How Do We Know What We Know? Teaching About the Scientific Process in Undergraduate Classes

RANDY J. NELSON
Professor and Chair, Department of Neuroscience
College of Medicine

Over the two decades of my professional career I have taught a variety of undergraduate and graduate courses in Behavioral Biology, and all of them have had an identical Lecture #2. In common with most courses, the first meeting is always an introductory lecture that outlines the learning goals of the class, the steps that the students must take over the course of the term to achieve those goals, and some overview of the topic—nothing out of the ordinary. But regardless of the course content area, Lecture #2 of every course I teach reviews what John Platt refers to as "Strong Inference".[1] Despite what some of my repeat students might think, the goal of this practice is not to reduce the total number of lectures that I must prepare for each course; rather, my goal is to teach the students the process of the scientific enterprise before I teach the "facts."

Strong Inference

In 1965 John Platt published an article in the journal *Science* entitled "Strong Inference." Platt declared that all scientific enterprises were not making equivalent progress. Some fields such as molecular biology and high energy physics were making rapid and impressive progress in solving the complex questions of the day, whereas other fields seemed to be stagnant. Platt argued that the difference in progress among scientific disciplines did not reflect the inherent complexity of the problems under study in each field, but rather reflected the consistent application of inferential logic trees or what he termed, "strong inference."

Essentially, the components of strong inference primarily reflect Francis Bacon's inductive inference method. There are four formal steps in this method: (1) devising alternative hypotheses to explain a phenomenon, (2) devising a crucial experiment (or series of experiments) that will exclude one or more of the competing hypotheses, (3) conducting the experiment to get a clear outcome, and (4) recycling the process to refine the remaining alternative explanations. Step 2 is the crucial move that is often overlooked in the scientific method—exclusion of explanations is the sausage grinder of the scientific enterprise.

At this point, I generally provide a number of examples in the content area of the course to the students. To make scientific progress, we must climb the logic tree in the interpretation of our data. At the first fork, if our experiments are well-designed and conducted, then the experimental results choose which alternative branch to follow. Of course, inductive inference is less certain than deductive reasoning made famous by Sherlock Holmes because inductive reasoning involves reaching into the darkness of ignorance to discover the "truth" underlying nature—at least the truth underlying nature under a very specific set of conditions. Platt believed that inductive reasoning was so powerful in understanding the complexities of nature that he attached the moniker "strong" to inference. He emphasized that science only advances by disproofs, that is, by rejecting sufficient alternative hypotheses until one develops a hypothesis that cannot be disproved despite numerous attempts. This is an important distinction that most students and many scientists fail to appreciate. Science only advances by disproofs. One cannot prove anything with the scientific enterprise.

Many students (and quite a few scientists) believe that a great theory or hypothesis is one that can explain everything. As Platt emphasizes, not a small number of eminent scientists become married to their pet hypothesis. They never design a critical experiment to attempt to falsify their hypothesis, but rather dither throughout their careers providing additional "support" for their hypothesis. In some cases, a hypothesis or theory is put forward that cannot be critically tested. Consequently, such theories almost force us into a "take it or leave it" response, and we opt to take it or leave it on the basis of being able to make a leap of faith or not. Freud's theory of the unconscious mind and Marx's theory of economics are examples of wide-ranging theories that, from my perspective—and Platt's—are fundamentally flawed because there are no experiments that could be conducted to rule out their basic hypotheses. Freud believes that as young child I wanted to have sex with my mother. "No I didn't," I protest, but neither of us can disprove our positions. What Freud could do is respond that I've repressed this desire, but that response is not persuasive because it depends on my accepting his general theory of the unconscious, and it is precisely that theory that is in dispute between us. The response also underlines the point that it is impossible to falsify Freud's hypothesis about the Oedipus complex. Consequently, I count myself among those unable to make the leap of faith necessary to accept his theory.

My goal is to teach the students the process of scientific enterprise before I teach the "facts."

Some folks believe that Darwin's theory of evolution through natural selection suffers from the same problems with tautology and nonfalsifiability. Why do birds have wings? Because having wings is adaptive—how do I know this? Because wings have evolved through the process of natural selection and all extant traits are adaptive. But Darwin has shown us a way out of the circular logic in his original book on the topic by stipulating what would count as evidence against his theory. Essentially, he states that if any adaptation can be demonstrated to have evolved solely to improve the fitness of another individual, then his theory of evolution through natural selection would be falsified. Nearly 100 years later, Darwin's views of individual fitness were modified, not falsified, when it was shown that altruistic behaviors appeared to have evolved to benefit other individuals. Through a brilliant analysis, W.D. Hamilton was able to rule out a number of competing hypotheses to determine that altruistic behaviors could evolve if the benefactor was close kin. He showed that a specific behavior could appear altruistic at the level of the individual, but selfish from the perspective of the gene. For example, if a female ground squirrel stands on her haunches and gives an alarm call in response to an attacking hawk and that alarm call draws attention to her resulting in her death, then you might think that such a maladaptive sacrificial trait could not have arisen through natural selection. Such a trait would seem to falsify Darwin's theory by using his own standard of evidence, viz., the existence of a trait that evolved to improve the fitness of another individual. However, this conclusion assumes that the individual, rather than the constellation of genes that comprise the individual, reflects the unit of selection.

Female ground squirrels (and other altruistic species) only engage in such altruistic acts when close kin are present. This provided some insight into the unit of selection. A female ground squirrel is 100% genetically related to herself, but shares some proportion of her genes with kin and few or no genes in common with strangers (including mates). If a female detects a predator in the presence of unrelated squirrels, then she never gives an alarm call. However, if she is around kin, then she behaves as if she performs a complicated calculation of relatedness. Imagine that two daughters (each related to her by 50% of their genes on average), a sister (50% related), a cousin (25% related), and an aunt (25% related) are present, then 200% of her genes (50% + 50% + 50% + 25% + 25%) are represented in the saved individuals. In that situation, she would give the alarm call and put herself at risk. Although her behavior appears altruistic, on the genetic level she is acting in the best interest of her genes' fitness. This new type of selective force, termed kin selection, modified Darwin's original hypothesis, but did not change our fundamental understanding of biology, since the apparently sacrificial behavior increases the chances that the female's genes will survive. Kin selection provides a biological foundation for the ancient Arabic proverb, "me against my brother, me and my brother against my cousins, me and my cousins against the world."

Strong inference moves us closer to the ultimate "truth" underlying nature, and moves us there faster than any other method. I ask my students to go to the library (yes, the big building with all those books) to choose 5-10 scientific papers of their choice and to count the number of papers in which the authors propose in the introduction alternative hypotheses and design an experiment to exclude as many possible competing hypotheses as possible. If such a statement is not found, then another useful exercise is to have the students think about what hypothesis the experiment in the paper disproves. For graduate students, it is useful to ask them to think when listening to a professional seminar, "what hypothesis is being disproved in the study" or when hearing a novel hypothesis, "what critical experiment or experiments would need to be conducted to attempt to rule out this hypothesis."

"Fifty percent of the information in the textbook is wrong; our problem is that it will take some time to know which 50%."
During my Lecture #2, I tell the students that much of the information in my college freshman biology textbook was simplified to the point of being inaccurate, and in many cases was proven wrong with additional testing. The same is true for current textbooks. In my course, "Hormones and Behavior," I am the author of the textbook so this is an especially painful revelation for me to provide to the students.

The students are taking my class to learn the "facts" regarding the interaction between hormones and behavior. To help the students understand this principle, I assign them two scientific papers that describe equivalent experiments on the same species, but report opposite, contradictory results. For example, one paper might suggest that estrogen is necessary to induce mating behavior in female rhesus macaques, whereas the other paper reports that estrogen has no effects on mating behavior in females of this species. Or, I assign a paper that reports that the mating behavior of male mice wanes 2–3 weeks after castration and withdrawal of testosterone, whereas a second paper reports that male mice continue to mate for months in the absence of testosterone. The students hate this exercise. Which results are true, the students ask. Great question, I respond. Which results do they think are true? Which results are true, indeed? My point in asking these questions is not to come to a definitive answer but rather to show the students that the declarative statements they read in most textbooks are attempts by the authors to make sense of messy, often contradictory information in the primary literature. We textbook authors leave out all the qualifications which make a declarative statement generally true, and, consequently, over time, as more studies are conducted, roughly half of those statements turn out to need substantial revision.

The more general point is that our students need to learn to be aware of the limitations of declarative statements and of the necessity of evaluating the primary evidence before they can take up reasonable positions on such matters as climate change, alternative fuels, endangered species, or nutritional recommendations. I feel that Lecture #2 has succeeded when I learn that a former student, who hears conflicting news reports that estrogen benefits and harms the hearts of women, responds by looking up the original studies online and asking, "What hypothesis was each study designed to disprove?" That student is now able to decide for herself what is "true."

True Facts and Weak Inferences

When I was a kid working on my grandfather's farm during the summer, I often remarked about the heat. When I mentioned to my taciturn grandfather, "It sure is hot today," he always responded, "That's a true fact, son." It is easy to teach facts, whether true or not. It is easy for students to learn facts sufficiently long to spit them out on an exam. However, I believe that we do a great disservice to our students by simply asking them to memorize facts. Not only is it remarkably easy to forget facts, but facts make sense only in relation to more general hypotheses that are subject to disproof. Students need to understand the proper methods to evaluate information. It is more than critical thinking. Importantly, it is a skill that can be taught.

When I began graduate school, I worked with a prestigious researcher who was a member of the National Academy of Science and had virtually established my field as a formal discipline. As a graduate research assistant, I spent four hours each morning in a windowless room with a stack of data sheets, a Monroe calculator (which could spit out simple statistics), and my advisor who slowly replaced the oxygen with the aromatic smoke from his pipe. He would lean back and say something like, "I wonder whether there was a difference in the number of growls shown by the females treated with testosterone early and late compared to males. Run a statistical analysis." As he puffed on his pipe, I'd add up the data detailing the number of growls produced by each type of animal and report back in 15 minutes or so that yes, there was a significant difference. He would nod approvingly and puff swirls of white smoke and after 15 or 20 more minutes, he'd say something like, "I wonder whether there was a difference in the number of tail wags among these groups? Do another stat test." And so it would go until lunch, during which I could replenish my lungs with oxygen. We'd repeat the process again from 1 to 5 pm each afternoon. It was maddening. Where were his ideas coming from? Where was I going to get these ideas? How was I going to become a scientist? When was he going to start teaching me how "to wonder whether"? To me, this felt like I was being taught to do research like an apprentice mechanic being taught his craft by washing cars (or more apropos, how to become a professor and give lectures by being a teaching assistant who made and graded exams)! But I was not being taught how to determine a true fact because I was not learning anything about the relations among facts, hypotheses, and falsifiabilty.

During this first year in grad school, I attended a proseminar led by four prominent researchers in my field of behavioral neuroscience, although at the time, it was called biopsychology. One faculty member introduced us to John Platt and Strong Inference. As he recounted all the benefits of this method, my advisor blurted out, "Well, I must be doing weak inference science all these years because I've never gone through such a process in my entire career." Although his remarks were met with laughter because he was so eminent and obviously knew how to do rigorous science, they were also a turning point for me. I understood immediately that I could learn to do science by using the strong inference process and that I would never learn the intuitive, so-called “weak inference” process advocated by my advisor. I see now that my advisor likely did some version of inferential logic in his head, but he could not teach this method. Eventually, I switched labs during graduate school and embraced strong inference as my scientific and teaching philosophy.

One of my teaching goals is to bring this process to the next generation, even if it means providing them with the tools to challenge the facts I teach in the classroom. Nothing is as rewarding as a teacher as having a student stop my lecture with a “But professor, aren’t there alternative interpretations of those results?”

References

Darwin, Charles. *On the Origin of Species by Means of Natural Selection: The Preservation of Favoured Races in the Struggle for Life*. London: John Murray, 1859.

Hamilton, William D. “The Genetical Evolution of Social Behavior.” *Journal of Theoretical Biology*. 7 (1964): 1 – 52.

Platt, John R. “Strong Inference.” *Science*. 146: (1964): 347 – 353.