HYPOTHESIS TESTING*

Students are taught the "scientific method" in middle school and again in high school. If you survey your students, all or nearly all will confirm that they learned the scientific method, although they may confess that they don't remember what it is or how to do it. Alison Gopnik (6 May 1999, Small Wonders, The New York Review of Books) offered a brilliant analogy between science education and baseball—Imagine if we taught baseball the way we teach science.

- Until they were twelve, children would read about baseball technique and occasionally hear inspirational stories of the great baseball players.
- They would answer quizzes about baseball rules.
- Conservative coaches would argue that we ought to make children practice fundamental baseball skills, throwing the ball to second base twenty times in a row, followed by tagging first base seventy times.
- Others would reply that the economic history of the reserve clause proved that there was, in fact, no such thing as "objectively accurate" pitching.
- Undergraduates might be allowed, under strict supervision, to reproduce famous historic baseball plays.
- But only in graduate school would they, at last, actually get to play a game.

If we taught baseball this way, we might expect about the same degree of success in the Little League World Series that we currently see in science performance.

One of the most difficult lessons for students to learn is that critical hypothesis testing should attempt to disprove ("falsify") hypotheses. A simple game can serve to illustrate the point. After you introduce this topic and explain about hypothesis testing, tell your students that you are going to have them play a guessing game to learn some methods of hypothesis testing. Tell them their job is to guess the "rules" you use to choose numbers that you'll write on the board. Then write three numbers on the board (e.g., 1, 2, 4). Tell them the game works as follows. A student can ask you whether a particular number fits your rule. You will answer truthfully, either "yes" or "no." If the answer is "yes" then you will add that number to the list on the board. If the answer is "no" then you will put that number in a separate place on the board. The questions you answer are analogous to the observations and experimental results generated by research, and the two list of numbers are the "data" the students have. Once a student thinks s/he knows the rule, s/he will "publish" it by describing the proposed answer to the class. The class will then discuss and evaluate the proposed rule. Anyone who disagrees with the "published" rule may publish one of her/his own by immediately proposing it to the class. Throughout these discussions, you will not say whether a proposed rule is correct, nor will you volunteer the true rule. These discussions are analogous to the process by which scientists communicate their ideas with others, and by which competing hypotheses are evaluated. If a proposed rule wins ready acceptance by the entire class, the proposer

* Thanks to Dave Tukey for teaching me the "numbers game" and to MarthaLeah Chaiken for telling me about Alison Gropnik's baseball analogy.

could win the Nobel Prize, but if the proposed rule is rejected by the class or disproved by a new "experiment" (by proposing a new number), the proposer's reputation is damaged and s/he must "sit out" for some short period (e.g. five minutes).

Note that the numbers in our example seem to follow an obvious rule: double the last number. In reality, the rule is: "positive integers" (1, 2, 3, 4, etc., but not zero and no fractions). In almost every case, students will pick an obvious rule and then propose numbers to you that fit the obvious rule. So, you are likely to have students propose 8, 16, 32, 64, etc. Let them keep going for a while; you keep saying "yes" and adding the next number to the list. Now, let's think about what the students are doing. In one respect, they are testing hypotheses. They have a particular rule (hypothesis) in mind, and they are proposing numbers (making predictions) that can be tested (you say "yes" or "no"). Unfortunately, they will never get any closer to identifying the rule because they are attempting to "prove" their hypothesis instead of trying to "disproving" it. What's the difference? All of their numbers FIT their hypothesis. You will say "yes" and add the next number to infinity (or at least until the bell rings).

To test the hypothesis critically, a student must propose a number that DOES NOT FIT the hypothesis. Why does that help? Because if the proposed number does, in fact, fit the rule, then the student has learned that the hypothesis is INCORRECT; the student has disproved the hypothesis. In contrast, even if your students double the last number one hundred times, they still haven't proven the "doubling" rule because other rules can still explain the numbers in the list. It's only by eliminating some possible rules (hypotheses) that we make any progress. So, in our example, if the student has in mind "double the last number" then s/he should propose a number that doesn't fit the proposed rule. For example, after 64, the student could propose 63. What happens when you then say, "yes, 63 fits!" Obviously, the rule is NOT "double the last number." In contrast, if the student proposes 128, you would also say "yes," but there are still other rules that also fit (e.g., a number larger than the last).

Interestingly, the process of hypothesis testing in animal behavior is exactly the same as in this number game (well, almost the same). If we continue to propose "tests" that simply confirm what we already know, we don't really get anywhere. It's only when we propose a test that does not conform to what we already know that we make any headway—we learn what explanations DO NOT APPLY to our behavior of interest.

Your students may come to the realization that testing to confirm hypotheses isn't getting them anywhere. If they do, be sure to point out their discovery and give them lots of reinforcement to this truly significant realization. If, however, they are still confirming hypotheses after several rules have been proposed, you can take a "time out" to discuss the differences in strategies of hypothesis testing with your students, then let them try out their new "method of disproof." (Note that this is the method advocated by noted philosopher of science, Karl Popper). Let them then resume guessing and you should begin to list lots of numbers that DO NOT fit your rule. Once they begin to get the idea, stop the game and resist the temptation (and the frequent requests) to tell them the rule. Remind them that this is science—we can NEVER find out the rule by asking a "higher authority."