mathrm{P} & \\
 & O \rightarrow {\sim}P & & & & O \rightarrow P & \\
(1) & \{T\} & & \text{or} & (2) & \{T\} & \\
 & \underline{O\qquad\qquad\quad} & & & & \underline{O\qquad\qquad\quad} & \\
 & \qquad\qquad {\sim}H & & & & \qquad\qquad H &
\end{array}$$

In these formulations the first premiss asserts that prediction *P* is a consequent of an amended theory, as described above. The second essentially asserts that *P* is disconfirmed by observation (argument 1) or that *P* is confirmed by observation (argument 2). The third premiss asserts the truth of theory $\{T\}$, and the fourth represents the observed data *O*. A horizontal line separates the argument's premisses and evidence from its conclusion, which is stated on the last line. Again, the symbol ~ in argument (1) is used to indicate logical negation, so that the expression $O \rightarrow {\sim}P$ means "the truth of *O* implies that *P* is false" (i.e., the observation indicates that the prediction is incorrect).

Though the 2 arguments above appear to be analogous in their forms, there is a crucial asymmetry in their logical content. Argument (1) is an example of the syllogistic form *modus tollens* (Copi 1982), wherein rejection of the conclusion in an argument of material implication implies rejection of the premiss:

$$\begin{array}{l} A \rightarrow B \\ \underline{\sim B \qquad} \\ \qquad \sim A \end{array}$$

(Copi 1982). Applying *modus tollens* to the scientific argument above, the observations $O$ do not correspond to what was predicted; thus, $O \rightarrow \sim P$ in the second line of the argument. But $\sim P$ implies $\sim\{T\}+H$ from the first line of the argument, which in turn implies either $\{T\}$ or $H$ (or both) is untrue. Since $\{T\}$ is assumed in the third line of the argument to be a confirmed and operational theory, this leaves the falsity of $H$ as a conclusion of the argument. Hence the conclusion $\sim H$. Simply put, this argument states that evidence contrary to a hypothesis is logically sufficient to disconfirm the hypothesis.

In contrast, argument (2) above has a very different logical content. Here the assertion is of the form

$$\begin{array}{l} A \rightarrow B \\ \underline{B \qquad} \\ \qquad \sim A \end{array}$$

Thus, the evidence $O$ in argument (2) confirms prediction P, the consequent of $\{T\}+H \rightarrow P$. The confirmation of $P$ in turn is held to confirm the amended theory $\{T\}+H$. Since $\{T\}+H$ is held to be true, the component $H$ in particular is presumed to be confirmed. This argument is common in scientific investigation, and dominates most of the research in wildlife biology. Unfortunately, it is logically invalid. Thus, the confirmation of $P$ and the truth of $\{T\}+H \rightarrow P$ cannot be used to assert the truth of $H$. Simply put, evidence supporting a hypothesis is logically insufficient to confirm that hypothesis: other factors than $H$ might well lead to the prediction $P$, independent of the truth or falsity of $H$. The fallacy of affirming the premiss of an implication based on its conclusion is an example of the fallacy of false cause known as affirming the consequent (Copi 1982).

Scientific investigation thus faces an asymmetry in the confirmatory role of experimental or field evidence. On the one hand, a hypothesis can be disconfirmed by evidence contrary to prediction; on the other, a hypothesis cannot be (logically) confirmed by evidence supporting prediction. It is in the context of this asymmetry that scientific hypotheses are held by some to be meaningful only if they are theoretically amenable to disconfirmation (Popper 1959).

The fallacy of false cause can be avoided in argument (2) only if the prediction $P$ can arise in no other way than by the truth of $H$, i.e., only when $P$ and $H$ have the same truth content (if $H$ is true, $P$ is also; if $H$ is false, $P$ is also). Under this much more restrictive condition the proposition $\{T\}+H \rightarrow P$ is replaced by $\{T\}+H \leftrightarrow P$, where the arrow pointing in both directions means that $P$ can serve either as premiss or conclusion in material implication. Thus, to avoid the fallacy of false cause all alternative hypotheses must be eliminated either through experimental design or otherwise must be identified, investigated, and rejected, so that by process of elimination only the hypothesis $H$ remains as an explanation of a confirmed prediction $P$. Hypothesis confirmation through the elimination of alternatives was termed "strong inference" in an important paper by Platt (1964). Although relatively simple in concept, such an approach obviously requires thorough field observations as well as careful analysis to identify and properly examine all reasonable alternative hypotheses.

## Inductive Logic and Statistical Inference

Inductive as well as deductive logic is required for hypothesis confirmation in wildlife science. That inductive logic is an essential feature of scientific enquiry is seen in the identification of hypothesized biological mechanisms, as well as the testing of these hypotheses with data. Indeed, a key activity in scientific enquiry is to identify, from a limited set of observations, hypotheses that explain more than the particular observations giving rise to them. Thus, a limited body of data generates possible explanations for their occurrence, and these are folded as hypotheses into an extant body of theory for elaboration and testing with additional observations. Because any particular set of data constitutes only a subset of all possible observations that could be used, testing procedures are designed to be robust to inherent variation in the evidence.

The formulation and testing of scientific hypotheses, based on only a partial record of potentially relevant observations, renders the practice of science inductive. Simply put, causal mechanisms are asserted to hold for a general class of phenomena, based on examination of only limited observations from that class. The inductive nature of the process inevitably gives rise to the possibility of incorrect inference, and necessitates the conservative rules of scientific and statistical inference that have been developed to accommodate, and protect against, such a possibility.

Statistical procedures are involved in hypothesis testing at the point at which data are collected and subsequently used for comparison against predictions. The principles of survey and experimental design serve to improve the efficiency of data collection, and to ensure that the data are relevant and useful in the investigation of predicted responses. Subsequent to data collection, procedures for statistical inference play a key role in determining whether the predicted responses are supported by the data. A correspondence between data and predictions provides evidence for hypothesis confirmation, and the lack of a correspondence leads to hypothesis rejection.

Statistical testing procedures typically are framed in terms of mutually exclusive and exhaustive "null" and "alternate" hypotheses (Mood et al. 1974). By null hypothesis usually is meant (1) an assertion of extant theory that includes an accepted, sometimes simplified, form of some relevant biological relation, or (2) a biological relation per se, to be considered for replacement in favor of an alternate hypothesis. By alternate hypothesis is usually meant a logically distinct, sometimes more complex, and often more appealing biological relation that potentially can replace a particular null hypothesis.

The mechanics of statistical testing involve the matching of observed evidence against predictions based on the null and alternate hypotheses, with the idea that both hypotheses cannot be true, but one must be. Thus, rejection of the alternate hypothesis leads automatically to acceptance of the null hypothesis as its only alternative. Several benefits accrue to the framing of test procedures in this manner. First, one retains the logical consistency afforded by *modus tollens,* whereby hypothesis rejection is inferred logically from the disconfirmation of a predicted response. For example, a lack of supporting evidence for predictions based on the additive mortality hypothesis leads to its rejection. Second, disconfirmation of predicted responses based on the alternate hypothesis leads automatically to acceptance of the null hypothesis. Thus, the rejection of the additive mortality hypothesis leads to the acceptance of the compensatory mortality hypothesis. Third, confirmation of predictions based on an alternate hypothesis leads automatically to rejection of the null hypothesis, and therefore to acceptance of the alternate hypothesis. Thus, confirmatory evidence for the additive mortality hypothesis leads to rejection of the compensatory mortality hypothesis and to acceptance of the additive hypothesis. In this case the test discriminates cleanly between hypotheses irrespective of test results, and thereby avoids the fallacy of false cause.

Although they appear to be analogous, acceptance/rejection of the null and alternate hypotheses suffer disproportionate burdens of evidence. Indeed, the use of statistical procedures in hypothesis testing expresses an asymmetry that parallels that of syllogistic logic, based on a requirement that evidence must be quite strong to reject a null hypothesis in favor of the alternate. Thus, testing procedures express a scientific conservatism in which amendment of an extant theory, or acceptance of a favored alternate hypothesis, is to be discouraged without strong evidence that it is warranted. In this sense the asymmetry in statistical testing is analogous to that of logical inference, whereby hypothesis confirmation accrues only through a preponderance of evidence, in striking contrast with the relatively modest evidentiary requirements for hypothesis disconfirmation.

## Complementary Hypotheses

Scientific methodology is framed above in terms of hypotheses about alternative mechanisms for an effect of interest, with the idea that only a single hypothesis is operative. Thus, one or more hypothesized mechanisms are considered as potentially explanatory, with repeated use of scientific methodology ultimately identifying the appropriate hypothesis. An underlying assumption is that there is only a single "appropriate" hypothesis, and that other hypothesized mechanisms under consideration will be found to be inadequate through proper use of scientific methodology.

Although this scenario no doubt applies to causal mechanisms in many disciplines, it fails to apply to many interesting problems in wildlife biology and ecology. In fact, wildlife science is replete with examples of complementary factors interacting in complex ways to produce observed effects. For example, it often is less a question of whether interspecific competition, predation, or habitat degradation can be viewed as the cause of declines in a wildlife population, but rather the contribution each factor makes in the declines. In this case all factors may be operating simultaneously, playing important but unequal roles in influencing population dynamics.

That issues involving simultaneous complementary factors arise frequently in wildlife biology is indicative of the complexity of the biological systems under investigation. Physical, ecological, and thermodynamic processes simultaneously influence these systems in a complicated network of interactions among populations and the communities and environments of which they are a part. A natural outgrowth of such complexity is the framing of many scientifically interesting issues about cause and effect in terms of the relative contribution of multiple causal factors (Quinn and Dunham 1984). A useful approach then may involve the estimation of parameters measuring the level of factor influence, based on statistical estimation procedures.

## DISCUSSION

Some researchers believe too much emphasis is placed on hypothesis testing as a signature feature of scientific methodology (Quinn and Dunham 1984, Loehle 1987). Indeed, many scientifically interesting questions, notably those involving the estimation of parameters such as abundance, location, and proportionate influence, can be handled conveniently outside the context of hypothesis testing. However, even in these cases the subject of scientific investigation concerns identification and/or parameterization of theoretically based relations. It is unclear how such relations can be recognized, or how assessed, separate from a foundation of theory.

Hypothesis testing per se is actually the testing of both the hypothesis and the theory it amends. This can be seen in the arguments for hypothesis confirmation presented above. Thus, the rejection of predicted response $P$ leads to rejection of the theory $\{T\}$ as amended by hypothesis $H$. The argument above concluded that because the theory was assumed to be true, the hypothesis was necessarily false. Of course, it is always possible that the theory itself is false and the hypothesis is true (or both are false). Indeed, the history of science contains many examples of accepted theories that were shown eventually to be false (Kuhn 1970). This ambiguity likely is an inevitable consequence of scientific methodology, whereby theories are constantly subjected to amendment and revision through the examination of hypotheses.

Scientific methodology as described above involves theory amendment either by the addition of hypotheses to a theory, or by the replacement of one hypothesis by another. Standard practices of statistical testing fit well with the latter description, primarily because they are framed in terms of the comparison of null and alternate hypotheses. Two exceptions to this framework should be mentioned. First, it sometimes is the case that of 2 hypotheses under consideration, neither is easily recognizable as established, and there is a question about which hypothesis is to be identified as the null hypothesis and which as alternate. The decision is obviously of some operational consequence, because of the differential burden of evidence for null and alternate hypotheses. Under such circumstances nonscientific criteria, involving potential costs and benefits of hypothesis acceptance, often influence the decision. When this occurs it is important to recognize, and acknowledge, that the investigation is guided by objectives that go beyond the objective pursuit of understanding. Testing of the compensatory and additive mortality hypotheses provides a good example, with hypothesis acceptance/rejection strongly influenced according to which hypothesis is identified as null, although neither hypothesis is unambiguously recognizable as null.

A second exception involves multiple comparisons of more than 2 hypotheses. Standard statistical procedures such as likelihood ratio testing do not lend themselves to the testing of multiple hypotheses, except with omnibus test procedures such as analysis of variance (Graybill 1976) or by the comparison of hypotheses taken 2 at a time (Mood et al. 1974). However, in recent years some promising approaches have been identified that allow for the comparison and selection of hypotheses from among multiple candidates. For example, model selection criteria proposed by Akaike (1974) have been used by Burnham and Anderson (1992) and others in the selection of biological relations, and adaptive resource management (Walters 1986) provides a promising approach to the identification of wildlife models from among multiple alternatives (Williams 1996).

A final point of emphasis is that scientific methodology as described above is fully complementary to the traditional goals and objectives of wildlife management. Indeed, many of the presumptive causes of wildlife patterns are recognized from observations made during the course of resource management, and in some instances management itself has been included in designs for their scientific investigation. The linkage between wildlife management and scientific assessment, in which management both supports and is supported by research, is definitive of an adaptive approach to resource management (Walters 1986). Adaptive resource management, in concert with the use of sound scientific methodology, holds great promise for accelerating our understanding of biological processes, while simultaneously improving resource management based on that understanding. In the long term, the melding of research and management may offer the only feasible approach to resolving many long-standing problems that confront wildlife managers.

## LITERATURE CITED

AKAIKE, H. 1974. A new look at the statistical model identification. Inst. Electr. Eng. Trans. Automatic Control 19:716–723.

ANDERSON, D. R., AND K. P. BURNHAM. 1976. Population ecology of the mallard. VI. The effect of exploitation on survival. U.S. Fish and Wildl. Serv. Resour. Publ. 128. 66pp.

BALDASSARRE, G. A., AND E. G. BOLEN. 1994. Waterfowl ecology and management. John Wiley and Sons, New York, N.Y. 609pp.

BROWNIE, C., D. R. ANDERSON, K. P. BURNHAM, AND D. R. ROBSON. 1985. Statistical inference from band recovery data—a handbook. U.S. Fish and Wildl. Serv. Resour. Publ. 2, Washington, D.C. 305pp.

BURNHAM, K. P., AND D. R. ANDERSON. 1992. Data-based selection of an appropriate biological model: the key to modern data analysis. Pages 16–30 *in* D. R. McCullough and R. H. Barrett, eds. Wildlife 2001: populations. Elsevier Appl. Sci., New York, N.Y.

———, D. R. ANDERSON, G. C. WHITE, C. BROWNIE, AND K. P. POLLACK. 1987. Design and analysis of methods for fish survival experiments based on release-recapture. Am. Fish. Soc. Monogr. 5. 437pp.

CONROY, M. J., AND D. G. KREMENTZ. 1990. A review of the evidence for the effects of hunting on American black duck populations. Trans. North Am. Wildl. Nat. Resour. Conf. 55:501–517.

COPI, I. M. 1982. Introduction to logic. Macmillan, New York, N.Y. 604pp.

AYBILL, F. A. 1976. Theory and application of the general linear model. Duxbury Press, Belmont, Calif. 704pp.

EEN, R. H. 1979. Sampling design and statistical methods for environmental biologists. John Wiley and Sons, New York, N.Y. 257pp.

IRSTON, N. G. 1989. Ecological experiments: purpose, design, and execution. Cambridge Univ. Press, New York, N.Y. 370pp.

HEMPEL, C. G. 1965. Aspects of scientific explanation. The Free Press, New York, N.Y. 505pp.

HULL, D. 1974. Philosophy of biological science. Prentice-Hall, Englewood Cliffs, N.J. 148pp.

HURLBERT, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecology 54:187–211.

KUHN, T. S. 1970. The structure of scientific revolution. Univ. Chicago Press, Chicago, Ill. 210pp.

LEBRETON, J.-D., K. P. BURNHAM, J. CLOBERT, AND D. R. ANDERSON. 1992. Modelling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecol. Monogr. 62:67–118.

LOEHLE, C. 1987. Hypothesis testing in ecology: psychological aspects and the importance of theory maturation. Q. Rev. Biol. 62:397–409.

MOOD, A. M., F. A. GRAYBILL, AND D. C. BOES. 1974. Introduction to the theory of statistics. Third ed. McGraw-Hill, Inc., New York, N.Y. 564pp.

NICHOLS, J. D. 1992. Capture-recapture models: using marked animals to study population dynamics. BioScience 42:94–102.

———, M. J. CONROY, D. R. ANDERSON, AND K. P. BURNHAM. 1984. Compensatory mortality in waterfowl populations: a review of the evidence and implications for research and management. Trans. North Am. Wildl. Nat. Resour. Conf. 49: 535–554.

PETERS, R. H. 1991. A critique for ecology. Cambridge Univ. Press, New York, N.Y. 366pp.

PLATT, J. R. 1964. Strong inference. Science 146: 347–353.

POPPER, K. R. 1959. The logic of scientific discovery. Basic Books, New York, N.Y. 479pp.

QUINN, J. F., AND A. E. DUNHAM. 1984. On hypothesis testing in ecology and evolution. Pages 22–37 *in* G. W. Salt, ed. Ecology and evolutionary biology. Univ. Chicago Press, Chicago, Ill.

ROMESBURG, H. C. 1981. Wildlife science: gaining reliable knowledge. J. Wildl. Manage. 45:293–313.

SKALSKI, J. R., AND D. R. ROBSON. 1992. Techniques for wildlife investigations: design and analysis of capture data. Acad. Press, New York, N.Y. 237pp.

THOMPSON, S. K. 1992. Sampling. John Wiley and Sons, New York, N.Y. 343pp.

U.S. FISH AND WILDLIFE SERVICE. 1988. Supplemental environmental impact statement: Issuance of annual regulations permitting the sport harvest of migratory birds. U.S. Fish and Wildl. Serv., Washington, D.C. 340pp.

WALTERS, C. J. 1986. Adaptive management of renewable resources. Macmillan, New York, N.Y. 374pp.

WHITE, G. C., AND R. A. GARROTT. 1990. Analysis of wildlife radio-tracking data. Acad. Press, San Diego, Calif. 372pp.

WILLIAMS, B. K. 1996. Adaptive optimization and the harvest of wildlife populations. Math. Bioscience 136:1–20.

*Received 11 September 1996.*
*Accepted 19 July 1997.*
*Associate Editor: Sinclair.*