Scientific Method

Robert A. Johnson, PhD

Research is a formal discipline that requires careful planning, vigorous discrimination and clarity of thought. High school and undergraduate curriculums are designed to provide information that is commonly regarded to be factual. Upon closer examination, that dogma may be questioned on a number of levels. But, the key to advancing our understanding of the world around us lies in the ability to find answers to those questions. This goal may be best accomplished by employing the long standing principals of the *scientific method*.

For hundreds of years the scientific method has evolved to serve as a framework to advance our understanding of the unknown. It serves as a means for us to organize our beliefs, identify challenges to the dogma, explore the viability of potential explanations, and to extend new findings into potential applications. While the scientific method used to be a fundamental concept for most academic programs, it is now virtually absent from most doctoral training programs. As a consequence many of today's emerging scientists -and even their mentors- are only vaguely familiar with the principals that serve as the backbone for scientific discovery.

The strengths of the scientific method are derived through the formation of tangible theories, the derivation of appropriate and testable hypothesis, and by an appreciation for the limitations that are inherent to experimental outcomes.

Premises and Conclusions - ***Developing scientific insight.***

In the philosophical arena, an *argument* is a conceptual entity that consists of an arrangement of *premises* and a logically derived *conclusion*. The ability to construct an effective argument is essential for building a strong scientific framework. For example:

> ***Premise 1*** - All dogs bark
> ***Premise 2*** - That critter is a dog
> ***Formal Argument*** - If all dogs bark, and if that critter is a dog, then that critter must bark!
> ***Conclusion*** - That critter must bark!

A premise is an assertion that is presumed to be "true". However, premises may not actually be true. In the above example, one could certainly argue that some dogs, like dead dogs, do not bark. It might also be argued that the appearance of the critter of interest may match a variety of characteristics (four legs, furry, with a tail etc) may still not be specific for a dog.

Regardless of the accuracy of a set of premises, these assertions can still be linked through collections of widely adopted logical systems. Some logical constructs are familiar to most high school curriculums and may consist of elements such as:

> If all "A" are "B", and all "B" are "C", then all "A" must be "C"
>
> If all "A" are "B" and no "B" are "C", then no "A" are "C" etc.

Premises may be arranged within a logical relationship to generate a valid argument, but it is important to realize that all arguments are inherently limited by the accuracy (or "trueness") of the premises. While one can never be absolutely certain of the accuracy of any premise, most premises rely on a formally outlined set of assumptions that are adopted from a working model of a proposed system.

Such a carefully arranged model is collectively known as a theory.

Construction of a Theory - *The foundation of any scientific argument.*

The scientific method relies on a declared set of personal beliefs and views commonly accepted by others, and to generate a novel explanation that might account for a poorly understood phenomenon. This logically arranged collection of premises may lead to a tentative explanation known as a *theory*.

The formation of a theory should be the first step towards any scientific study. The theory has two primary purposes. It should 1) identify and support the framework of knowledge that will be used to construct and interpret an investigation and 2) incorporate a novel element that is consistent (or tolerable) with the asserted dogma.

A theory is born from extensive reviews of the literature, group discussions and from personal experiences. As a start, a simple textbook or comprehensive review may help to identify the current dogma. Confidence in the dogma may be explored via a more comprehensive literature search which should incorporate original publications (not reviews) that span a variety of models and a diversity of disciplines. Publications should be read in full, and with the highest attention to details.

Often the literature search will reveal several related theories, but it also may identify observations which are conflicting. The formation of a theory requires careful personal conviction for discriminating studies that will be embraced from those that will be ignored. Any theory which overtly opposes mainstream beliefs should address the controversy and/or provide experimental data to support the decision. Right or wrong, the theory declares what is assumed to be true.

Since a theory serves as a means for defining terms and conditions surrounding an investigation, it will also outline the premises and establish the "rules" for experimental designs and for interpreting experimental outcomes.

Delineation of a Hypothesis - *Putting a theory to the test.*

Once a theory has been outlined, then it is necessary to create a series of tests that will support or refute the elements of a theory. This test, termed the *hypothesis*[1], must be based upon an argument derived from the theory. It must be empirically testable and is most convincing if it addresses an unlikely feature of the related theory. The hypotheses and the theory are intimately related. A theory must delineate the rules that will be used to interpret any outcome; it must not be internally contradictory. The formation of a strong hypothesis is dependent upon solid reasoning and a comprehensive, carefully crafted theory.

The development of an effective hypothesis is usually a dynamic process. An initial hypothesis is typically straight forward and often unsophisticated. This serves as a framework to design experiments that will test the hypothesis. Almost inevitably, closer examination of the experimental design will reveal potential limitations in testing the original hypothesis. A hypothesis may be deemed untestable because of limitations in technology, resources, or by other practical

[1] In recent years several groups (including NIH and American Heart study sections) have created the "global hypothesis", which is not a hypothesis at all. The term when used in those circles is actually meant to be a "theory" proper. For the purpose of my monographs, a hypothesis will exclusively refer to a testable question that is derived from a theory.

features. This leads to an adjustment in the hypothesis so that it may be transformed into one which can be empirically tested. This streamlining is sometimes referred to as "couching the hypothesis".

As an example, one may hypothesize that a "pure incentive" will inspire rats to learn. Quickly it becomes clear that one cannot really measure "learning" or precisely define a "pure incentive". However, one may reform the theory to speculate that chocolate is a "pure incentive", and that an animal who learns to run a maze will have a reduced transit time (do it faster). Accordingly, one may then reform the hypothesis to speculate that "Rats running a maze for chocolate milk will display a reduced transit time, with respect to those who are running for white (plain) milk". While it may seem a little discontinuous from the original belief, it is a testable hypothesis and could fit into a well integrated theory.

Are all theories scientific? No. In order to be scientific, a theory must be suitable for scrutiny via the scientific method with tangible, empirical outcomes.

As an example, this is a feature which distinguishes between "evolution" and "creationism". While both may constitute theories -or stories that potentially explain an event- the principals underlying the theory of evolution can be clearly and reliably demonstrated. For example, a culture of *staphylococcus aureus* may be gently and intermittently treated with a train of antibiotics over the course of several weeks. In the face of rapid cell turnover, resistant mutants (evolved offspring) can survive. Consistent with the theory of evolution, a new antibiotic- resistant life form, "*bacillus godzillus*" emerges. In the case of evolution, the finding constitutes the examination of a testable hypothesis which is derived from the theory, and the outcome supports the theory. In contrast, the elements underlying the theory of creationism are not testable. Nothing proves creationism to be wrong, it is simply untestable. As such, it cannot be properly classified as "scientific".

Limitations of the Tested Hypothesis: *Recognizing inherent boundaries for interpretation of the results.*

The over-interpretation of data is alarmingly frequent in the literature. Once a theory has been developed, a hypothesis has been formed, and a well designed experiment completed, it is easy to determine if the findings support or refute the theory. However, it is essential that the interpretations from the findings be clearly appreciated.

When applying for funding and in pitches for project support, it is vital to know the potential importance of a study. This ability to project a valuable application is a mark of long-term organization and insight. The ability to project a study for the advancement of mankind is admired by funding agencies, peers and the public. However, the actual progress of scientific projects is dependent upon ones ability to recognize the limitations of any study.

With respect to this point, Sir Karl Raimund Popper (1902-1994) released numerous treatises, in various disciplines that were designed to re-evaluate science. Through his works originating in the 1930's, Popper attempted to separate science from non-science and pseudo-science. His efforts revolutionized the way that experiments could be interpreted. While recognized as one of the most brilliant philosophers of the last century, his fundamental view was alarmingly frank; "We are all indefinitely ignorant".

Popper had tremendous respect for the scientific method, but he also pointed out that the greatest knowledge is obtained by realizing how things cannot be; but, this insight comes at an unavoidable cost of never knowing how things really work. While the concept may seem like "psycho-babble philosophical clap-trap", it is rather a vital point for scientific advancement.

From the earlier sections, it should be clear that the scientific method involves 1) development of a theory, 2) derivation of a hypothesis, 3) design of a controlled experiment, 4) execution of the experiment and 5) objective interpretation of the outcome.

Also previously noted, the theory is a collection of believable (or pseudo-believable) assertions that serve as a working frame work for a project. A strong theory consists of elements that are mostly accepted as "facts". It should also include novel elements that are in question, but potentially "true". Once an experiment is performed to test the hypothesis, the findings will either support the theory or not. Findings which support the theory can ***never*** be proof that a theory is correct. In contrast, an experimental finding that is inconsistent with a theory is ***definitive*** proof that something is wrong. Whilst one can be confident about the failure of the test, they are still left with the task of identifying which element(s) is (are) wrong. Was it an error in logic, formation of the premises, weak experimental design? Perhaps the theory is wrong! Experimental evidence that is found to be inconsistent with the parent theory is a signal to re-evaluate at all levels. It was Karl Popper's assertion that the greatest discoveries arise from re-evaluation of such "failures".

The nature of this complex process of progress through failure may be illustrated a though some colorful (and practical) examples.

Example 1: *The role of religion as a promoter of violence.*

As a ludicrous example of the way that scientific studies can be dangerously over interpreted, but ultimately resolved via the scientific method, lets consider an obviously flawed theoretical example. Imagine that an extremist group of atheists speculate that "Religion may promote violence". As a starting point they may note that many violent conflicts, like the Crusades (11^{th} - 13^{th} centuries) and the Reconquista (722-1492), have been driven by religious interests. They also assume that churches are the driving force behind religions so, ultimately, churches may cause violence. A little more mainstream, they also speculate that murder is an act of violence. They form a theory that includes:

Premise 1: Churches promote religion.
Premise 2: Religion promotes violence.
Premise 3: Violence promotes murder.
Formal Argument: If churches promote religion, and if religion promotes violence, and if violence promotes murder, then churches will promote murder!

Conclusion: Churches will promote murder!

Arising from this theory, they generate a hypothesis.

Hypothesis 1: An increased number of churches will be accompanied by an increase in the number of murders.

To test this hypothesis, they decide to select a number of cities. In each city, they plan to count the number of churches and the number of murders to determine if they display a positive correlation.

The genuine data in Figure 1 show that

there is a strong trend between the number of churches and the number of murders in a sample group of cities. While endless criticisms may be mounted of the precise definition of a “church” or “murder” or the findings further stratified for types of murders

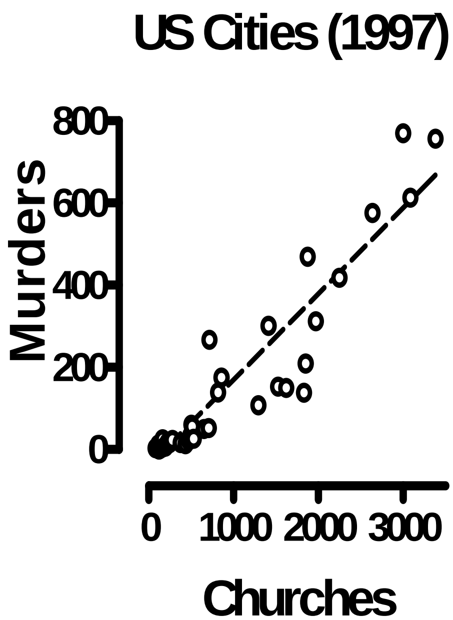

Figure 1 Plot displays number of murders as a function of number of churches in a pseudo-arbitrary selection* of 49 U.S. cities ($r^2 = 0.885$), data from 1997 state issued public records)

(accidental murder, 1st degree murder, second degree and etc.,) or for types of churches, the trends might likely hold. In addition, the corollary nature of this study cannot possibly a determine a causative element (churches promoting murder or murder promoting churches?) or that there is ANY cause and effect relationship between the two.

Even so the finding does, in fact, oddly support the theory that religion promotes violence. Before the investigators promote a national drive to burn all the churches, their enthusiasm must be tempered because -as noted earlier- a supportive finding cannot be proof of validity. But, the group potentially might also speculate:

Hypothesis 2: A maneuver to reduce the number of churches should reduce the number of murders in a town.

If that group were to decide to burn all of the churches, they would likely find that they will not reduce the number of murders in that town. Obviously, such a finding would refute the underlying theory. That would be a moment that the scientific method would yield information that is definitive. That is, something must be wrong.

Extensive critical challenges could be reviewed about the “real definition” of a church, what exactly constitutes a “murder”, the precision of the record keeping and the statistical handling of the data. Eventually, they would hopefully realize that “Premise 2" may be flawed. Maybe religion really does not promote violence. At such a point it would be prudent to revise the theory. Perhaps the number of churches and the number of murders might be related to something else; maybe they should consider the city populations?

This absurd example is meant to highlight the assertion that the validity of a theory can never be proven with absolute certainty. However, a finding that refutes a theory is absolute proof that something must be wrong. The onus of identifying the particular faulty element falls upon the investigator.

While the church/murder example exploits foolish short sightedness, it illustrates the potential flaws in over interpretation and highlights the merits of judicious application of the scientific method. The next example is a genuine example of the inevitable uncertainty of any theory.

Example 2: *The role of blood volume in diuretic-induced fall in blood pressure.*

Some additional lessons can be learned from history of diuretics. Diuretics have been used for the treatment of edema since the 16th century. In 1930, Swartz discovered that the antimicrobial sulfanilamide could be used to treat edema in patients, and it was soon realized that similar agents that promoted water loss (urine formation) could also promote a fall in blood pressure.

The underlying theory was very straight forward. It was known that diuretics could cause an increase in urine formation and that these compounds similarly caused the blood pressure to drop. The basic premises of the theory:

Premise 1: Diuretics reduce blood volume.
Premise 2: Blood pressure is dependent upon blood volume.
Premise 3: Diuretics will lower the blood pressure.
Formal argument: Since diuretics reduce the blood volume, and since blood pressure is dependent upon blood volume, diuretics will lower the blood pressure because of their ability to reduce blood volume.
Hypothesis: A diuretic will increase urine formation and lower blood pressure.

What followed was decades of studies that supported the theory, and thus led most (if not all) to accept the theory as a straight forward fact. There was dispute that diuretics promote urine formation. Ask the countless millions who have used them. In addition, the link between blood pressure and blood volume is reliably confirmed by incidents of accidental hemorrhage and by emergency room personnel who struggle with maintaining blood pressures in patients who have had extensive bleeding. The theory was so simple and so readily applicable to a number of conditions, that it quickly became accepted as a "fact".

Controlled measurements of circulating blood volume were awkward, both in clinical and in animal models. However, improved techniques and ambitious studies arising in the sixties yielded curious findings.

When a hypertensive patient is placed on a diuretic, urine formation increases and the circulating blood volume DOES decrease as the blood pressure falls. However if one takes the same patient, who is still on the diuretic for six months, they are likely to find the blood pressure is still reduced towards acceptable levels. However, at this extended period, the blood volume returns to its original (pre-diuretic) level. In short, long-term blood volume is unchanged while the blood pressure remains reduced.

This single observation is an outstanding example of Popper's assertions. There is a problem and the original theory is <u>refuted</u>. In support of Popper's views, the formation of a theory is the first step towards an inevitable failure. How does one proceed?

The first step is to determine if the drugs were correct, if diuresis really occurred, if the blood pressures and blood volumes were accurately measured and other aspects of the experimental designs, executions, measurements and summaries of outcomes appear to be correct. If those issues survive critical scrutiny, then one should consider abandoning the original theory or (in many cases) it can be modified.

As a result of this discovery, most textbooks and medical lectures have been modified to reflect that "Diuretics promote urine formation and volume loss, and that these effects promote a fall in blood pressure. However, diuretics may also exert direct effects on the vasculature that can contribute to the therapeutic target; just don't ask how that happens.

Have the studies in the past been without merit? Absolutely not. Even though the rationale may not be all accommodating, the abundance of studies arising from the theory have greatly supported the contention that diuretics may be used for hypertensive therapy.

The original theory is still widely promoted to reflect the elements that would be of primary concern in a clinical setting. A patient being treated with diuretics may complain, "I want to stop taking my medicine because it makes me use the bathroom a lot". However, a health care attendant will likely explain that the urine formation is fluid leaving the body, and that is what keeps the blood pressure under control. "Yes, the medicine makes you urinate a lot, but that is the drug working to keep you healthy." It is a simple, viable explanation that may help a patient to be compliant. As such, it is still very useful.

The failed theory serves as a new point for forming original theories and hypotheses. The next step in this particular example might be to determine (speculate, really) the way in which long term diuretic therapy promotes vasodilation.

Summary

The scientific method provides a valuable framework for developing an understanding of the unknown. The root of a scientific investigation originates within a theory, which is carefully and selectively derived from the dogma. The process continues through a series of objective and testable challenges that are logically constructed from the underlying theory. These hypotheses are then examined through a series of experiments which are designed to support or refute the parent theory. A finding which supports the theory can still be subjected to an infinite array of theoretical challenges, but an outcome which refutes a theory is an undeniable indication that some element of the theory, the asserted premises or the logical relationships, must be wrong.

In the end, it is this ebbing of such outcomes which builds a sense of reliability for a particular theory and identifies new areas of uncertainty.