A Novel Approach to Understanding the Process of Scientific Inquiry

Mark H. Anders
Columbia University, New York, NY, USA
manders@ldeo.columbia.edu

Abstract

Many of the basic concepts involved in the process of scientific inquiry can be represented by analogy to a simple game called Battleships. The same processes used in this child's game demonstrate what role hypothesis generation and testing play in the search for truth in nature. The analogy can also be extended to demonstrate how scientists can unknowingly be led to faulty conclusions about the structure and patterns of nature.

There is a rather simple children's game called Battleships that requires only two pieces of paper and two pencils. Each player grids his or her paper with an identical pattern of columns (A to Z) and rows (1 to 26). Each player is allowed a particular number of certain "ships," each of which is defined by connected squares. The bigger the ship, the more squares it occupies (e.g., an aircraft carrier would be seven connected squares, a battleship six, a cruiser five, and so on). A person might start the game by calling out for a square, say P-12. His or her opponent then responds by announcing whether this was a "hit" or a "miss." The first player to hit all the squares of their opponent's ships wins.

This game has much in common with the process of scientific inquiry. By analogy with the game, one might perceive nature's hidden truths as ships arrayed in some complex pattern, and it is the role of scientists to find out what these patterns are. Probing for patterns in nature is done by experiment. Some patterns are obvious from inspection, while others require intricate schemes to decipher.

In the early stages of the game of Battleships, calls are random (e.g., P-9, B-13, Q-12) until hits are recorded. This stage has similarities to inductive science; often given short shrift by the self-appointed guardians of the "scientific method." Inductivism can play an important, though not necessary, role early in the process of scientific discovery. For example, an earth scientist might come up with an idea, such as continental drift, by staring at a globe in the corner of his office and noting, as Alfred Wegener (1915) did, the fit of the African and South American coastlines--but you must first have a globe. Although having some observations is a good place to start from, science philosophers such as Karl Popper (1959) argue that one can generate a hypothesis independent of observations.

In Battleships, when hits are recorded they are followed by "what ifs"--what if there are connected squares that are part of a larger ship? Nearby coordinates are called out, and the first hint of a pattern starts to unfold. A true inductivist would continue to call out numbers randomly until all the squares are chosen and then identify a pattern to the ship placement. However, using deductive reasoning one can postulate, or hypothesize, a pattern based on the initial random probing, and test the hypothesized pattern by calling out B-14. If B-14 comes in as a hit, the player has achieved what Gottfried Leibnitz would call "the greatest commendation of a hypothesis (next to truth), if by its help predictions can be made even about phenomena not yet tried" (cited in Kneale, 1967, p. 29). Simple? Yes and no. In 1973, Nobel laureate Harold Urey suggested that a large meteorite hit the earth and killed the dinosaurs. Seven years later, Berkeley professor Walter Alvarez (Alvarez, Alvarez, Asaro, & Michel, 1980) finds high levels of the rare element iridium most commonly found in meteorites, at the paleontologically defined boundary thought to coincide with the extinction of the dinosaurs. *Quod erat demonstrandum*, right? Well, not really--many at the time questioned the meaning of the discovery by Alvarez. For example, could the iridium be concentrated at the existing boundary by ground waters unrelated to a meteorite impact? Here the tightly-scripted game Battleships must be manipulated to parallel the vagaries of scientific inquiry. Imagine, for example, that in your game of battleships the coordinates you read out pass through some unseen filter that, unknown to you, reroutes your choices, so that your call of B-14 is recorded as Q-19 and your call yields a hit when it should have produced a miss. Philosophy-of-science writers like Karl Popper prefer to avoid such gray areas. Rather, they would like a theory to be proposed and a test performed that either is consistent with it or falsifies it. In other words, a theory holds until one decisive test of its predictions fails. But who defines what is "decisive"?

Filtering which results in an erroneous test can either be systematic or accidental. The systematic filtering is the most dangerous because repeated experiments will produce repeated incorrect conclusions. In my own field of structural geology, for example, granite mylonites--rocks with distinctive "augen," or "eyes," of white feldspar minerals in a sea of dark, fine-grained minerals--were once thought to be the distinctive product of intense frictional grinding and thus indicative of earthquakes. Detailed microscopic and theoretical studies of these rocks later showed that they actually deformed ductily, like stretched taffy, and thus they have no relation to earthquakes. Clearly, there existed a systematic filter that corrupted tests for the seismogenic behavior of many faults.

Even if there is no systematic filtering effect, your experiment can mislead you because it is too narrowly restricted--either by your lack of imagination or some inbred, heartfelt, gut feeling that you "understand" nature. Such a pre-experiment prejudice can often lead to a selecting of calls that are likely to yield a pattern that you interpret to be consistent with your cherished hypothesis about the shape of the ships. Human nature is such that we tend to like the theories in which we have our time and reputation invested. Max Planck once wryly commented that "a new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it" (cited in Kuhn, 1970, p. 151). This kind of restricted thinking in the scientific game of Battleships results in making the same calls you've made before, because the ships were there in the past. However, if you envisage several possible ship configurations in advance and pick the best call that can be made to distinguish between them, you are then much more likely to discover unpredicted patterns that may exist in ship placement than by just testing for a single ship configuration--and thus the multiple working hypothesis model of T. C. Chamberlin (1897).

There is commonly an erroneous assumption that scientists go about gathering facts objectively until they have overwhelming evidence that “proves” a particular scientific principle. Scientific discoveries are often made beginning with this process, but little advancement can be made without application of the process of proposing hypotheses and testing them. In a classroom setting, the best way to convey how this process might work in the Battleships analogy is to draw a grid on the blackboard. The grid doesn’t need to be extensive (e.g., A through G on one axis and 1 through 7 on the other). The instructor chooses a pattern of squares that is unknown to the students. This can be something simple like a cross or an X shape. The squares don’t have to be connected, although it works best for demonstration purposes if they are. Ask students to pick various squares, one at a time. To demonstrate the approach that is commonly thought to be the “scientific method,” have the students call out a series of grid squares until enough have been called to expose all the grids making the pattern shape you have selected. A shape will emerge if enough squares are called out. This is a simple inductive method, which can produce results, but is limited to the number of possible grids that can be called. In nature, that would be like measuring everything possible to identify patterns--in other words, an endless process that might advance our understanding of natural processes or might just waste our valuable time.

On a second grid, again select a pattern unknown to students. Have them, as before, randomly select grid squares, but this time when the first hits are made the instructor could hint that the shape might be the same as before (which it is not) and have the student select a grid square consistent with the previous pattern. Here the instructor should point out to the students that they are developing a hypothesis and that it can be tested simply by choosing a single square. Since the instructor has chosen a new pattern, the square the students pick should record a miss. So at this point the instructor can discuss what some philosophy of science writers call falsification--that is, the students hypothesized a certain shape and the ground testing of that hypothesis showed it to be incorrect. Now have the students continue with a few random selections until one or more hits are made. At this point have the students develop a hypothesis. In fact, have two or more students develop different hypotheses--again, the concept of multiple working hypotheses. Have the students agree as to where the next best square selection might be made to distinguish between the various hypotheses. After the selection is made, have them discuss which hypotheses are still viable and which ones should be discarded. Continue this process with accompanying discussion until the instructor feels they have successfully communicated to the students what is essentially the process of scientific inquiry.

Scientific inquiry, though, is not limited to some imaginary 26 by 26 grid. Even with the best intentions, we can never perform enough experiments to completely describe nature's patterns. We are always extrapolating from a limited number of calls. Some extrapolations are certainly better founded than others, but even the best have some level of ambiguity. Furthermore, some experimental results are not clear; but are obfuscated by poor design, poor communication, or intent. This lack of clarity creates another kind of accidental filter that is equally detrimental to identifying patterns.

In the paper game of Battleships, one eventually gets to find out where all the ships were placed. Scientists aren't so lucky. Karl Popper (1959) commented that, in valid scientific inquiry, only testable hypotheses have meaning, and even a poorly supported hypothesis is acceptable as long as a test can be constructed which could show it to be wrong. Moreover, he felt that no theory should be considered "true" in the strictest sense. From this it follows that commonly used words like *prove* and *confirm* are inappropriate language to be applied to a hypothesis, or a test of a hypothesis.

Since a theory can never be proved, one might conclude that truth in science is by plebiscite, which is as likely a way of defining "truth" as twelve jurors are likely of determining guilt beyond a reasonable doubt. As with guilt, truth in science by consensus can be, and often is, wrong. Yet, even if the search for the ultimate truth in nature is indeed futile, scientist wouldn't dream of giving up their quest. We may never hit all of nature’s battleships, but at least we can narrow the field of their locations--and, it is hoped, leave the world a little more interesting a place than we found it.

References

Alvarez, L., Alvarez, W., Asaro, F., & Michel, H. (1980). Extraterrestrial cause for the cretaceous-tertiary extinction. *Science*, *208*, 1095-1108.

Chamberlin, T. (1897). The method of multiple working hypotheses. *Journal of Geology*, *5*, 837-848.
Kneale, W. (1967). Scientific revolution for ever. *The British Journal for the Philosophy of Science*, *19*, 27-42.
Kuhn, T. (1970). *The structure of scientific revolution*. Chicago: The University of Chicago Press.
Popper, K. (1959). *The logic of scientific discovery*. New York: Basic Books.
Urey, H. (1973). Cometary collisions and geologic periods. *Nature*, *242*, 32-33.
Wegener, A. (1915). *Die entstehung der kontinente und ozeane*. Braunschweig: Friedrich Vieweg & Sohn.

Critical Incident

An Invitation

Readers are invited to send, to the Editor at editor@ScienceEducationReview.com , a summary of a critical incident in which you have been involved. A critical incident is an event, or situation, that marks a significant turning point, or change, for a teacher. The majority of critical incidents are not dramatic or obvious, but are rendered critical through the analysis of the teacher (see Volume 3, p. 13 for further detail). You might describe the educational context and the incident (please use pseudonyms), analyse the incident (e.g., provide reasons to explain your observations), and reflect on the impact the incident made on your views about the learning and teaching process. Upon request, authors may remain anonymous.

We have undoubtedly all done things about which we were very pleased, and perhaps done other things about which we did not feel so pleased, and we all need to remain reflexive of our practice. While teachers will view an incident through the lenses of their own professional experiences, and may therefore explain it differently, this does not detract from the potential benefits to be gained from our willingness to share our experiences and thus better inform the practice of other teachers.

So Where is Your Homework?

By: Erica M. Brownstein, Capital University, Columbus, OH, USA
ebrownst@capital.edu

I was checking homework. This was a regular routine in my high school chemistry class. I believed that homework was essential to helping students not only learn the subject at hand, but also be successful in college. And then there was Brian.

> "Brian, where is your homework?"
> "I didn't do my homework."
> I mustered up my most matronly voice: "I'm concerned. I see that you haven't turned in homework for awhile now."
> "Mrs B, I've already failed your course."
> I was stunned. "I don't understand."
> "Well, if you fail four of the six grading periods, you fail a course. So, I figured it doesn't matter if I do homework. I like your class and I like chemistry. I don't like doing homework anyway."

I was in shock. How could Brian have failed my class? Sure, he didn't do homework, and yes, he performed poorly on tests. But, he was often the one in class that was helping other students do their homework. He would tutor students during class. He would lead the class in understanding laboratory applications. He would volunteer to answer questions in class.