动物学报　54(2):383－386，2008
Acta Zoologica Sinica

Hypothesis testing and the scientific method revisited

Jerry O. WOLFF[1*], Charles J. KREBS[2]

1. Department of Biological Sciences, St. Cloud State University, St. Cloud, MN 56301, USA

2. Department of Zoology, University of British Columbia, Vancouver, B. C. V6T 1Z4, Canada

Abstract Our intention in this essay is one of consciousness raising for investigators to revisit the scientific method, make good use of hypothesis testing in research design, and have a good understanding of natural history in the selection of model species. Our overall premise is that science could be advanced more quickly, objectively, and conclusively by following the logic of "Strong Inference", which posits rejecting alternative hypotheses and not just supporting favored hypotheses. We emphasize the importance and provide a logical sequence of steps from the observation, i. e. identification of a problem or unknown that has biological significance, a listing of all credible alternative hypotheses that could explain the observation, each with its own set of testable and falsifiable predictions, followed by an experimental or research design that is a logical outcome of these predictions. We also emphasize the importance of selecting model species that are appropriate for addressing a specific theoretical question and drawing inferences. What we present is not new, rather a reminder of the value of following basic scientific approaches for the advancement of science as well as individual careers. [*Acta Zoologica Sinica* 54 (2): 383－386, 2008].

Key words Hypothesis testing, Scientific method

假说检验和科学方法

Jerry O. WOLFF[1*]　Charles J. KREBS[2]

1. Department of Biological Sciences, St. Cloud State University, St. Cloud, MN 56301, USA

2. Department of Zoology, University of British Columbia, Vancouver, B. C. V6T 1Z4, Canada

摘　要　本文的意图是让研究者审视研究方法，并在研究设计中充分使用假说检验，并在选择模式物种时充分理解其自然史。我们的总前提是，按照“强推论”（指假定拒绝某一假说而不是支持某一偏爱假说）的逻辑，科学能够进展得更快、更可观、更有确定性。我们强调并提供了符合逻辑的一系列步骤，即确定科学问题或确定具有未知生物学意义的问题；列出所有可靠的、能解释所观察现象的假说，每个假说列出其可检验的、可证明其无根据的预测；然后是符合预测检验的实验或研究设计。我们也强调，模式物种对于解决科学的理论问题以及得出推论是很重要的。本文所展示的不是新思想，只是提醒研究者要注意遵循的基本研究途径［动物学报 54（2）：383－386，2008］。

关键词　假说　检验　科学方法

We wish to raise the question about how ecologists carry out their scientific investigations today in the light of the recommendations of the classical scientific method. Over the last several years we have presented invited talks and workshops at many universities, research institutions, and conferences in North America, Europe, Southeast Asia, and Australia on our research, but also on a more general approach to the use of theory and hypothesis testing in research design in behavior, ecology, and evolution. One thing that became apparent is that we seem to be raising a cohort of young investigators that have lost sight of the scientific method and how to use an inferential approach to their own research or in critiquing that of others. We provide here an overview of how to correct this imbalance.

Good science must start with an observation; one that has biological significance and is founded on current evolutionary theory. An observation becomes the problem or question to be asked. For instance, populations seem to be declining, why? Females occupy exclusive territories and males show considerable overlap, why? Young males have a tendency to disperse from the natal site whereas females are philopatric, why? These are rather universal questions that have far-reaching implications for a large

---

Received Oct. 29, 2007; accepted Nov. 08, 2007

* Corresponding author. E-mail: jowolff@stcloudstate.edu

number of species and therefore are relevant to a large part of ecological research. Starting with a good question prevents the audience or reader from asking the "so what" question about your results. A sound theoretical framework is essential to asking relevant questions that advance science and move the discipline forward.

Once the observation or question has been clearly articulated, the second question you, as a young ecologist should ask is "what is my hypothesis". A scan of 264 articles published in this journal over the last 5 years revealed that only 28 (11%) used the term "hypothesis" in their objectives. The reason for this, you might argue, is that the hypothesis is implicit in the study and does not need to be stated. This presumption is a slippery slope because two people reading the same paper will not necessarily interpret the "hypothesis" in the same way. Rule # 1: be explicit about your hypothesis. This approach means, as every statistician will tell you, that you must state exactly the ecological situations to which your hypothesis will apply. Thus, it is important to state clearly what inferences you will draw upon completion of your study. Will your results apply only to passerine birds? Or only one species of bird? Or only one population of a species?

Given now that you have a hypothesis, you cannot just dive into the ecological details of your study. You must also list the alternative hypotheses that are available for the particular problem under study. If there are no alternative hypotheses, there is no need to do the research because the initial hypothesis must be correct if there are no possible alternatives.

One common mistake in ecological research is to state one hypothesis and no alternatives, and then to interpret your data in terms of support for your hypothesis. We remind you of the Popperian approach to research that was clearly articulated by John Platt (1964) in his excellent article published in *Science*, "Strong Inference". Karl Popper reminded us that science advances not by gathering data that support a hypothesis, but rather by rejecting alternatives. The scientific method (Figure 1), as taught in nearly all introductory general biology courses, includes an observation, listing of all alternative hypotheses that could explain that observation, design of experiments or an indication of the kinds of data necessary to accept or reject each hypothesis independently, conducting the experiments or gathering the data, data analysis, conclusions, and the drawing of inferences back to the theory associated with the original observation. This dry analysis of what science is all about tends to become lost when graduate students begin their research careers, partly because the excitement in ecology is all about the biology of the system under study. The important point is to maintain this excitement while tempering it with a rigorous approach through hypothesis testing.

We have argued before (Wolff, 1995, 2000; Krebs, 2000) that the presentation of observational data without hypotheses does not advance science. We distinguish information from science. Much research published today might provide information; that is, data, but not necessarily knowledge or science. We must think of science as a *system* of acquiring knowledge based on the scientific method as well as the organized body of knowledge gained through such research. What separates science from information is that science is knowledge that makes falsifiable predictions. We made this distinction in Wolff (2000) and Krebs (2000) and illustrate the difference here with a couple of examples. Research that measures the food habits of the muntjac or nest locations of the Mandarin duck provide *information*, but not *knowledge* and therefore is not science. These studies are not science because they do not set up hypotheses and do not make falsifiable predictions. Another study can find that muntjacs eat different foods or Mandarin ducks nest in different locations, neither of which would affect the conclusions of the first two studies. The first two studies do not make any predictions regarding cause and effect. In contrast, research such as "the relationship between nitrogen content and food habits of the muntjac" or "the effects of predation on nest site location of Mandarin ducks" makes falsifiable predictions in that they conclude cause and effect relationships that can be falsified with additional research. In fact, this critical testing of prior information is what advances science, rather than confuses it with just more information.

At the present time there is considerable interest in the potential impacts of climate change on populations and communities. To begin to answer this broad question, ecologists are tempted to go into the field and measure many things in the absence of any clear hypotheses. If we follow this approach, we will have much information about climate change effects but no science on the ecological consequences of climate change. The temptation to measure many variables in ecological systems should not be confused with a scientific approach to this important issue of our times.

Another confusion in attempts to apply hypothesis testing to ecological and behavioral research is in making a clear distinction between *hypotheses* and *predictions*. A hypothesis is a suggested explanation for an observation, phenomenon, or scientific problem that can be tested by further investigation. We can think of a hypothesis as an educated guess of some outcome based on scientific logic. The best use of hypotheses is when they are developed directly from theory, that is, scientific hypotheses quickly become entrenched within some theoretical framework and then are tested over and over again with different model species and systems until they are accepted as fact. Implicit within scientific hypotheses is that they make predictions than can be tested with experimental or

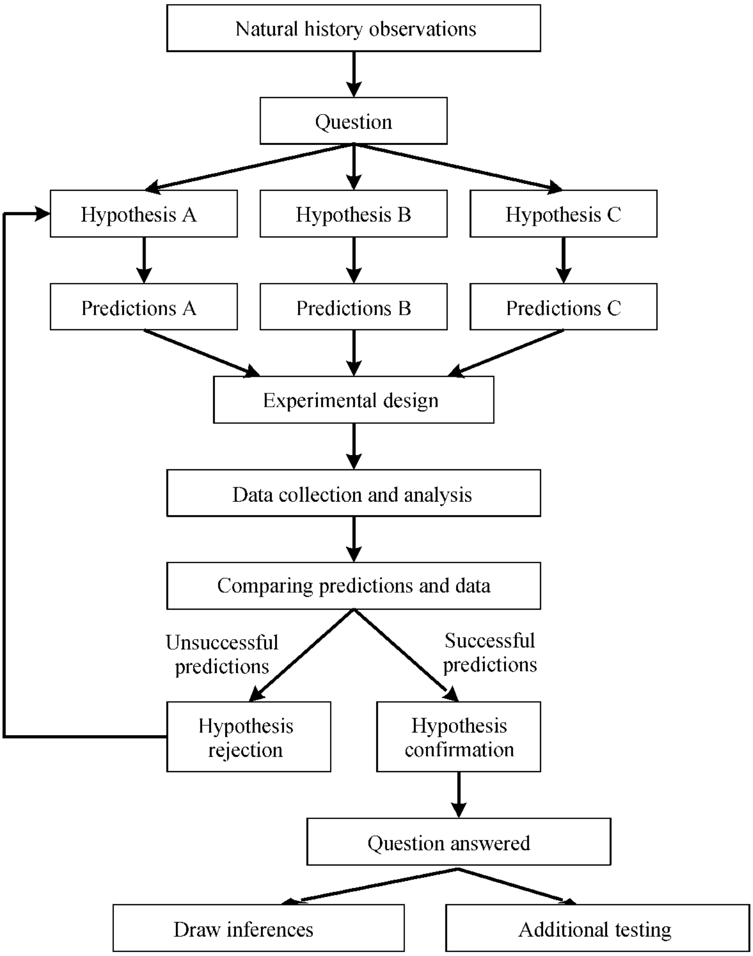

Fig.1 A schematic illustration of the scientific method with the use of multiple working hypotheses as recommended by Chamberlin (1897) and Platt (1964)

observational data. For example, there are three hypotheses for sex-biased natal dispersal in mammals, avoidance of resource competition, avoidance of reproductive competition, or avoidance of inbreeding. These hypotheses are well ingrained in the scientific literature (e.g., Wolff, 1993; Pusey and Wolf, 1994). For research purposes, each of these hypotheses makes a series of predictions that can be tested experimentally. We briefly provide one example for each. The resource competition hypothesis predicts that resources (such as food) are limited and thus if food is added to the natal site, juveniles will not disperse. The reproductive competition hypothesis predicts that adult males will expel sons and adult females will expel daughters to avoid competition for breeding opportunities. The inbreeding avoidance hypothesis predicts that opposite sex relatives (parents) are the stimulus for dispersal of each sex juvenile. The hypotheses are indirectly validated by testing these predictions with experimentation. In this situation, advancing science does not come from gathering data that support any one of these hypotheses, rather rejecting the two alternative hypotheses.

Another classical example of the application of strong inference was used by Tinbergen et al. (1963) in which they demonstrated that removal of egg shells from gull nests decreased predation by crows. The key point of this study was that Tinbergen rejected two alternative hypotheses, that egg shells attracted parasites (which infected the chicks) or that chicks might cut themselves on the shells. A caution is in order here because we should always remember that our list of alternative hypotheses may not be complete because scientific knowledge is never complete. We may have completely overlooked a factor that is the primary explanation for our particular question. A simple example would be to test the two alternative hypotheses that population changes are caused by changes in reproduction or in mortality, when in fact it could be that the changes are due to immigration or emigration.

Science is an on-going process and our ecological world view is based on many particular studies on many species and ecosystems. When we set up a general hypothesis, it is important to keep in mind that failure to support or reject predictions from hypotheses does not necessarily falsify the hypotheses overall. We may find that a specific study area or species population does not fit the general hypothesis. The hypotheses might still be valid for many other situations and that is why it is important to clearly articulate assumptions and predictions of hypotheses for the specific study question and design. Understanding the limits of the original observation, or purpose for doing the study, will help to determine what inferences can ultimately be drawn from accepting or rejecting a given hypothesis. We set limits to hypotheses by finding exceptions, so that for example we might find that moisture limits plant growth in arid ecosystems but not in temperate rainforests. Every hypothesis has its limitations, and the art of doing science is to recognize these limits within the generality of cause and effect that defines a scientific approach.

We close this essay with a brief discussion on how we choose research questions. There is a clear distinction here between pure and applied research in ecology and behavior. In applied problems we are often given a species or ecosystem situation that needs resolution. For example, panda populations may be declining in a particular region, and we need to find out why. In pure research, on the other hand, we should be question-driven and not species- or methods-driven. Some species or ecosystems are not good for asking particular research questions and thus selection of the research question should always precede that of the species or ecosystem when one is doing pure research. Also, the association between the natural history of the species and the research objectives must be well understood and clearly articulated to justify using the chosen species as a model for testing theory.

A model species must be just that, a good model for testing some biological phenomenon that is a natural and integrated part of the species' natural history. One certainly would not use an asocial species for studying cooperative behavior, a desert mammal for studying adaptations to cold, or a homeothermic mammal for studying hibernation. Good research involves a full understanding of the natural history of the model species such that the investigator knows what inferences to draw with the conclusions. In other words, a researcher must know the universe of the study, i. e., the results can be applied to what larger group of organisms? North American ground squirrels appear to be good models for alarm calls given by all mammals, but are not good models for testing hypotheses about infanticide. Meadow voles are good models for female territoriality, but not for monogamy. Wolves are good models for studying cooperative hunting but not promiscuity. Thus, a species' natural history such as what it eats, where it lives, how does it avoid predation, what is its social system, does it migrate, what is its life (reproductive) history, and so forth are all important in selecting a model species and knowing what inferences can be drawn from the results.

In conducting research that makes a contribution to science, we are reminded of the four essential components to all research, theory, data, evidence, and mechanism. Good science requires all four components. In some cases, especially in ecological and behavioral research, the mechanism might not always be known, but at least it must be something feasible and testable. Our intention in this essay is one of consciousness raising to new, young (and some not-so-young) investigators to revisit the scientific method, make good use of hypothesis testing in research design, and have a good understanding of natural history in the selection of model species. Although we emphasized ecological research in this essay, hypothesis testing and the scientific method are applied to all aspects of science and made some of the greatest strides in the molecular sciences (e. g., Platt, 1964). Science is advancing so rapidly in the 21st century that it is important for everyone to get on board and take advantage of some of these basic scientific approaches for the advancement of ecological science as well as individual careers.

Acknowledgements We thank Guiming Wang for his encouragement in writing this essay and for assisting with gathering data on hypothesis testing from recent issues of Acta Zoological Sinica.

References

Chamberlin TC, 1897. The method of multiple working hypotheses. J. Geology 5: 837 – 848 (reprinted in Science 148: 754 – 759 in 1965).

Krebs CJ, 2000. Hypothesis testing in ecology. In: Boitani L, Fuller TK ed. Research Techniques in Animal Ecology. New York: Columbia University Press, 1 – 14.

Platt JR, 1964. Strong inference. Science 146: 347 – 353.

Pusey AE, Wolf M, 1994. Inbreeding avoidance in animals. Trends Ecol. & Evol. 11: 210 – 206.

Tinbergen N, Broekhuysen GJ, Feekes F, Houghton JCW, Kruuk H, Szulc E, 1963. Egg shell removal by the black-headed gull *Larus ridibundus* L: a behavior component of camouflage. Behaviour 19: 74 – 117.

Wolff JO, 1993. What is the role of adults in mammalian juvenile dispersal? Oikos 68: 173 – 176.

Wolff JO, 1995. On the limitations of species-habitat association studies. Northwest Science 69: 72 – 76.

Wolff JO, 2000. Reassessing research approaches in the wildlife sciences. Wildl. Soc. Bulletin 28: 744 – 750.