SCIENTIFIC METHODOLOGY IN ENTOMOLOGY

RANDY THORNHILL

SYNOPSIS

Considerable research on insects is not directed by evolutionary theory. Fortunately, this is changing. The staggering diversity of insects will be rendered intelligible only by the explanatory and predictive power of the theory of evolution by natural selection. As an evolutionary understanding of insects increases, important clarifications and extensions of the basic theory probably will occur.

Entomologists that employ evolutionary theory should consider the basic nature of the hypothetico-deductive model of science and the four methods for applying it: lab experiments, field experiments, observational analysis and the comparative method. Here I outline the general model of science and the pros and cons of each of the four methods for applying it. I argue that the Popperian view of science is useful because it causes scientists to consider predictions that may lead to elimination of hypotheses. An extreme interpretation of this approach, however, is inappropriate because it ignores the value of positive evidence for understanding nature. All four methods of applying the scientific model can lead to precise understanding of cause and effect in biology, especially when alternative hypotheses with mutually exclusive predictions are carefully considered. Lab and field experiments and observational analysis can provide answers to questions about the nature of selection presently acting on traits of organisms. But only the comparative method can yield answers to questions about evolved function (selective history) of traits.

There are evolutionists who argue that the comparative method is at best a method of obtaining correlative patterns that provide tentative hypotheses that must be subjected to experimental or observational tests in order to determine their value. But regardless of the method, the goal is to discover significant differences between sets of data that allow construction and refinement of correlations. I argue that scientific knowledge actually accumulates and science progresses toward understanding simply as a result of improving correlations between presumed cause and effect.

INTRODUCTION

"For answering questions on function in biology, comparative evidence is more reliable than mathematical reasoning." (Williams 1975, p. 7)
"Only experiments can truly test theory." (Stearns 1976, p. 42)
"The only valid method by which the adaptive significance of a feature can be determined is by direct analysis which includes observing the animal in its natural environment and direct determination of the biological roles and selection forces." (Bock 1977, p. 79)
"There is no fundamental difference between the comparative method and the experimental method in biology." (Alexander 1978, p. 95)
"... observation and comparison are methods in biological research that are fully as scientific and heuristic as the experiment." (Mayr 1982, p. 76)

*Randy Thornhill is an Associate Professor in the Department of Biology at the University of New Mexico. His research focuses on insect (especially Meliponine) and human biology. His book with John Alcock, The Evolution of Insect Mating Systems (Harvard University Press, 1983) provides a modern analysis of the diversity of mating behavior in insects. Current Address: Department of Biology, University of New Mexico, Albuquerque, New Mexico 87131.
What is considered appropriate scientific methodology varies among areas of science, between disciplines of biology, and even among practitioners within a biological discipline. The above quotes by evolutionary biologists reveal that within this field of biology large differences in opinion exist about appropriate methodology. Some critics of the comparative method (e.g., Stearns 1976, Rock 1977, Reznick 1989) argue that experiments represent the only methodology that can yield an understanding of cause and effect. This view also is held by some biologists who are unfamiliar with evolution (e.g., many "laboratory" biologists) and by many physical scientists. On the other hand, some evolutionists state that the comparative and experimental methods are not fundamentally different in quality (e.g., Alexander 1978, 1979, Mayr, 1982) or imply this in their work (e.g., Darwin 1859, 1874, Williams 1975, Maynard Smith 1978, Alcock 1979, Clutton-Brock and Harvey 1979, Mayr 1982, 1983).

The view that the comparative method is inferior stems in large part from lack of recognition that each method has strengths and weaknesses and is consistent with criteria of the general model of science. I will argue in this paper that the comparative method is the only method we have for determination of the selection that has shaped characteristics of organisms (as opposed to the selection presently acting on traits) and it is as precise as any other method in science.

The study of the ultimate meaning of living things was brought into focus by Charles Darwin. Not only did he provide the general theory of life but he also gave us powerful methodology comparative analysis for examining and clarifying the theory. The profundity of Darwin's contributions was incompletely understood until recently. Hamilton's classic papers (1964 I and II) and Williams' book *Adaptation and Natural Selection: A Critique of Current Thinking in Evolution* (1966) had a revolutionary impact on biological investigation. They caused investigators in many biological disciplines to focus their thinking on reproductive competition between individuals and thus return to a Darwinian framework.

Like molecular biology (Lewin 1981) and certain other fields devoted to the study of life, entomology has largely lacked a strong evolutionary foundation. This is changing in molecular biology (e.g., Orgel and Crick 1980, Lewin 1981), and there is indication of such a shift within entomology as a result of efforts by several individuals, most notably Lloyd (see symposia volumes 1970-1982), but there is need for a stronger evolutionary basis in most insect studies. The lack of rigorous hypothesis testing in many areas of entomological research is a reflection of the absence of the powerful theoretical basis that evolution provides. Though I direct this paper primarily toward an entomological audience, the message I am trying to convey is general and applies to all areas of biology that do not presently rely on modern evolutionary analysis for research direction.

Undoubtedly, some of my entomological colleagues will reject my view of the current state of entomology. Such colleagues might point to some experimentally elegant piece of entomological research done in total absence of any understanding of evolution. However, the question is not how elegant research is in terms of control of confounding variables, or how clever the idea behind the work is, but, does the research yield a better general understanding of life? We see over and over again in the history of biology that specific results do not make any general sense until placed in the framework
of evolutionary theory (Dobzhansky 1973, Mayr 1982). An excellent example is the present attempt to couple molecular studies with evolutionary analysis (Lewin 1981). As Dobzhansky (1973) put it, "Nothing in biology makes sense except in the light of evolution."

CAUSATION AND THE GENERAL THEORY OF LIFE

The effects of interest to biologists are the traits or features of organisms. Biologists consider causation of any trait from two perspectives: proximate and evolutionary. Proximate explanations for the existence of biological traits deal with genetic, biochemical, physiological, developmental, social, or other immediate causes leading to the expression of the characteristics. Evolutionary explanations address causes that operated during evolutionary history to lead to present biological phenomena. The evolutionary approach focuses on the relationship between biological characteristics and the selective forces that produced them—that is, the contribution of traits to differential reproduction of individuals in the environments of evolutionary history.

Consider the warning coloration of the monarch butterfly. The colors of adults and juveniles are caused by genetic, biochemical, physiological, and developmental proximate factors. In terms of ultimate causation, the colors probably stem from a history of nonrandom differential reproduction of individuals in the context of avoiding visual predators. The two levels of explanation are complementary, but as discussed below, they are not alternatives.

It has been claimed by some that viewing traits in terms of evolutionary causation solely as the product of selection ignores the roles of agents other than selection in shaping them. For example, drift and mutation cause changes in gene frequencies from generation to generation (evolution). But, relative to selection, mutation and drift are impotent as evolutionary forces, because they act randomly with regard to fitness and thus are unlikely to bring about significant cumulative change (see Alexander 1979 for detailed discussion). Viewing features of life as shaped by selection provides the best general working hypothesis. Any other view is completely impervious to test, as critics of the selectionist approach admit (Lewontin 1978).

The selectionist approach is referred to as the ultimate approach because it is selection that accounts for the existence of proximate mechanisms. Thus, the two forms of causation are not alternatives in any sense. This means that there is really only one approach in biology and that it is the ultimate one. The theory of evolution by selection is not a theory of life, it is the theory of life. Charles Darwin did not invent the idea of selection; he discovered the process, as did Alfred R. Wallace independently at the same time. Selection has acted continuously on all living things throughout the history of life and continues to do so today. Thus selection is omnipotent. The features of life are what they are because of selection in the past, and thus all features of all living things are expected ultimately to promote reproduction or genetic propagation of individuals in evolutionarily relevant environments. This provides the foundation for scientific study of all life and tells the biologist how to proceed in order to gain further understanding of life through experiment, observation, and comparative analysis, regardless
of whether he is interested in molecules, behavior, physiology, morphology, proximate or evolutionary causation, or beetles or human beings.

This is not to say that study of proximate causation is unimportant. A complete understanding of any feature of life includes elucidation of both proximate and ultimate causation, I am saying that the theory of evolution by selection provides the best direction for investigating proximate causation. This approach has recently been used in the study of insect pheromones. Until recently, pheromone structure was approached only from a descriptive perspective, but now ideas are being developed regarding chemical composition of insect pheromones expected on the basis of sexual selection theory (Marshall 1982, Arnold and Houck 1982, Thornhill and Alcock 1983). Selection theory is being used successfully to predict the actual ingredients of insect pheromones (Marshall 1982, in preparation). In the sense I have outlined, the “evolutionary” in evolutionary biology is redundant. Biology is the scientific study of the evolution of life. No other definition of biology makes any sense.

THE GENERAL METHOD OF SCIENCE

Science is a very complicated endeavor, and definitions of science are probably as numerous as scientists. Although it is perhaps a waste of time to attempt a perfect definition of the enterprise, I like Peter Medawar's (1967) definition: the art of the soluble. Most scientists that I am acquainted with feel that things are ultimately knowable, and when we do not know, or only know incompletely, it is because we have not asked the right questions. The hypothetico-deductive method provides the way for scientists to ask questions and seek to answer them. The hypothetico-deductive model consists of the following stages: observation, hypothesis formation, identification of predictions, and testing predictions. The stages are interactive and all stages are creative. One can do science by generation and testing of hypotheses and/or by testing assumptions of hypotheses. These endeavors include locating errors in the observations, hypotheses, and tests of others, and for this reason science is often defined in terms of its repeatability and self-correcting nature (e.g., Simpson 1964).

An observation in the right hand leads to scientific inquiry. Hypotheses begin with speculation, a hunch about the cause of some effect of interest. Scientists should speculate. Without imaginative speculation there would be no hypotheses and thus no direction for seeking understanding (see Lloyd's discussion of speculation in this volume).

Gould (1978) and Lewontin (1978) have argued that evolutionary hypotheses are often ad-hoc (only specific to the trait in question and without sufficient generality to allow testing). They consider evolutionary hypothesizing the art of creating just-so stories the same way Rudyard Kipling did in explaining the leopard's spots, the camel's hump, etc. Ad-hoc arguments do not represent valid speculation. The initial speculations of the working biologist are post-hoc, not ad-hoc, and there is nothing wrong with using post-hoc leads for generating true hypotheses that are general enough for testing. As Clutton-Brock and Harvey (1979) have pointed out, post-hoc explanations represent an inevitable first step in any observational science. The suggestion that monarch butterfly coloration is an anti-predator trait can be used to generate a real hypothesis, which is testable by observational
analysis or experiments dealing with effects of the colors on predator preferences, or by comparative analysis across species of animals with and without warning colors.

For a hypothesis (or theory) to be scientific it must be testable. This means it must be both predictive and empirically falsifiable. The requirements for predictions are that they be logically derived from the hypothesis and be statements about the unknown. It has been argued that evolutionary theory is not truly predictive because it focuses on historical causes and not future events (Peters 1970). But, as numerous people have pointed out, prediction of future events is not a requirement of a scientific theory. Prediction of the unknown, whether past, present, or future, is the important issue.

A hypothesis that predicts everything imaginable is not within the realm of science because it cannot be falsified. That is, a hypothesis that explains everything explains nothing. The theory of evolution by selection erroneously has been called a nonfalsifiable theory because of this (Popper 1994). It is puzzling how Popper and others could read Darwin's work and conclude that the theory of evolution is nonfalsifiable. Darwin's writing reveals his care in identifying observations that would falsify his theory (for discussion see Ghiselin 1969, Alexander 1977). Popper (1984) is usually given credit for recognizing that in order for an idea to be considered scientific it must be falsifiable in principle (the criterion of demarcation), but this procedure is apparent throughout Darwin's work and in the research of other early scientists using hypothetico-deductive analysis (e.g., Mendel, Newton, Pascal, Pasteur).

One way to look at the value of the criterion of demarcation is in terms of what logicians call the fallacy of affirming the consequent. Consider a hypothesis and its derived predictions. Assume that the predictions are found to be true. Now consider the following argument: if the hypothesis is true, then the predictions must also be true; the predictions are true, therefore, the hypothesis is true. According to logicians the conclusion is not valid even if the hypothesis is correct. If the predictions of a hypothesis are confirmed, it is logically invalid to conclude that the hypothesis is correct, because some other hypothesis(es) might yield the same predictions. Thus, logicians argue that attempts to falsify hypotheses avoid this fallacy. It is logical to conclude that a hypothesis is false when its predictions are false.

But the major value of the criterion of demarcation is simply that it causes scientists to look for predictions that potentially can eliminate a given hypothesis rather than only evaluating predictions that, regardless of whether they are true or false, will support the hypothesis or multiple hypotheses. The importance of falsification of hypotheses to the advance of a science was eloquently stated by Darwin: "...false views, if supported by some evidence, do little harm, for everyone takes a salutary pleasure in proving their falseness; and when this is done, one path towards error is closed and the road to truth is often at the same time opened" (C. Darwin 1874, p. 606). Envisioning empirical observations that would disprove hypotheses is vital for scientific advance. However, extreme applications of Popperian philosophy and the fallacy of affirming the consequent in scientific endeavors are inappropriate because they deny the significance
of positive evidence in the achievement of true understanding and promote
the erroneous view that knowledge is an illusion (see below).

The strongest test of any hypothesis involves identification of competing
hypotheses that predict mutually exclusive, empirically falsifiable outcomes.
Such a test in its strongest form supports one hypothesis and falsifies the
alternative hypotheses (see Platt 1964). Another aspect of strong scientific
testing is that predictions be precise, which increases the likelihood of
falsifying the hypothesis generating the predictions. A hypothesis that has
passed many crucial tests involving precise mutually exclusive predictions
from alternative hypotheses can be said to be corroborated (Popper 1964).
The number of tests is not the important factor for evaluating the reliability
of a hypothesis. Instead, it is the number of severe or crucial tests that de-
termine our confidence in a hypothesis. The degree to which hypotheses are
corroborated varies, and science can lead to a degree of corroboration ap-
propriately labelled certainty.

The use of alternative causal hypotheses also avoids the natural tendency
of investigators to become attached to a pet hypothesis. Of course, one can
favor a pet hypothesis by setting up weak alternatives that have no chance
of matching observation. But this will fail ultimately given the self-correct-
ing nature of science.

By alternative hypotheses I mean alternative routes of causation, and
not simply a test of a null hypothesis and its alternative, because only one
causal hypothesis is identified in such a test.

Before outlining an example of the use of alternative hypotheses, I will
address the view that knowledge is an illusion. That is, the notion that,
given the necessity of employing the criterion of demarcation, we can never
know anything with certainty (e.g., Harris 1982), which stems from ex-
reme interpretations of the Popperian philosophy of science. An extreme
Popperian accepts only negative results. With this view only falsified
hypotheses are valuable; positive results do not count and certainly do not
imply understanding, because to accept positive evidence commits the
fallacy of affirming the consequent. This is where the actual practice of
good scientists deviates from a Popperian perspective. Although one cannot
actually prove a hypothesis to be true in the sense of logical proof, science
can lead to proof of a hypothesis in the sense of meaningful understanding
of natural phenomena. For example, the planets actually orbit the sun
and not the earth, and the earth is roughly spherical and not flat. The
hypotheses that the earth is the center of the universe and that the earth
is flat turned out to be totally wrong. The notion of pangenetic inheritance
is completely wrong. We now know that, barring cultural inheritance,
genesis (in interaction with environment) are responsible for parent-offspring
correlations in similarity. The list of what we know for sure is long and
growing. Any view of science that pretends we don’t know for sure that
insulin is produced in the isles of Langerhans, that bacteria and viruses
can cause disease, that cells are the building blocks of higher organisms,
that natural selection acts incessantly, that chromosomes house genes, etc.
is nonsense.

**Testing Alternative Hypotheses**

In this section I provide an example from my own research of how use-
ful alternative causal hypotheses with mutually exclusive predictions can
be. I use my own work because of my intimate familiarity with it. It deals with the nature of selection currently acting on a morphological feature of scorpionflies rather than the selection that has produced the structure; only the latter selection directly addresses the evolved function of a character. Later I discuss the distinction between selective maintenance and selective history and the methodologies for studying each of these categories of selection.

DORSAL CLAMP OF Parnara: SPERM COMPETITION OR FORCED COPULATION

Sperm competition is competition between ejaculates of two or more males for the fertilization of eggs of a single female. That sperm competition can be a potent selective force leading to male morphologies and behavior was first discussed by Parker (1970). This now classic paper has led to many studies of sperm competition as the selective force that has molded male reproductive characteristics. However, sperm competition is the only context considered when many investigators study insect characteristics such as copulation duration, copulatory frequencies, interactions of males and females during copulation, and post and precopulatory interactions of females and males. There are alternative hypotheses to explain these characteristics, and they require examination.

I began work on scorpionflies (Mecoptera) in 1971, the year following the publication of Parker’s seminal paper. The males have behavioral and morphological features that I initially interpreted as evolved in the context of reducing sperm competition. As my studies developed it became more and more difficult to accept this interpretation in all cases. This led to experiments beginning in 1977 designed to analyze alternative explanations of the traits.

There is a clamp-like structure on the dorsum of the male’s abdomen in scorpionflies of the genus Parnara. The dorsal clamp is formed from parts of the dorsum of the male’s third and fourth abdominal segments. The clamp holds the female’s wings during mating. Solitary males often attempt to disrupt copulating pairs and such males are occasionally successful. This led me initially to the interpretation that the clamp is important in preventing the female from being usurped and inseminated by an intruder, reducing the probability of the ejaculate of the usurped male being the fertilizing ejaculate. This interpretation was in keeping with Parker’s (1970) view of male-grasping morphologies—he saw them as evolved to prevent “take-overs”. (As it turns out, the dorsal clamp appears to be used for forced copulation; see below).

Male Parnara exhibit three alternative forms of mating behavior that are present within the behavioral repertoire of each individual. Two alternatives employed by males to obtain copulations involve nuptial feeding—the male presents a food item to the female during courtship and the female feeds on it throughout copulation. (1) A male may secrete a salivary mass. After saliva secretion, males stand near their salivary mass and disperse distance sex pheromone. A female attracted by the pheromone feeds on the saliva. (2) A male may feed a female a dead arthropod. In this case a male locates a dead arthropod, feeds on it briefly, and then dispenses sex attractant while standing next to it. (3) A male may employ forced copulation in which a male without a nuptial offering (dead insect or salivary mass) rushes toward a passing female and lashes out his mobile
abdomen at her. (Males engaging in forced copulation do not release pheromone.) If such a male successfully grasps a leg or wing of the female with his genital claspers, he then attempts to position her to secure the anterior edge of her right forewing in his dorsal clamp. Then the male attempts to grasp her genitalia with his genital claspers. The male retains hold of the female's wing with the dorsal clamp throughout copulation. Forced copulation in *Panorpa* is not an abnormal or "aberrant" behavior, but an aspect of the evolved behavioral repertoire of individual males that is widespread among species of the genus *Panorpa* (Thornhill 1980a, 1981, 1984a).

The behavior of females toward males with and without a nuptial offering is distinctly different. Females flee from males that approach without a nuptial offering but approach males that have nuptial offerings and behave "coyly" toward them. Females struggle to escape from the grasp of forceful males, but do not resist copulation with resource-providing males.

I have shown in laboratory and field experiments and observations involving several species of *Panorpa*, that the extent of use of each of the three behavioral alternatives by males is related to the availability of dead arthropods, which is determined by absolute abundance of arthropods and by male-male competition for the arthropods. Individual males prefer to adopt the three alternatives in the following sequence: dead arthropod > salivary mass > forced copulation. That is, when males are excluded from dead arthropods via male-male competition, they secrete saliva if they can (a male's ability to secrete saliva is determined by his recent history of obtaining food), and males only adopt forced copulation when the other two alternatives cannot be adopted. A male's body size influences his ability in male-male competition, and large males tend to adopt the use of dead arthropods as nuptial gifts, medium-sized males most frequently use saliva, and forced copulation is adopted most frequently by small males.

The behavioral alternatives contribute differently to male fitness. The preference of alternatives employed by males is consistent with female choice, and thus with male mating success. Females prefer males with arthropods over males with salivary secretions and actively attempt to avoid force copulating males. Also, the alternatives appear to be associated with different male mortality probabilities. Relative to large and medium-sized males, small males tend to lose in the competition for food, and thus are forced to feed on dead arthropods in the webs of spiders, which results in high mortality. Finally, force copulators have relatively low reproductive success compared to resource-providing males because females lay few eggs following forced copulation (Thornhill 1980a, 1981, 1984a).

Lab experiments have revealed that the dorsal clamp is essential for forced but not unforced copulation (Thornhill 1980a, 1984a). In one experiment beeswax was used to cover the dorsal clamps of males that had been starved, which prevents them from secreting saliva. The dorsal clamp of starved control males was left functional. Treated and control males attempted copulation with equal frequency, but only control males succeeded. Treated males tried to reposition females so as to secure their forewing in the clamp, but the females escaped by struggling. In other experiments males with saliva or dead crickets were treated with beeswax. In these tests treated males readily copulated with females. Furthermore, insemination rates for treated males in unforced copulations was 100%.
The experimental results confirm predictions of the hypothesis that the dorsal clamp is important in the context of increasing the success of forced copulation attempts when sexual competition forces individuals into this alternative behavior. Yet, despite the apparent uselessness of the clamp in unforced copulation, the female's wing is placed in the clamp during both forced and unforced copulation. Could the dorsal clamp be important solely or in part in some other context? The experiments only superficially address this question. They were designed to test predictions of a forced copulation hypothesis. The predictions that were tested could be consistent with those stemming from an alternative hypothesis(es) for the role of the dorsal clamp.

A reasonable alternative was identified earlier. It views the dorsal clamp as important in the context of sperm competition—as a structure that prevents disruption of copulating pairs and the insemination of the female by an intruding male. Aspects of the reproductive behavior of male *Panorpa* are consistent with this hypothesis—pair disruptions by intruding males are not infrequent, disruptions sometimes result in the intruder copulating with the female, and the clamp is used in both forced and unforced copulations. Mutually exclusive predictions from the two alternative hypotheses are easily identified. If the clamp serves a male's reproduction by reducing the probability of the takeover of a copulating male's mate, one would expect treated males (dorsal clamp occluded with beeswax) to experience higher takeover rates than untreated males. But if the clamp is used solely for something else (i.e. is important in the context of forced copulation, or some other context consistent with the predictions examined earlier from the forced copulation hypothesis), one would expect the treated and untreated males to experience similar rates of takeover. If the former prediction (takeover high when clamp functionless) is supported, the forced copulation hypothesis would be eliminated as the sole explanation of the selection presently acting on the dorsal clamp. If the latter prediction (takeover rate not influenced by clamp) is supported, the takeover hypothesis would be falsified. The predictions are strong in the sense that they offer potential for falsification.

The predictions from the two hypotheses were tested with a lab experiment. The results reveal that copulating males whose clamps were covered with beeswax had the same takeover rate as untreated copulating males (Thornhill 1984a). Thus, the clamp apparently does not presently aid a male's reproduction in the context of take-over attempts—i.e. there is no selection on the clamp in the context of take-over. If the experiment had shown that treated males experienced significantly higher take-over rates, the forced copulation hypothesis would remain potentially very important because of the findings from experiments described earlier.

**DORSAL CLAMP OF Panorpa: OTHER HYPOTHESES**

There are other alternative hypotheses that can be considered in an attempt to understand the selective maintenance of the dorsal clamp. Felt (1895) observed female *Panorpa debilis* palpatating the dorsal clamp of males during courtship. I have observed palpation of the dorsal clamp during courtship by females of several *Panorpa* species. Felt concluded that the structure probably secretes a volatile oil that attracts the female to the male and primes her for mating. Felt's hypothesis is incorrect. I have
examined the histology of the dorsal clamp of several species of Panorpa, including debilis, and found no associated glandular tissue. The hypothesis is also inconsistent with experimental results. When the dorsal clamp of resource-providing males is covered with beeswax, presumably preventing any odors from being released, females mate readily with the males.

The dorsal clamp could be of value to males in species recognition. The dorsal clamp varies in morphology across Panorpa species. The examination (olfactory and apparently visual) by the female prior to copulation could reduce the probability of an interspecific mating error. This hypothesis must be examined because it has been, and still is, a widely used evolutionary explanation for species differences in courtship and mating behavior and associated morphological features. Darwin argued that sexual selection was the most important context for the evolution of sexual differences in sexual behavior and morphology, but Wallace identified species and sex recognition as the more likely contexts (see Thornhill 1980b). After Darwin and until recently, premating and mating behavior and associated morphological features have been generally viewed as functioning as reproductive isolating mechanisms—that is, as adaptations that prevent wasted reproductive effort by individuals in heterospecific interactions (Thornhill and Alcock 1983, West-Eberhard 1983). The theory of sexual selection appears to have far greater predictive power for understanding the diversity of these traits (Thornhill and Alcock 1983, West-Eberhard 1983). But the species' and sex identification hypotheses can serve as alternative hypotheses in studies of traits presumed to be important in the context of sexual selection or reproductive competition in general.

Although interspecific mating errors could potentially occur because of the co-occurrence of sexually active adults of several species of Panorpa in time, evidence indicates that females do not depend on cues from the dorsal clamp for species (or sex) discrimination. Male Panorpa produce species-specific sex pheromones that attract females from a distance and may also serve in close-range interaction (Thornhill 1979). Males exhibit species-specific courtship actions (wing and body movements). Also, when females approach a heterospecific male they do so to obtain a meal—i.e., they attempt to feed on the nuptial offering. Females never behave coyly toward heterospecific males as they always do toward conspecifics. Females of large species attempt to usurp, via aggression, the nuptial offerings of males of small species, and are sometimes successful. Resource-holding males behave aggressively toward heterospecific females that approach them. These considerations indicate that species and sex identity are discerned prior to close-range courtship by both sexes. Finally, the experiments discussed above are strongly inconsistent with the sex—and species—discrimination hypotheses. In these experiments the dorsal clamp was covered with opaque beeswax and thus it would not apparently emit normal visual (or olfactory) cues. Yet, females readily mated with treated males who offered nuptial gifts.

At this point in my research on the dorsal clamp of Panorpa, results suggest that this structure is maintained solely by selection in the context of forced copulation. The alternative hypotheses I have considered were either falsified (the sperm competition hypothesis and the pheromone-emission hypothesis) or inconsistent with existing evidence (the sex and species identification hypotheses). Only further tests would falsify the sex
and species identification hypotheses. The general point I want to emphasize is that alternative hypotheses should be impartially examined in relation to all available evidence. In a subsequent section I further consider the dorsal clamp in relation to comparative evidence that could provide the best understanding of its evolved function (as opposed to its selective maintenance).

**Methods of Applying the Hypothetico-Deductive Model**

The hypothetico-deductive system for ascertaining cause and effect can be applied in four ways: lab experiments, field experiments, observational analysis, and comparative analysis. All four methods are equally valid scientific procedures, all four have strengths and weaknesses and they are based on different premises. The first three methods primarily yield information about present selection maintaining a trait of interest (i.e. about a trait's present contribution to reproduction), but only the comparative method can provide information about the selective history or evolved function of a trait.

**Experimentation**

The essence of experimentation is manipulation. The experimental method involves some systematic variation of a variable of interest. But lab and field experiments represent very different scientific procedures. Lab experiments typically take the form of attempts to control all variables but one. Field experiments often do not involve control of variables that may confuse results; instead, all parameters except the manipulated one are allowed to vary naturally. Randomization of treatments in field experiments can produce limited control of confounding variables (e.g., site effects) but always many variables remain uncontrolled. Typically, one is less certain about the influence of other variables on the result with field experiments compared to lab experiments. Even with lab experiments, however, all potentially confusing variables cannot always be controlled. The number of confounding variables that could cause a given result is potentially infinite. Also, lab and field experiments are very different in that the lab is at best seminatural, and thus one can never be sure that lab results address nature.

A major problem with the experimental method is that manipulation effects often confuse results. It is always difficult to determine if one's result is due to the presumed cause, the manipulated variable, or is an experimental artifact unrelated to presumed causation. For example, it would seem from many lab and field experiments that nitrogen content of soil is a cause of plant growth and health (the effect). However, the correlation between nitrogen content and plant growth may be spurious (Grover 1982), because when nitrogen is added other changes occur (e.g., redox potential). These changes may be the actual cause of improved plant growth (effect). Other experiments will be necessary to control parameters that confuse the result.

Another problem with experimentation is the difficulty of determining the appropriate parameters to manipulate in order to yield biologically relevant information. It may take years of natural history observation on a biological system before an appropriate experimental procedure can be identified. For example, the availability of the limited resource (dead
arthropods) is a major determinant of the behavioral variation among male Panorpa (see above). Which alternative behavior is adopted by a male depends on his ability to obtain resources. Resource abundance for a male is a function of absolute abundance of dead arthropods in the habitat, sex ratio (number of competing males), and size of conspecific and heterospecific male competitors (Thornhill 1981). It took several field seasons to obtain sufficient natural history information to identify the field and lab experiments that would provide results relevant to understanding Panorpa in nature, and it is my experience that biologists are often willing to initiate experimental work without a proper understanding of the natural history of the system they are interested in. Until recently experimental psychology was the epitome of such an approach. Elegant, elaborately controlled lab experiments were conducted with only repeatability in mind and without concern for the biological meaning of the results.

Observational Analysis

By observational analysis I mean the testing of predictions with observational data in the absence of manipulation or comparative analysis. The absence of manipulation is one of its strengths, but its weakness is a lack of rigorous controls. Confounding variables plague observational work, some can be eliminated by refining observations via further observational analysis pertaining to some question about a trait (e.g., see Clutton-Brock and Harvey 1979 and Skinner and Charnov in this symposium volume for examples), but many problem variables will remain. Despite this problem, crucial tests can be conducted via observational analysis (i.e., tests involving mutually exclusive quantitative predictions from alternative hypotheses).

Parker's (1978) work on the dung fly Scatophaga stercoraria provides an example of observational analysis. Initial experimental research on sperm competition revealed that the longer a male mates the more eggs of his mate he fertilizes, up to 100 minutes of mating which results in all eggs fertilized. But the egg gain for the copulating male rapidly diminishes at about 40 minutes, a duration which yields 80% fertilization. With this information and an understanding of the average time necessary for a male to guard his mate from rivals until she lays her eggs, and then begin searching for another mate, Parker predicted that the optimal copulation time should be 41 minutes. This prediction is precise and thus easily disproved, and is logically derived from the hypothesis that males behave so as to maximize the number of eggs fertilized per minute. The actual average time in the field turned out to be 36 minutes. The prediction about copulation duration might be improved by considering the cost of extended copulation in relation to male size, rather than as an average for all males, because small vs. large male dung flies have different opportunities for access to multiple mates (see Borgia 1980).

Although explicit alternative hypotheses were not considered in the dung fly work they could be in future studies. It is not correct to argue that the use of alternatives is unnecessary because the prediction is precise enough to exclude a great many alternative hypotheses. Both the Bohr theory of the atom and the Schrödinger theory predict exactly the same Rydberg constant!

Observational analysis can be conducted in the same fashion as evolution-
ary comparative analysis (below). By using observational comparisons which by their nature and number randomize and thus control confounding variables, one can arrive at robust conclusions. A good example from everyday life is the work on the role of seat belts in preventing injury during automobile accidents (see Alexander 1979, p. 12). Another example is the research on the role of genetically inherited tendencies in criminal activity. With regard to the latter, sociologists have realized that with appropriate comparisons the quality of the result obtained is as good as in a rigorously controlled lab experiment (see Ellis 1982).

COMPARATIVE ANALYSIS

The comparative method involves species or population comparisons conducted so as to randomize the influence of confusing variables on the effect of interest (Alexander 1978, 1979). This method is based on convergent and divergent evolution. The former involves distantly related forms converging on an adaptation because of similar selection pressures. The latter pertains to closely related forms diverging in adaptation to come condition because of different selection pressures.

One real challenge in biology is understanding events of the distant past without which our understanding of evolutionary history would remain incomplete. The comparative method is analogous to a time machine, and with it we can ask what selective force operated to lead to the present expression of a characteristic of interest?

The comparative method randomizes the influence of confounding variables on a result in such a way that single presumed causal forces can be examined, though complete randomization is difficult, perhaps impossible. The difference between the comparative and experimental methods is not in controls, both methods include controls, but in the presence or absence of manipulation. Because manipulation effects can confuse results, the absence of manipulation in comparative analysis is one of its strengths.

With the appropriate comparisons, the comparative method can lead to results as precise as those obtained through other methods. Because life is incredibly diverse it provides vast numbers of appropriate comparisons for almost any question about the selective history of any biological characteristic. The ingenuity in use of the comparative method involves recognition of appropriate comparisons, i.e., those comparisons that because of their number and diversity are likely to randomize and thus control the influence of other variables on the result.

None of the methods I’ve discussed is perfect; all have inherent problems. Experiments designed to ascertain cause and effect are improved through time by investigators interested in a given cause-effect relationship. Experimenters strive for refinement; better controls and manipulations are created. Likewise, investigators using comparative analysis to elucidate a given cause-effect relationship make new and better comparisons. It is incorrect to suggest that a finding derived from the comparative method is less accurate than a finding derived from the experimental method. One’s confidence in a given finding obtained by either method should depend on the quality of controls employed.

The erroneous opinion that only experiment provides reliable results has led to inappropriate conclusions and research directions in biology. After
Insect Behavioral Ecology—’83 Thornhill

discussion of a few well documented cases of this in the history of biology, Mayr (1982, p. 856) said, “It would be interesting to go through the history of science and see how often a misplaced insistence on experiment has caused research to move into unsuitable directions.”

It has been said that the comparative method can provide insights about general patterns in nature but cannot elucidate cause and effect (see Reznick 1982). That is, the comparative method can generate correlations but is not a method for examining causation; its results must be tested by other methods. This view is puzzling. All scientific findings are correlations. I don’t mean that all scientific findings are represented by Spearman’s or Pearson’s correlation coefficients. I mean that all scientific knowledge is based on relationships between variables—presumed cause and effect relationships between variables. Regardless of whether we use experiments, observational analysis, or the comparative method to determine significance between sets of observations, and whether we use regression, t-test, etc., we are examining and attempting to construct and refine correlations.

Examples of Comparative Method

Two studies that should drive home the essence of comparative work are one dealing with the function of sexual reproduction and another with sexual dimorphism in vertebrates. Williams’ (1975) analysis of the selective background of sexual reproduction is classic in this regard. Williams hypothesized that sex is a parental adaptation to the likelihood that offspring will face changed or unpredictable conditions. He made several predictions from this hypothesis about the occurrence and nature of sexual and asexual reproduction in organisms and tested these via the comparative method. One of the most important predictions he made was: in organisms that employ both sexual and asexual reproduction, sex should occur in the life cycle prior to changed or unpredictable conditions. This prediction was met. In effect he was attempting to randomize the influence of confounding variables on the timing of sex in life cycles of organisms with both modes of reproduction. In this analysis he examined the evolved function or selective history of sex. His analysis allows us to go back in time and begin to understand factors influencing differential reproduction of individuals as a result of variations in sexuality. His analysis is not the final answer, but it provides a beginning answer to an important question. Present attempts to elucidate the function of sex appropriately focus on tests of alternative hypotheses via comparative analysis (Bell 1982).

It has been assumed since Darwin that a positive relationship exists between adult sexual size dimorphism (with males larger than females) and degree of polygyny in vertebrate mating systems. Darwin reasoned that as the degree of polygyny increased, more and more males would be excluded from reproduction, which in turn would cause selection for male combative traits, including large body size. Comparative evidence suggested that the predicted relationship holds for birds and mammals, but no rigorous comparative analysis had been conducted until recently. The most rigorous tests have been done by Alexander and colleagues (1979), Clutton-Brock and Harvey (1977, 1979), and Payne (1983).

The approach is to make comparisons of vertebrates that differ in sexual dimorphism and breeding systems. If appropriate comparisons can be made the influence of confounding variables on the result can be controlled,
Alexander and colleagues found significant positive relationships between degree of sexual size dimorphism and degree to which breeding systems deviate from monogamy toward extreme polygyny in ungulates, primates, and pinnipeds. Closely related species diversified, and distantly related species, even across orders, converged in sexual dimorphism in relation to the extent of sexual competition in evolutionary history.

Clutton-Brock and Harvey analyzed 42 species of primates and found the same relationship that Alexander and colleagues found for this taxon. Clutton-Brock and Harvey’s work is stronger in the sense that they used more species, carefully considered which taxonomic level should be used for analysis (e.g., species, genus, or family), paid careful attention to alternative hypotheses, and attempted to eliminate allometric effects, but their analysis is weaker because it did not cut across distantly related taxa, as did the analysis of Alexander and his coworkers. Payne’s work on a diversity of bird taxa provides a comparative test of alternative hypotheses for the relationship between mating systems and sexual dimorphism; his results also support Darwin’s hypothesis for sexual dimorphism.

The comparative method has made some sense out of seemingly chaotic natural variation but there is still unexplained variance in the correlations that have been discovered. The presence of unexplained variance in a significant relationship between two variables does not in itself mean that the relationship or the prediction behind the relationship are questionable. I know of no presumed cause and effect relationship in any area of science in which all the variance is understood. This is why statistical analysis is used to detect significant differences in results. Regardless of one’s procedure—experimental, observational, or comparative—there are always exceptions to expected patterns.

*Other Applications of the Comparative Method*

I have focused my discussion of the comparative method on its value in studies of evolved function. This is where this method has been most successful. But Darwin, the inventor of the comparative method, used it to examine phylogeny, speciation, biogeography, soil formation by earthworms, and coral reef formation. Darwin even took a stab at community structure using the comparative method. Ghiselin (1968) and Gould (1982) address in detail the advances in the scientific study of historical phenomena that Darwin provided us with. The comparative method should be equally potent for studying all long-term events.

All methods need constant refinement, and it is likely that some considerations useful in analyzing evolved function will not work for biogeography or community structure and vice versa. Some of the problems that ecologists have encountered recently in attempts to apply the comparative method to community structure (Case and Sidell 1983 and references therein) may stem from not separating community organization into evolutionary and ecological components. The latter may change rapidly and lead to lack of fit between comparative predictions and pattern. The evolutionary component of community organization, perhaps synonymous with biogeography, is historical and thus subject only to comparative analysis. Clutton-Brock and Harvey (1976, 1979), Clutton-Brock (1982), Harvey and Mace (1982), and Jarman (1982) have provided some refine-
Selectivc History vs. Selective Maintenance

Perhaps the most rigorous approach for examining a hypothesis is the use of critical tests involving all methods, because each method is based on different assumptions. When results from all methods point to the same conclusion, one usually achieves considerable confidence about cause and effect. But the combination of methods to examine a question must be done in a fashion that considers the relative potency of methods for ascertaining information about the role of selection in the history vs. the current maintenance of a trait.

Both experimental and observational analysis may cause one to reject a correct hypothesis about evolved function when predictions are not met because the tests are done in an evolutionarily novel environment. One may show via experiment or observation that a prediction from a hypothesis about the evolutionary history of some presumed adaptation is not correct, but one does not know if the validity of the hypothesis has really been examined.

Suppose a hypothesis predicts that the reproductive success of males of a species should be positively correlated with body size (e.g., the sexual dimorphism hypothesis discussed earlier). When one tests this prediction, no correlation is found. Thus, it seems that the hypothesis has been falsified. But is the hypothesis wrong or is the species living in an evolutionarily novel environment in which body size is not related to sexual access, despite the correlation between fitness and male size in the evolutionary history of the species. The novel environment might involve artificially high population numbers or very low food supply brought about by human activity. Under the former condition even the strongest and largest males may be unable to control the activities of other males that persist in copulation attempts. Under evolutionarily abnormally low food levels, large males may be too weak to fight for females or to intimidate males through display (because of the greater energetic cost associated with large body size). Sometimes the evolutionary novelty in the environment is apparent and research can be modified to take this problem into account (Jarvis 1974). But in most cases novel circumstances are unknown or incompletely understood.

Although novel environments may allow detection of counter selection on a trait—e.g., natural selection may reduce body size and sexual selection increase it—such circumstances present problems for testing functional hypotheses via experimental and observational methods. Furthermore, there is the difficulty of measuring fitness in a way that will provide meaningful answers via short-term and site specific experimental or observational analysis. On the other hand, observational and experimental results provide the only means of obtaining information about which selective forces (a) are presently favoring a trait or if the trait is being selected against. This information is of great interest.

But only the comparative method can yield answers to questions about selection pressures that have led to present features of living things (also see Curio 1973 and Mayr 1982, pp. 855-6). If a functional hypothesis predicts a correlation between sexual size dimorphism and breeding system,
the more species that fit the assumptions of the hypothesis and also fit the prediction, the more one has confidence in the hypothesis. The more evidence for predicted convergence and divergence, the more assured the investigator is that the hypothesis is correct.

Thus the comparative method can corroborate results pertaining to possible evolved function obtained by other methods, but other methods cannot question results from comparative analysis. Observational and experimental studies can be coupled with comparative analysis in order to determine if a force presently favoring a trait is the same as the selection that produced it.

For these reasons, when possible in my own work on scorpionflies, I employ all methods in an attempt to examine major questions about specific traits (see Thornhill 1981). The experiments and observations I described earlier provide understanding of selection maintaining the dorsal clamp of male Panorpa. I consider the following prediction an important test for elucidating the evolved function of the dorsal clamp: the frequency with which males of the various species of Panorpa exhibit forced copulation should be positively related to the size of the clamp (or other morphological correlates of the effectiveness of the structure as a clamp). The alternative hypotheses pertaining to the dorsal clamp would not generate this prediction, but they would yield other mutually exclusive ones that could be tested against comparative data. Panorpa spp. and species in related genera exhibit considerable diversity in shape and size of the dorsal clamp (Byers and Thornhill 1983), but too little information is presently available for appropriate comparative work.

It is argued that the nature of selection that has shaped a trait can never be ascertained for sure because multiple selection pressures probably operated on the trait and in the same direction. This view is implicit in any attempt to rank selection pressures in importance in studies of the evolved function of biological characteristics. This argument is invalid. If two selection pressures are acting in the same direction, and one is stronger than the other, the degree of evolutionary modification of a trait will reflect the stronger of the two, because the organism cannot “feel” the weaker pressure. This holds regardless of number of selective pressures. It is highly unlikely that two selection pressures acting on a trait will be equally potent, but even if this does occur the degree of adaptive modification will reflect the strength of any one of them alone. Ecc Curio (1973, p. 1046) for a full discussion of this. Any trait has many effects, some beneficial and others harmful to fitness, but only one fitness effect is its function—its reason for being (Williams 1966). This is why it is appropriate for biologists to attempt to locate a single evolved function for each trait they are interested in. Ecological phenomena, including selective maintenance of a trait, may have multiple causes, but a question about evolved function has only one causal answer.

**Scientific Progress: Improving Correlations**

I think it is safe to say that there is no satisfactory account of how science progresses or accumulates knowledge. Some argue that science does not progress; that any impression of progress in understanding is an illusion. Others argue not only that science doesn’t progress but that it changes in an arbitrary way. This stems from the view that, at any time,
the theories being examined in a science change like automobile styles. There is no right or wrong theory, only popular and unpopular theories. And some feel that the major determinate of popularity of an idea is cultural attitude. For example, when football and other sports are popular and during World Wars ecologists are expected to believe that competition is important as an organizer of community structure! Undoubtedly, our beliefs, which are culturally inherited from parents, friends, etc., influence how we view the world scientifically. This is especially apparent in the writing of scientists who adhere to Marxist philosophies, but is also present in more subtle form in the work of anyone who adheres to any strong ideology. But as Ruse (1982) has pointed out, ideas in science that stem from ideology, cryptic or blatant, are not necessarily bad, because they are subject to the same criterion as other ideas. When ideological perspectives are paraded as science they will eventually meet the criterion of demarcation and be eliminated if they are wrong.

I suggest that scientific progress ultimately has nothing to do with arbitrary popularity of ideas, but instead, advances happen via improved understanding of the relationship between variables—that is, by increased understanding of correlations between presumed cause and effect. Every scientific hypothesis (or theory) describes an expected correlation between a cause(s) and an effect. To elucidate the variance in such relationships via the hypothetico-deductive model is the occupation of scientists. In no case is all the variance understood, even for simple relationships. I'll take a relatively simple example from my own work to illustrate my view of scientific progress and the achievement of greater understanding in general.

I began work on female choice in the scorpionfly Hylobittacus apicalis in 1971. H. apicalis exhibits nuptial feeding; the male feeds a female a prey arthropod during courtship and throughout copulation. The sizes of prey carried by males vary. Initial observations led to the hypothesis, based on evolutionary theory, that females will value material and genetic benefits of males in mate choice. I assumed that the material benefits (nuptial prey size) and genetic benefits (offspring quality) a male could deliver to a female would be positively correlated. I examined two qualitative predictions from this hypothesis: 1) females will sometimes refuse to mate with males with small prey, and 2) the duration of mating will be positively related to nuptial prey size. The general correlation I sought to define and explain was choice behavior of females (effect) in relation to the prey size males possess (presumed cause).

Initial research revealed that the predictions were upheld (Thornhill, 1976, 1977). Many females refused to copulate with males with small prey (i.e., less than 16 mm² in surface area) and there was a positive relationship between mating duration and nuptial prey size. I also discovered that copulations involving large prey (≥ 16 mm²) are terminated by males, whereas copulations involving small prey (when a female allows copulation) are terminated by the female. Furthermore, by interrupting lab copulations involving virgin females I found that in the first 5 minutes of copulation few or no sperm are transferred to the female. From 5 to 20 minutes there is a direct positive relationship between number of sperm transferred and mating duration, and beyond 20 minutes of mating no further sperm are transferred. Finally, studies revealed that females lay eggs and do not mate again following matings with males with large prey, but females
remain sexually receptive and do not lay eggs when they mate with a male with small prey.

The relationship between copulation duration and sperm transfer yielded the prediction that when females mate with males with small prey they will terminate copulation after 5 minutes; that is, at about the time when sperm begins to flow. This way a female obtains a brief meal but receives few or no sperm from inferior hunters. Field observations revealed that mean time of mating involving small prey was 5.8 minutes (Thornhill 1980c). Although this is close to the predicted 5 minutes there was still much variance in mating times when small prey were involved. But I was making progress in terms of understanding the initial relationship of interest to me. I now knew that 50% of the time females reject males with small prey and when females allow such males to mate they terminate mating on average at about the time sperm begins to flow from male to female. The correlation between female mate choice and male nuptial prey size was being improved.

At this point in the study I modified the hypothesis in an attempt to understand more variance in the relationship, but the modification did not change the general correlation I was studying. The modified hypothesis is:

The relative value of material and genetic benefits for females in choice decisions depends on female conditions of body size, feeding history and mate availability. Basically, I was proposing that every female will strive to maximize material and genetic benefits received from males. But I need to clarify how I derived the modified hypothesis. Energetic cost of body maintenance is positively related to body size; the larger the animal the more nutrients required. Also, feeding history of females was expected to vary and thus some females should be more willing to mate with males with small prey. Mate availability should also influence female choice; the more males available the more choosy females might be.

The major predictions the modified hypothesis yielded were: 1) large females will be more likely to mate with males with small prey than will small females, 2) there will be a positive relationship between female body size and duration of matings involving small prey, 3) large females will behave like small females when fed prior to placing them with males possessing small prey, and 4) independent of female body size, females will become more choosy as potential mate availability increases.

I have begun testing these predictions. All four are supported. The first three are well substantiated, but the fourth will need more testing in order to clarify completely the role of male density in female mate choice (Thornhill 1984b). Thus, I have considerable understanding of the variance in the correlation between female mate choice and mate nuptial prey size. I do not have all the variation explained; there are still exceptions. But if I know a female's body size, her recent feeding history and the mate availability I can predict with considerable accuracy not only whether the female will mate with a male with small prey but also the duration of the mating if it occurs.

I emphasize that the the sequence my work followed is distinctly different from endeavors involving the addition of parameters that explain more and more variance in a data set. A common procedure in areas of social science and biology is to employ multiple regression analysis to determine the ability of presumably important parameters to explain variation. In some cases, parameters are added until most or all of the variance is ac-
counted for. This procedure involves ad-hoc explanation. It is not valid
science to modify a hypothesis to account for unpredicted observations and
then claim that the hypothesis is confirmed. Likewise, it is invalid to
construct a model from data and claim that the model is confirmed. It is
scientifically accurate, however, to use unpredicted observations to modify
or eliminate a hypothesis and then test the predictions from the modified
or alternative hypothesis with new observations (or simply suggest the
direction that testing should involve).

I feel that the sequence of events the *Hylobittacus* work went through is
the appropriate and typical sequence for scientific hypotheses and even
general theories when they are successful. (See Skinner and Charnov's
paper in this symposium volume for an additional example of this sequence.)
Hypotheses other than the one outlined that I considered in my research
sequence with *Hylobittacus* are discussed in Thornhill (1980c, 1984b). A
particular sequence may involve only one or all methods of applying the
hypothesico-deductive model. Also, a particular sequence may involve one
or multiple investigators examining the same presumed cause-effect
relationship. At any stage in a sequence a hypothesis (or theory) may be
disproved and replaced by another hypothesis (or the original hypothesis
modified) that attempts to account for the relationship.

Even great theories like those of Darwin, Einstein, and Newton portray a
relationship or correlation between variables. Darwin's theory is by far the
most comprehensive theory in science in the sense that it is directed at ex-
plaining life, the most complex and diverse phenomenon known to hum-
ankind. But still even this theory rests on the relationship between the
diversity of life (effect) and a history of differential reproduction of in-
dividuals (cause). Since Darwin, biologists have been attempting to under-
stand the variance in this relationship. Biologists ask questions about the
relation between imagined selection pressures and diversity in sexual di-
morphism, life history, chromosome structure and number, mating be-
behavior, etc. The imagined selection pressures of biologists serve as alter-
native hypotheses which succeed or fail to explain subrelationships of the
general correlation Darwin's theory generated.

**Footnote**

*Karl Popper (1934), a philosopher of science, has argued that only falsifiable ideas are
within the realm of science, and that tests of scientific hypotheses (or theories) should
focus on attempts to falsify them. His ideas have been very influential in many areas of
science, including biology.*

**Acknowledgements**

I thank John Alcock, Bill Kulpers, Jim Lloyd, Larry Marshall, John
Sivinski, Nancy Thornhill, John Wiens, and Bruce Woodward for construc-
tive criticisms on the manuscript. The Population Biology and Psychobiology
Programs of the National Science Foundation have generously supported
the research from which this paper is derived (grants BNS-7912208, DEB-
7910293, BSR-8219810).

**Literature Cited**

Sinauer and Assec., Sunderland, Massachussette.


Insect Behavioral Ecology—'83 Thornhill


———. 1983. How to carry out the adaptationist program? American Nat. 121: 324-34.


———. 1980a. Rape in Panorpa scorpionflies and a general rape hypothesis.


