

**Sara M. Hiebert**

*Advan Physiol Educ* 31:93-96, 2007. doi:10.1152/advan.00034.2006

**You might find this additional information useful...**

---

This article cites 11 articles, 3 of which you can access free at:

<http://ajpadvan.physiology.org/cgi/content/full/31/1/93#BIBL>

This article has been cited by 2 other HighWire hosted articles:

**Teaching simple experimental design to undergraduates: do your students understand the basics?**

S. M. Hiebert

*Advan Physiol Educ*, March 1, 2007; 31 (1): 82-92.

[Abstract] [Full Text] [PDF]

**Are chicken embryos endotherms or ectotherms? A laboratory exercise integrating concepts in thermoregulation and metabolism**

S. M. Hiebert and J. Noveral

*Advan Physiol Educ*, March 1, 2007; 31 (1): 97-109.

[Abstract] [Full Text] [PDF]

Medline items on this article's topics can be found at <http://highwire.stanford.edu/lists/artbytopic.dtl> on the following topics:

Physiology .. Galliformes

Research Methods .. Statistical Analysis

Veterinary Science .. Chickens

Updated information and services including high-resolution figures, can be found at:

<http://ajpadvan.physiology.org/cgi/content/full/31/1/93>

Additional material and information about *Advances in Physiology Education* can be found at:

<http://www.the-aps.org/publications/advan>

---

This information is current as of June 11, 2007 .

# The strong-inference protocol: not just for grant proposals

Sara M. Hiebert

Department of Biology, Swarthmore College, Swarthmore, Pennsylvania

Submitted 13 May 2006; accepted in final form 13 October 2006

**Hiebert SM.** The strong-inference protocol: not just for grant proposals. *Adv Physiol Educ* 31: 93–96, 2007; doi:10.1152/advan.00034.2006.—The strong-inference protocol puts into action the important concepts in Platt's often-assigned, classic paper on the strong-inference method (10). Yet, perhaps because students are frequently performing experiments with known outcomes, the protocols they write as undergraduates are usually little more than step-by-step instructions for performing the experiment. The strong-inference protocol, however, includes an explicit statement of possible experimental outcomes and the interpretation that would follow from each. This approach encourages thorough planning, enhances the efficiency of experimental designs, and increases the power of statistical analysis by explicitly stating a priori predictions as well as the statistical methods that will be used to test them. A sample protocol for an experiment investigating temperature-metabolism relations in chicken embryos is provided to illustrate the important components of the strong-inference protocol and to encourage instructors to incorporate this powerful research tool into undergraduate laboratory courses.

experimental design; multiple hypotheses; teaching

FOR MOST STUDENTS, writing a protocol means copying the instructions from their laboratory manual into their laboratory notebook, embellished with a short description of the purpose of the experiment and the student's hypothesis about the outcome. Arguably, such protocols are missing the most powerful component of any protocol: a list of the possible outcomes of the experiment and how each of those outcomes would be interpreted in light of the question that the experiment is designed to answer.

Why are the outcomes and interpretations crucial components of an effective protocol? In his often-cited paper on the subject of strong inference, Platt (10) outlined a series of steps that should be followed systematically in scientific investigation:

1. Devising alternative hypotheses;
2. Devising a crucial experiment (or several of them), with alternative possible outcomes, each of which will, as nearly as possible, exclude one or more of the hypotheses;
3. Carrying out the experiment so as to get a clean result;
4. Recycling this procedure, making subhypotheses or sequential hypotheses to refine the possibilities that remain; and so on. Any college student would recognize this series of steps as the scientific method (9). Few, however, are called on to make full use of strong inference because they have not been taught the simple steps that turn a glorified recipe into a powerful tool for inquiry. A list of possible outcomes and interpretations will stop a poor experiment in its tracks, pointing the way to important refinements that will keep the researcher from wasting effort on an inefficient experimental

design (4–6, 14). The outcomes and interpretations also help to ensure that the student/scientist keeps in mind multiple hypotheses rather than a single pet hypothesis for which he or she has developed undue affection (2).

Platt was not the first to propose many of the ideas in his 1964 paper, but he gave voice to them in a colorful way that brought them to the attention of, and subsequently influenced, researchers in a variety of fields (3). Although some have argued with some of Platt's points (3), it is generally accepted that the practices of identifying multiple hypotheses (when available) and planning data collection with statistical analysis and interpretation in mind will result in a higher quality of research, even though the words "strong inference" are not always applied to this method (14). Investigations planned in this way are characterized by efficient use of time, money, and animals (4–6, 14) because they reduce the chances of drawing incorrect conclusions (4, 5, 14), increase the statistical power derived from a given sample size (4, 6, 11), and ensure that appropriate conclusions can logically be drawn from the data (14). Good strong-inference protocols are an important element of successful grant writing in experimental sciences. Why, then, would we not want to make this tool a part of every student's scientific training? We do not need to wait until graduate school to introduce it (14). My purpose is to encourage instructors to incorporate strong-inference protocol writing into the undergraduate curriculum as early as possible.

## *A Strong-Inference Protocol for a Physiology Experiment*

The strong-inference protocol is a concise document that summarizes why and how an experiment will be done, describes how the data will be analyzed and interpreted, and lists potential difficulties in experimental design or in drawing logical conclusions from the data. The general form of the strong-inference protocol is applicable to any experiment. Below, a general description of each section of the protocol is provided. In the *Appendix*, more specific content is described for the case of a student-designed experiment described in the companion article "Are chicken embryos endotherms or ectotherms?" (7). In this experiment, students investigated the thermoregulatory mode of chicken embryos by measuring respiration rates ( $\dot{V}O_2$ ) at different ambient temperatures and comparing their results with the signature relations between temperature and metabolism in endotherms and ectotherms.

## *Sections of the Protocol*

**Introduction and background.** In one or two paragraphs, this section explains why the study is interesting and how it relates to existing knowledge. Students are asked to cite one to five published papers in this section. Unlike the introduction to a scientific paper, it is not meant to be an exhaustive review. This section should also explain how the data that will be

Address for reprint requests and other correspondence: S. M. Hiebert, Dept. of Biology, Swarthmore College, Swarthmore, PA 19081-1390 (e-mail: shieber1@swarthmore.edu).

collected relate to the physiological variables of interest and to the hypotheses being addressed.

**Methods.** This section describes the experimental procedures, including an explanation of controls, a timetable of measurements, and salient details of the protocol that are not found in the laboratory manual because they are features of the students' own experimental design. I encourage students to omit standard details of the procedure such as how to use the equipment by referring instead to the appropriate sections of the laboratory manual. In my experience, students do not respond favorably to tasks that seem like busywork; for this reason, requiring them to copy procedures that are already printed in the manual tends to disengage students from the intellectually important features of the strong-inference protocol.

**Data analysis.** In a few sentences, students are asked to name the statistical test that they will use and exactly which groups of measurements will be used in the analysis. A statement such as "We will use a *t*-test to analyze our results" is inadequate because it does not provide evidence that the student understands how the test will be applied.

**Outcomes and interpretations.** This section contains one entry for each possible outcome of the experiment, followed by a brief description of what the student could conclude if this outcome occurred. One of the greatest conceptual difficulties students have with this section is distinguishing between an outcome and an interpretation. Outcomes are summaries of results, whereas interpretations represent the way in which the results address the question that the experiment was designed to answer. Students are encouraged to mention how probable they consider a particular outcome or interpretation and to provide evidence to support their assessment from the literature or from a priori reasoning. They may also briefly suggest further tests that might resolve ambiguities or further corroborate the result.

**Caveats.** This section should include a brief description of untested assumptions and identify a study's shortcomings. Although students are often hesitant to expose any weaknesses in their design, the ability to identify caveats indicates clarity of thought. Few experiments are perfect.

### Classroom Implementation

In the laboratory section of my intermediate-level Animal Physiology course, students participate in two 3-wk modules, in each of which the 12 members of the laboratory section design, as a group, an experiment they will perform together to answer a physiological question that I pose. These two 3-wk modules are followed by a 7-wk independent project, in which students work in small groups to address a question of their own choosing. At the conclusion of each module or project, students are required to write a laboratory report in the style of a scientific paper.

In the first year that I incorporated the strong-inference protocol into the laboratory for this course, I discovered that students needed guided practice before they could produce an effective strong-inference protocol on their own. This was especially evident in the *Data analysis* and *Outcomes and interpretations* sections, even after we had discussed the ex-

perimental design together in depth (but without specific reference to how the different elements of the design would be represented in the strong-inference protocol). A method that produced much better results in subsequent years was to design the experiment in the first module as a class and then talk very specifically about how this experiment would be addressed in each section of the protocol. In this discussion, we considered wording that might be appropriate for each section. The significantly higher scores that students earned on the protocol for the first module in the next 4 yr ( $17.7 \pm 0.3$  compared with  $15.2 \pm 0.2$  out of a total possible score of 20 in the first year) are consistent with this interpretation (one-way ANOVA,  $F_{4,111} = 13.2$ ,  $P < 0.0001$ ).

For the second module, we design the experiment together but I leave to students the task of figuring out the essential content of each section of the protocol as it applies to this experiment. For the independent project, students design the experiment and write the protocol themselves after discussing their ideas with the instructor. Protocols are graded on an increasingly stringent rubric as the semester progresses, such that a particular kind of mistake on the first protocol might elicit a comment but not a point deduction, whereas the same error on a later protocol would incur a point penalty. Thus even a slightly better score later in the semester reflects a substantial improvement. Even with this rising scale of expectations and with progressively less instruction, protocol scores (out of a total of 20 possible points) increase by an average of 0.5 points with each successive protocol (repeated-measures ANOVA,  $F_{2,179} = 26.3$ ,  $P < 0.0001$ ). This progression suggests that students can learn and are later able to apply the general form of the strong-inference protocol to specific experiments, with the proviso that they benefit greatly in the early stages from seeing a specific example on which to model their own protocols.

Aside from technical questions on how to use the statistical software, students in the Animal Physiology course almost never ask questions about data analysis and logical interpretation after they have written their strong-inference protocol. Rather, our discussions regarding interpretation focus almost exclusively on the larger questions of why chicken embryos might be ectothermic and what implications ectothermy has for the life and energy economy of the embryo and the incubating parent. The strong-inference protocol provides a basic outline of the contents of each section of their laboratory report, beginning with the introduction and continuing with the methods, results, and the first paragraph or so of the discussion. Having the protocol in hand as a road map for the laboratory report appears to have the added advantage of reducing the anxiety associated with having to produce the entire laboratory report at once. If, after they collect their data, students forget the reasoning behind the experiment or need to be reminded of how they will use their data, they have only to consult their own protocol.

Adding a data analysis and an outcomes and interpretations section to student protocols means that instructors will need to read and provide feedback on a slightly longer document. However, in my opinion, the ultimate benefit to students weighed against the relatively small increase in an instructor's reading time gives this section one of the best per-word pedagogical values in the laboratory classroom.

### Appendix: Specific Content for the Chicken Embryo Metabolism Experiment

**Introduction and background.** A chicken embryo will eventually grow into an endothermic adult and must necessarily start life as an ectothermic zygote, but the point at which the transition takes place is intuitively ambiguous and could be argued either way from first principles. To put the methods into context, this section must explain how endotherms and ectotherms differ in their metabolic responses to environmental temperature and must state, in general terms, how this difference will be used in the experiment to answer the question of whether chicken embryos are ectotherms or endotherms (see Fig. 1 in Ref. 7). Specifically, two features of the relation between temperature and metabolic rate differ between endotherms and ectotherms: the shape of the curve and the range of absolute values for  $\dot{V}O_2$  (an index of metabolic rate), which is substantially higher overall in endotherms than in ectotherms. Finally, students need to explain how what they will measure ( $\dot{V}O_2$ ) is related to metabolic rate, the physiological function that they are investigating in this experiment (see Ref. 7).

**Methods.** This section should include details that could vary from experiment to experiment, such as the temperatures at which measurements will be made, where and for how long eggs will be placed to equilibrate to the temperature treatments, the volume of the test chamber (respirometer), and the duration of each measurement period.

**Data analysis.** The content of this section will, of course, vary with the experimental design that students choose. It might read as follows: "A two-sample *t*-test will be used to compare  $\dot{V}O_2$  between experimental and control eggs after a 90-min equilibration at 23 and 38°C, respectively."

**Outcomes and interpretations.** For an experiment in which  $\dot{V}O_2$  is measured at 23 and 38°C, there are three possible outcomes.

**OUTCOME 1.**  $\dot{V}O_2$  for eggs maintained at 38°C is significantly greater than that for eggs incubated at 23°C.

These results suggest that the embryos are ectothermic, because  $\dot{V}O_2$  is directly proportional to ambient temperature. A similar outcome could be obtained if all three of the following conditions are true: 1) the embryos are endothermic, 2) 38°C lies above the thermal neutral zone, and 3) 23°C lies within or just below the thermal neutral zone (see Fig. 4 in Ref. 7). However, this interpretation is considered highly unlikely because 38°C is the normal incubation temperature provided by the hen. If the embryos were endothermic with a body temperature set point of <38°C (*condition 2*), incubation by the hen would cause the developing chicks to expend energy to cool themselves to the body temperature set point. Because this would consume energy otherwise available to fuel development, an incubation temperature greater than the body temperature set point is unlikely to be favored by natural selection.

**OUTCOME 2.**  $\dot{V}O_2$  of eggs maintained at 38°C is significantly less than the  $\dot{V}O_2$  of eggs incubated at 23°C.

These results would suggest that the embryos are endothermic, because  $\dot{V}O_2$  is inversely proportional to ambient temperature (see Fig. 1 in Ref. 7).

**OUTCOME 3.**  $\dot{V}O_2$  of eggs maintained at 38°C is not significantly different from the  $\dot{V}O_2$  of eggs incubated at 23°C.

Because preliminary measurements of egg cooling show that the equilibration times we chose are adequate to cool an ectothermic embryo to 23°C, it is unlikely that this result could be due to a failure of the embryos inside the eggs to cool to 23°C if they are ectothermic. A second possibility is that the embryo is endothermic and that 23 and 38°C both lie within the thermal neutral zone (Fig. 3 in Ref. 7), but this is considered unlikely because the thermal neutral zones of similarly sized endotherms (~16 g on incubation *day 16*) typically span a much smaller range of ambient temperatures (see Fig. 1 in Ref. 7). A third possibility is that a 14°C temperature difference is insufficient to provoke a detectable change in  $\dot{V}O_2$ . This is also considered unlikely because the typical  $Q_{10}$  values for aerobic metabolism predict that a two- to threefold increase in metabolic rate should accompany a change in body temperature of 10°C (1) in an ectotherm. A fourth possibility is that there is so much individual variation in  $\dot{V}O_2$  that sample sizes are too small to detect a statistically significant effect. In this case, the experiment could be repeated with additional eggs. A fifth possibility is that the embryos are only partially endothermic. In this scenario, the embryos have some capacity to slow the rate of cooling but not enough to cause an increase in metabolic rate (7). Further experiments in which younger, prethermoregulatory embryos are tested or in which embryos are subjected to more prolonged bouts of cooling could address this possibility (13).

**Caveats.** First, the conclusions apply only to embryos of the age actually tested in the experiment. Second, we must assume that all embryos are healthy and at the same developmental age and that the vendor shipped embryos at the age stipulated in the order. (On one occasion, we received embryos that were 6 days old rather than 16 days old, a fact discovered when a student accidentally dropped and broke one of the eggs.) Third, the choice of temperatures reflects the reasoning that if the chick were an endotherm, the normal incubation temperature of 38°C would likely lie within the thermal neutral zone and the lower temperature (23°C) would likely be less than the lower critical temperature, which is the lower limit of the thermal neutral zone. If this assumption is incorrect, an endothermic embryo could show a positive relation between temperature and  $\dot{V}O_2$  that might be interpreted as an ectothermic response. However, the range of values measured for  $\dot{V}O_2$ , typically much higher in endotherms than in ectotherms, should prevent such a misinterpretation (see Fig. 1 in Ref. 7).

### ACKNOWLEDGMENTS

Primary thanks are due to Irving Zucker, who first introduced the author to the benefits of the formal strong-inference protocol in his beautifully managed lab at University of California-Berkeley. Support for writing and publication was provided by a Lang Fellowship from Swarthmore College and by Fritz Geiser and his laboratory at the University of New England.

### REFERENCES

1. **Bartholomew GA.** Body temperature and energy metabolism. In: *Animal Physiology: Principles and Adaptations* (4th ed.), edited by Gordon MS, Bartholomew GA, Grinnell AD, Jørgensen CB, White FN. New York: MacMillan, 1982, p. 333–406.
2. **Chamberlin TC.** The method of multiple working hypotheses. *Science* 148: 754–759, 1965.
3. **Davis RH.** Strong inference: rationale or inspiration? *Perspect Biol Med* 49: 238–249, 2006.

4. **Festing MF.** Guidelines for the design and statistical analysis of experiments using laboratory animals. *ILAR J* 43: 244–58, 2002.
5. **Festing MF.** Principles: the need for better experimental design. *Trends Pharmacol Sci* 24: 341–345, 2003.
6. **Festing MF.** Reduction of animal use: experimental design and quality of experiments. *Lab Anim* 28: 212–221, 1994.
7. **Hiebert SM, Noveral J.** Are chicken embryos endotherms or ectotherms? A laboratory exercise integrating concepts in thermoregulation and metabolism. *Adv Physiol Educ* 31: 97–109, 2007.
8. **Institute of Laboratory Animal Resources, National Research Council.** *Guide for the Care and Use of Laboratory Animals*. Washington, DC: National Academy, 1996.
9. **Kinraide TB, Denison RF.** Strong inference: the way of science. *Am Biol Teacher* 65: 419–424, 2003.
10. **Platt JR.** Strong inference. *Science* 146: 347–353, 1964.
11. **Sokal RR, Rohlf FJ.** *Biometry* (2nd ed.). New York: Freeman, 1981.
12. **Van Wilgenburg H, van Schaick Zillesen PG, Krulichova I.** Sample power and ExpDesign: tools for improving design of animal experiments. *Lab Anim* 32: 39–43, 2003.
13. **Whittow GC, Tazawa H.** The early development of thermoregulation in birds. *Physiol Zool* 64: 1371–1390, 1991.
14. **Zolman JF.** Teaching experimental design to biologists. *Adv Physiol Educ* 22: 111–118, 1999.

