GIA
GEMOLOGICAL INSTITUTE OF AMERICA®

How Science Works: Understanding the Scientific Method *Summary Version*

By Mary Johnson, Ilene Reinitz, and James Shigley

Understanding the scientific method has become increasingly important for members of the diamond industry. Manufacturers, dealers, and retailers continually make important business decisions, and these decisions are often based on competing claims about diamond appearance, especially as it relates to the diamond's proportions. These competing claims make the scientific method essential, because only through research conducted using the scientific method can we critically evaluate the basis of these claims (i.e., the assumptions or hypotheses on which these claims rely). By comparing different claims in this way, we are better able to determine which claims have the highest validity, and this allows for more reliable business decisions.

The scientific method (as it is now understood) was described by Karl Popper in 1934. Before then, the method of induction often was used to find truths about the world. Induction assumes facts about the world solely through the *observation* of instances, that is, empirical evidence (e.g., we notice that all of the grapes we have eaten in the past had seeds, and therefore decide "all grapes have seeds"). Unfortunately, induction has several weaknesses. For example, researchers may never know if they have seen a complete sample of the instances in question (e.g., we may never have eaten seedless grapes). Also, there is a possibility that conscious or unconscious observer biases may affect the results (e.g., we may think that all grapes should have seeds, and so all seedless "grapes" are not really grapes at all).

Nevertheless, although one can never conclusively prove the truth of a claim through empirical observations, one can *disprove* it through observations of negative instances, which is called *falsification*. Falsification leads to critical and rigorous science by testing for ways in which a currently held or proposed claim might be false or inaccurate (e.g., we purposefully examine all types of grapes that we can find to see if we can locate a type of grape without seeds).

The root of any claim or theory about phenomena in the world is an idea. The idea is stated as a *hypothesis* (i.e., a clear prediction of a phenomenon's nature or behavior in a given situation). This hypothesis must be formulated in such a way that suggests a test or experiment that can prove it false (e.g., "All grapes have seeds").

In addition, the hypothesis must have specified testing procedures that are detailed and repeatable. This allows future falsification tests to be conducted by any group of scientists; the more independent falsification tests a hypothesis withstands, the more reliable and valid it becomes. Once a hypothesis has been validated, another important step in the scientific process is to publish it, along with its testing procedures, in a peer-review journal. This encourages further scrutiny of the hypothesis and its results.

Computer modeling sometimes may be used to provide another level of predictive power. In these cases, a falsifiable hypothesis would be formulated to test whether a computer model provided accurate results (e.g., whether it gave similar results to another form of testing, such as observation tests).

Additionally, mistakes are often a crucial part of the scientific method. Rigorous testing sometimes brings to light new aspects of a hypothesis or aspects that need to be tested further. These further tests may lead to amendments to the hypothesis or even its replacement with a new, and hopefully more accurate, version.

The only absolute certainty one can have in science is that a particular hypothesis is false. However, the more falsification tests a hypothesis withstands, the more justified one is in believing it to be valid. A hypothesis that withstands this scrutiny becomes known as a *theory*. However, even accepted theories are provisional, and eventually may be proven false and replaced.

Because all truths are "provisional" in science, scientific truth often creates difficulties for the practical world of business. Businesses are forced to make policy decisions based on theories that are reliable (i.e., hypotheses that have withstood numerous and rigorous falsification tests), but are never absolutely certain. Such hypotheses must be submitted to the most rigorous falsification tests possible, to insure that policy decisions are based on valid and reliable hypotheses.

Originally published: September 27, 2002
GIA on Diamond Cut

How Science Works: Understanding the Scientific Method

By Mary Johnson, Ilene Reinitz, and James Shigley

First Thesis: We know a great deal. And we know…things which are of considerable practical significance and, what is even more important, which provide us with deep theoretical insight, and with a surprising understanding of the world.

Second Thesis: Our ignorance is sobering and boundless. Indeed, it is precisely the staggering progress of the natural sciences (to which my first thesis alludes) which constantly opens our eyes anew to our ignorance, even in the field of the natural sciences themselves…With each step forward, with each problem which we solve, we not only discover new and unsolved problems, but we also discover that where we believed that we were standing on firm and safe ground, all things are, in truth, insecure and in a state of flux.

K. R. Popper ("The Logic of the Social Sciences" in *The Positivist Dispute in German Sociology,* 1976)

In our previous articles in *GIA On Diamond Cut*, we examined a variety of information that is essential for understanding issues related to polished diamond appearance. For the most part, this information would come as little surprise to those in the scientific community who are familiar with diamonds; it forms the basis of our current knowledge of diamond appearance. But how do we move from the *unknown* to the *known* when trying to understand this or any other subject of investigation? At some point in the past, all of the facts presented in our previous articles had to have been "discovered," and then tested and found to be true and (perhaps) useful. But how do we discover new "truths" about diamond appearance? Or, alternatively: How do we replace current ideas about diamond appearance with new facts and statements that represent a *better understanding* of reality? In moving from a *lack of knowledge* about diamond appearance — or perhaps, a state of *incomplete* or *inaccurate knowledge* — to a state of greater and more accurate knowledge, we must rely on the *scientific method*.

KARL POPPER: THE SCIENTIFIC METHOD EXPLAINED

The scientific method was first described only in 1934, long after Newton (and even Einstein) had done their most significant work. In *The Logic of Scientific Discovery*, Karl Popper (the original articulator of the scientific method)[1] outlines what he thinks is the only practical method for scientists to use in examining and understanding how the world works in order to gain new knowledge. Although many scientists previous to Popper used some form of the scientific method to make their discoveries, Popper was the first to outline logically the necessary steps and requirements of this method, as well as to analyze clearly how it was distinct from other scientific approaches.

Popper's first step is to separate facts found through the process of scientific method from those found through the act of induction alone. *Induction* assumes that general facts about the world can be learned just through the observation of specific instances. Simply put, if we wanted to learn something about the world using induction (that is, *empirical* evidence, or evidence gained from our own experience), we would need only to observe as many instances of this phenomenon or event as possible and then draw our general inference from the average results of these observations. Although this is a reasonable first method for finding out about things in the world, this process of determining truth can have several weaknesses.

One weakness is that we always have to question whether we have seen a *complete sample* of the phenomenon we are studying. For example, the first synthetic rubies (e.g., Vernueil) were grown by melt processes; they all had curved growth layers and gas bubbles similar to those seen in glass. Therefore, if gemologists wanted to know whether a ruby was natural or synthetic, they examined its growth layers. Any ruby with straight growth layers that formed angles, and with fingerprint-like inclusions, could be assumed to be natural. If we were relying solely on induction, we could have created a theory that stated: All synthetic rubies have curved growth layers (and at this point stopped looking at additional synthetic rubies). Unfortunately, this criterion alone became an obsolete truth when flux and hydrothermal synthetic rubies were developed, because they do not contain curved growth layers. If we had relied only on a "truth" or theory created solely through induction — i.e., the observation only of melt-grown synthetic rubies — we might still be identifying these new synthetics as natural.

Another possible weakness in a theory created solely through induction is that we might consciously or unconsciously have a propensity to believe "that which makes sense to us." Unfortunately, what could be termed *common sense* is often limited by our experiences and desires, for example: (1) a limited exposure to an adequate sample of instances (i.e., because we haven't seen the "new truth" ourselves), (2) a general desire not to change from a belief that has been held for a long time (especially if there are practical concerns, such as the financial well-being of a company), or (3) a reticence to admit believing something later found to be incorrect. All of these motives have the power to limit our desire and ability to learn new things about the world through induction; they limit our ability to learn facts and create theories that are *closer to reality*.

So, if a multitude of observations about something in the world cannot reveal an absolute truth about it, what method remains to study the world around us? Popper suggests that although we may never conclusively prove the truth of a theory solely through empirical observation (i.e., induction), we can often *disprove* a theory based on empirical observations of *negative instances*. In other words, it only takes one synthetic ruby with straight growth layers for us to disprove the theory: All synthetic rubies have curved growth layers.

The method of using negative instances is called *falsification*, and it is the essence of the scientific method. A science based on falsification is a "critical" science because it:

1. Purposely attempts to find ways in which current theories fall apart or are no longer seen to be true (because they no longer can explain new observations), and
2. Creates new theories about the world by consciously structuring the proposed theory's building blocks (i.e., its *hypotheses*) in such a way that they can be tested to see if they are false.

Let's examine the general outline of a scientific experiment that follows the guidelines of the scientific method.

APPROACHING THE TRUTH

The root of any hypothesis or theory is *an idea*. Through general observation or original insight, an individual or group of individuals decide that there is something in the world around them, a *phenomenon*, that they would like to investigate. Popper is very clear to point out that the origin of the idea does not need to have been created in a logical or scientific manner. It is not the process of creating the idea, *but the method with which that idea is tested*, that defines robust science.

The next step in the process is crucial: The idea must be formulated into a hypothesis that can be *falsified* (i.e., shown to have instances different from those formulated) by rigorously conducted experiments and testing. In most cases, then, a hypothesis will be a clear *prediction* that a phenomenon will always be true or not true (e.g., all synthetic rubies have curved growth layers) in specified circumstances. These specified circumstances must be *fully detailed and repeatable*, so that initial testing — as well as future testing — of the hypothesis has agreed-upon *boundary conditions* (i.e., type and number of samples, instruments used, testing procedures, etc.) within which tests of falsification may be made. If a hypothesis is either (1) not falsifiable (e.g., if the hypothesis allows for negative instances of its own theory), or (2) not repeatable (e.g., if the conditions within which the hypothesis is tested cannot be reproduced by the original individuals or other experimenters), then the process by which the theory is created is not following the scientific method.

A research technique that may be inserted in these early stages of the scientific method is modeling. In our last *On Diamond Cut* article[2], we discussed the history and general methods of modeling (especially as related to diamond appearance). In general, modeling attempts to recreate phenomena or processes (e.g., either with mathematical formulas or through computer programs) that occur in the real world. As a stage of the scientific method, modeling fits somewhere between the step of getting the initial idea and the step of formulating a hypothesis (although sometimes researchers may move back and forth between the model and the hypothesis before they finally decide on the appropriate method and boundary conditions of the experiment).

An example of modeling would be the way that GIA arrived at its metric for "fire" in round brilliant cut (RBC) diamonds. If accurate, GIA's fire model — and more specifically, its end result, the fire metric — should be able to predict which round brilliant diamonds will appear to have more or less fire than other diamonds (within the boundary conditions established in the model). Also, once a model's predictions have been verified by experimentation and observation of the real system, the model itself can be used for further experimentation of *other boundary conditions*. This allows researchers to save time and money by using the model to identify significant or valuable areas of the phenomenon that they can then explore using physical experimentation.

Here is a diagram of the steps in the scientific method that we have examined thus far (figure 1). If a model is part of the scientific process with which we are attempting to find a "truth" about something in the world, then the resulting hypothesis should in some way formulate a prediction, based on the results of the modeling, that is *falsifiable*. For example, if the model predicts the relative amounts of fire in RBC diamonds, then one hypothesis could state: A group of observers, viewing the diamonds within the same boundary conditions as the model, will rank the relative amounts of fire in the diamonds *in the same order as predicted by the model*. As long as the boundary and testing conditions of the model are repeatable, multiple tests by various experimenters can be undertaken in an attempt to prove the hypothesis false.

If the hypothesis is shown to be false in tested instances, and the boundary conditions of these tests are shown to be reliable and accurate, then the hypothesis is false (as is the theory that might have been formulated from it). Interestingly, this is the only certainty that the scientific method provides: The knowledge that something is not true. Note that even when a hypothesis withstands continual tests of

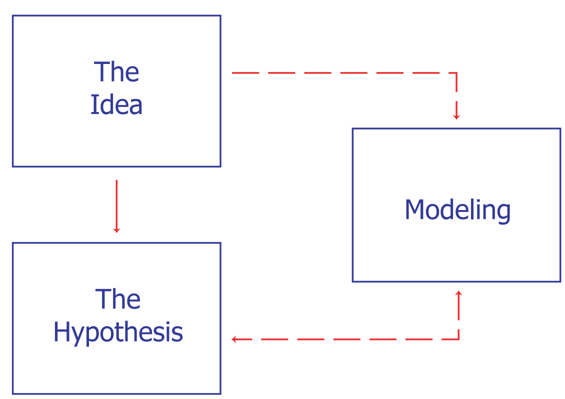

Figure 1. The scientific method begins with an idea about a certain phenomenon, and then proceeds to create a hypothesis about that phenomenon that is stated in such a way that it can be tested to be proven wrong. Researchers may initially use modeling (before developing a hypothesis) to understand the phenomenon; this model is then tested by means of a hypothesis and physical experiments.

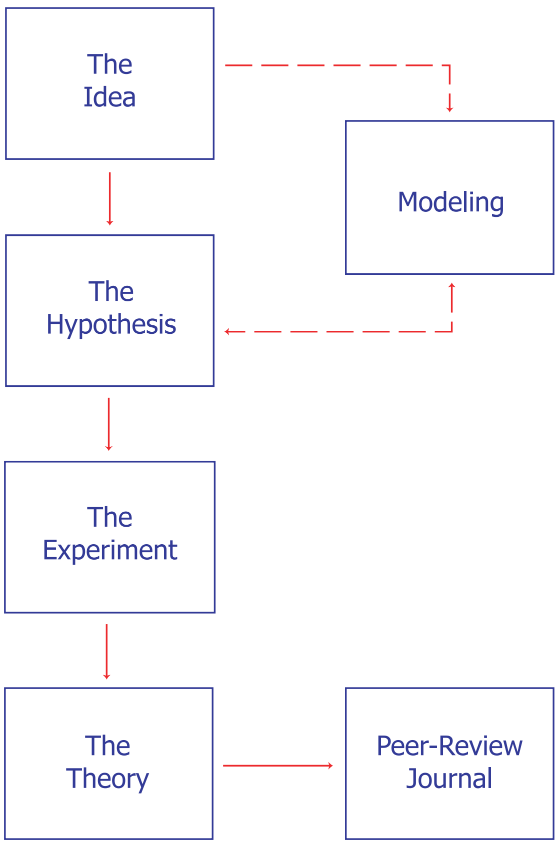

Figure 2. The completed scientific method.

falsification, which make it a valid hypothesis, it is still not an absolutely reliable statement of fact or truth. There are many ways in which empirical observations, or the inferences made from observations by induction, can be unreliable. However, the more falsification testing a hypothesis withstands, the more likely (and justified) we are in believing that the hypothesis is *valid and provisionally true*. Every theory that science holds true about the world, it only holds as being provisionally true. Although we may never know the absolute truth about some phenomena, the more a theory (and the hypotheses from which it is built) withstand the tests of time and falsification, the more justified we are in thinking that we can rely on it.

Once a hypothesis has been rigorously tested, and found to withstand falsification, it can be combined with similar hypotheses on related phenomena and be formulated into a theory. An important step is to disclose or publish the tested hypothesis, or theory, so that others may carry out their own falsification tests. (The completed scientific method is summarized in figure 2).

Again, it is essential that all testing and boundary conditions be disclosed truthfully so that tests can be performed accurately and repeatedly by others in the scientific community. An important part of this disclosure and publishing process is the role of the peer-review journal.

Publication in a peer-review journal accomplishes two important functions: (1) it provides a framework within which new hypotheses and test results can be explained and evaluated by other experts in the field ("peers") before they are published; and (2) it makes these hypotheses and test results (along with the details of their test and boundary conditions) publicly available.

Peer-review journals have been around since the mid-1700s; even before that time, the value of submitting papers to colleagues for informal or formal review was considered an integral part of the scientific process. Articles that are submitted to peer-review journals typically are reviewed anonymously by three or more respected experts and scholars in the same field of study. Reviewers generally consider several aspects when evaluating a paper:

- Is the information an original addition to a general topic?
- Does the author demonstrate a thorough understanding of previous work on this general topic?
- Are the assumptions, hypotheses, test samples, and experimental conditions adequately explained? Also, are the samples and tests used appropriate and/or sufficient to the hypothesis?
- Is there accurate use of logic and the scientific method?
- Are the discussion and conclusions of the experiment consistent with the data results?
- Is the information appropriate for the intended audience?

If all these concerns are adequately met, the article will usually be published. In most instances, reviewers (and journal editors) provide valuable comments to the authors of an article that will help them strengthen the presentation of information.

SCIENCE IN THE "REAL WORLD"

We all make mistakes; to err is not distinctively human. But although many other living things, animals and even plants, do have a partial ability to anticipate some of their mistakes, to recognize them and even to learn from them, only human beings, it seems, actively assert themselves in this direction. Rather than wait for errors to reveal themselves, perhaps with disastrous consequences, we consciously and deliberately seek them out: we put our ideas and inventions to the test, we probe critically, we scrap what we find to be wrong and try again…There is no way known of systematically avoiding error; no way known, in particular, of avoiding it in our exploration of the unknown. Thus a reluctance to make mistakes typically degenerates into a wariness of new ideas, into a distaste for any kind of bold initiative. If we are in earnest to discover what the world is like, we must be fully prepared to correct mistakes; but if we are to correct them, we must be fully prepared to commit them first.

David Miller (taken from Editor's Introduction to *Popper Selections*, 1985)

ENTERING THE PRACTICAL WORLD

The consequences of a system that only acknowledges provisional truths are easier to accept in a hypothetical situation than in a practical "real-world" situation where serious tangibles such as money (as in business) or lives (as in medicine) are at stake. How can an individual or company make rational decisions based on truths that are never completely reliable or absolute? Karl Popper phrased this problem as two distinct but related questions:

1. On which theory should we *rely* for practical action, from a rational point of view?
2. Which theory should we *prefer* for practical action, from a rational point of view?

Popper's answer to the first question is that it is not rational to rely on any theory, since no theory can be shown to be undeniably true in all cases. However, if we are deciding rationally between theories, we should *prefer* the best-tested theory that has withstood all attempts to falsify it. In a situation in which we must act (keeping in mind that inaction is also an act), the theory that has withstood multiple tests of falsification is our best choice.

One more characteristic of the scientific method should be described. Many natural phenomena can be explained by more than one hypothesis or theory. It sometimes takes scientists many months and even years to design and implement experiments to test which theory is more correct. When there are competing hypotheses or theories that attempt to explain a certain phenomenon, the simplest explanation (until proven otherwise) usually is considered to be the best choice. Even when a person has successfully relied on the scientific method to provide results, in many cases the ultimate decisions will depend as much on policies and business practices as on science. We have already discussed the difficulty of wedding the practical world to that of theoretical science: the unavoidable need to make decisions based on conclusions that are only provisionally true (and that represent incomplete knowledge). This, in fact, leads to the second difficulty: *dealing with mistakes (and uncertainty).*

Mistakes are a crucial part of the scientific method. Mistakes are part of discovery; they are part of the provisional nature of truth itself. *Ultimately, in the world of science, mistakes are unavoidable. However, we can lessen their impact, and hopefully their regularity, by increasing the number and range of falsification tests to which we originally subject our hypotheses*. Although this may increase the time it takes to release results, it allows us to feel increasingly justified in our belief that a hypothesis is valid.

CONCLUSION

Although the scientific method may not be the easiest or most comforting method of trying to approach the truth about phenomena in the world, it *is the most reliable*. Karl Popper's keen insight was that because any truths we may hold about the universe can only be provisional, our only hope is to rely on a method that earnestly attempts to falsify those truths in an effort to find the most reliable of them. This was an essential discovery in the world of philosophy and science. Although scientists had previously discovered many "truths" about the universe, they now had a general road map for finding still better, more robust theories. Fortunately for the gemological world, the scientific method can be applied as easily to gem research as it can to other scientific endeavors. ■

We hope that you found this article useful, and invite any feedback or comments that you may have. You may contact us by e-mail at DiamondCut@gia.edu.

Originally published: September 27, 2002
GIA on Diamond Cut

REFERENCES

Popper, Karl (1934) *Logik der Forschung* [*The Logic of Scientific Discovery*, translated in 1959], Routledge, London, 480 pp.

Popper, Karl (1985) *Popper Selections*, ed. By David Miller, Princeton University Press, New Jersey, 479 pp.

Keller, Alice (2000) Editorial: Not invented here. *Gems & Gemology*, Vol. 36, No. 2, p. 95.

Johnson, Mary (2000) Letter to Editor. *Rapaport Diamond Report*, Vol. 23, No. 27 (July 28), pp. 11, 103.

[1]Karl Popper was born in Vienna in 1902. He received his doctorate in philosophy from the University of Vienna in 1928, and qualified to teach mathematics and physics in secondary school in 1929. His 1934 publication of *The Logic of Scientific Discovery* heralded Popper's start as a philosopher of science.

[2]"Diamond Appearance: The Components of a Computer Model"

