How Concepts of Self-Regulation Explain Human Knowledge

by **Daniel E. Williams, Ph.D., P.E.,** *Florida Zeta '95*

Epistemology is a field within philosophy seeking to understand the nature of knowledge, with perhaps the most significant contributions made by Immanuel Kant. Kant's epistemological contribution was inspired by his attempt to understand how the deterministic physics of Isaac Newton was so effective in describing reality. A notable physical scientist in his own right, Kant is almost impossible to read, at least in translated English, but fortunately there are many insightful interpretations. In essence Kant believed the absolute reality of the universe was unknowable and that the mind imposed order on inputs from the senses.

Immanuel Kant's influence can be seen on Thomas Kuhn's modern day *Structure of Scientific Revolutions* that introduced the term *paradigm shift* into our modern vernacular while casting doubt on the absolute truth of science. Kuhn believed the scientific community organizes around a theory and develops instrumentation and experiments based on that theory to verify the theory. Kuhn's notion of the paradigm was significant because it is opposite of inductive scientific reasoning first suggested by Francis Bacon, where particular observations lead to general theory. The following work contends that truth can be found only by the dual processes of man's contemplation of and interaction with his environment. These dual processes can be thought of as a closed loop, and concepts familiar from control theory are applied to this epistemological model. When basic system-theory concepts are applied to the closed-loop scientific process, various strands of epistemology can be pulled together and historical events put in context.

A Control-Theory Primer

Control theory is a body of knowledge that describes the interaction of dynamic systems and guides the engineer to design such systems to achieve a desired result. In general, a typical control-theory problem would be to design a computer algorithm so that information from various sensors (devices that glean information from the physical world) is mathematically manipulated to produce signals that drive actuators (devices that act upon the physical world). The interesting and powerful thing about control theory is that actuators have an effect on sensors, so that the control system is self-regulating. It is able to check itself and make sure that the desired result is being achieved in the physical world.

The controls engineer concentrates on mathematical manipulation of sensor data. Drawing an anthropomorphic comparison to the human body, the sensors are the five senses, the actuators are the muscles, and the control algorithm is the brain. The remainder of this work will develop this comparison and apply system theoretic principles to human understanding of the universe.

There are two main ways a controls engineer goes about developing an algorithm. The first, useful on only the simplest of systems, is to use complicated instrumentation (an extension of the senses) to drive the actuator in a certain way and see what happens to the sensor signal. Using this actual data produced from driving the system, rules are used to determine the control strategy that will achieve the desired result. This method is normally useful only on a simple system with a Single sensor Input and a Single actuator Output (SISO), but, for this common class of simple systems, it is quite powerful.

More complex systems feature Multiple sensor Inputs and Multiple actuator Outputs (MIMO), such as the flight-control system of an airplane. It is not possible to drive these systems to acquire data and use empirical design rules to determine the optimal control. In this case a model of the physical world is constructed that relates the actuator effects to the sensor measurement. In the flight-control systems example, the model is the dynamics of the airplane, the inputs are the positions of the flight-control surfaces, and the outputs of the model are sensors that measure the resulting motion of the airplane. Using this model a closed-loop control strategy can be developed to transform the sensor measurements to actuator inputs.

The model used by control theorists is comprised of—appropriately enough—blocks. These are not children's building blocks, however; they are mathematical characterizations of how an input signal is transformed to an output signal. In general, the input and output signals can be arrays of multiple variables. For example, the input signal

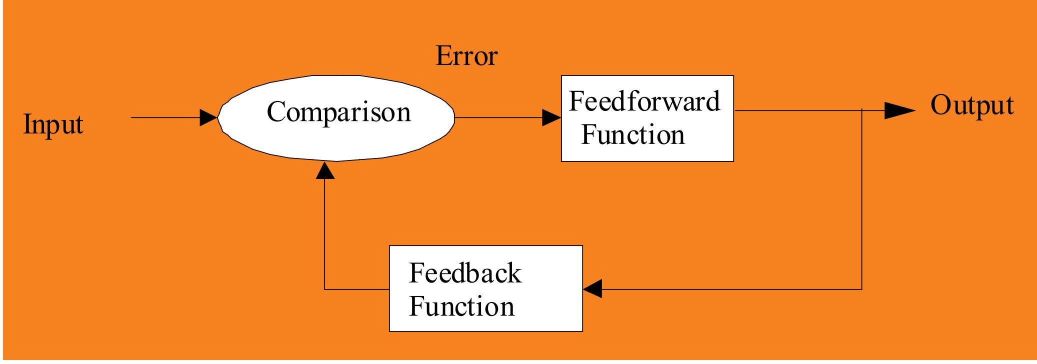

Figure 1: The closed-loop process

to a mechanical spring is a force, the output is the spring displacement, and the transfer function is the inverse of the spring stiffness, known as compliance. Instead of acting on a stiffness, a force could act on a mass to produce an acceleration inversely proportional to its inertia. The power of this type of modeling is that an electrical circuit produces a model with stiffness and inertia replaced by capacitance and inductance. So the same differential equation can represent vastly different physical realities, providing an extremely parsimonious mental model of the physical world.

A simple control block is an example of an open-loop system characterized by an input that is transformed to become an output. A very interesting behavior occurs when the output of such an open loop or *feed forward* block is measured and *fed back* to be compared with the original input, as shown in Figure 1. When this feedback occurs, the controls engineer designs the dynamic properties of the feedback block so that the error between the input signal and the feedback signal is minimized. It is important to appreciate that the output of the feedback block will track the input if the error between the input signal and feedback signal is minimized. This may seem a trivial result, but it has great practical application. One way to control the speed of your car is to give it a constant open-loop accelerator input. The problem with this method is that if you happen to be going downhill the constant accelerator input will allow the car to go faster, and going uphill it will go slower. Your car's cruise control does a better job of controlling speed because it is a closed loop. The controller compares your desired speed with the measured speed and, if you are going too slow, increases the accelerator input and, if too fast, reduces it. This property is known as disturbance rejection. Of course this example is the simplest of control systems. The main point is that all well designed control systems minimize the difference between the input signal and the feedback signal.

Sir Francis Bacon

The Importance of a Model

Sometimes it is required that the control not merely minimize the difference between the input and feedback signals, but completely eliminate this error signal—achieved by accumulating or numerically integrating it. The absolute value of the error signal might be small, but as it persists in time the integral becomes quite large. By seeking to minimize the integral of the error, precise agreement between the input and feedback signal is achieved.

Once a dynamic model of the physical world exists, there is a rich and highly developed body of mathematical techniques that will result in an optimal control algorithm for MIMO systems. The most powerful of these techniques assume that the system is linear in a dynamic sense. By this it is meant that in some way the output is proportional to the input. Consider a spring. The output, or spring displacement, is proportional to the input, or the force that compresses the spring. If you press twice as hard the linear spring is squeezed twice as much.

A nonlinear dynamic system can be seen in a block (in this case an actual toy building block) sliding on a surface with some friction. Initially a small input force produces no output motion. Only when the input force is sufficient to break the static frictional connection is there motion. Nonlinear dynamics can sometimes follow certain trends and patterns, popularized as chaos. Controlling such systems, however, must be accomplished on a case-by-case basis, and lessons learned in one nonlinear system might not apply to another.

State of the System

Happily, most systems are close to being linear in some sense. Cars and airplanes at certain speeds can be well described by a linear dynamic model, which is often sufficient if the task is to design a speed controller. Because the controller's purpose is to keep the system operating in the neighborhood where the linearity assumptions are valid, this can work for systems that are generally nonlinear. With an adequate linear model in hand, an optimal control algorithm is possible; however, it is more an art than a science to determine an adequate model. Consider the open-loop speed control of a car. One physical model is that when the accelerator pedal is pushed, the car goes faster. Another model would be that as the accelerator is pushed, the engine generates more torque, and that torque accelerates the driveline and wheels, and the car goes faster. Yet another model would be that the accelerator is pushed, more fuel enters the cylinders, more powerful explosions drive the pistons with greater force, which generates more torque on the driveline, which accelerates the wheels, and the car goes faster. This deterministic process could theoretically be extended, but there is a practical limit as the Heisenberg uncertainty principle does not allow the precise simultaneous knowledge of where a particle is and what it is doing, known as the state of the system.

For Newtonian dynamics to completely describe the behavior of a dynamic system, complete knowledge of the initial state is required, and this includes all positions and

NATURALIS
PRINCIPIA
MATHEMATICA

velocities. So the real world is not always linear, and, even if it were, it would not be possible to completely model a dynamic system. The model that a controls engineer uses to establish the optimal control is only an approximation of reality. The model of a dynamic system is a Kantian phenomena. It is not reality, and yet it is the way the controls engineer can conceive reality that can be used to manipulate it.

Using theoretical design techniques, an optimal controller can be designed that mirrors the model on which it is based. If the theoretical model of the physical system is complex, the controller will be equally complex. Associated with this complexity is the requirement that more parameters of the physical system be valued. Thus, there are counteracting tendencies at work with increasing model complexity. First, as complexity of the model is increased, more of the dynamics present in the real world are accounted for. Second, for each degree-of-freedom added to the dynamic model, perhaps three or even more parameters need to be established—each parameter the result of an assumption of linearity that is true to varying degrees.

Sir Isaac Newton

The Art of Modeling

An increase in complexity has the potential both to improve and to degrade the overall performance of the control system—thus requiring the art of modeling. Sufficiency of a model is contingent upon its intended purpose. Optimality of a model is the simplest sufficient model. Following Einstein's advice, the model should be "as simple as possible, but no simpler."

This conceptual parsimony is perhaps most famously illustrated by the Copernican heliocentric-solar-system model supplanting Ptolemy's earth-centered concept. Both models sufficiently explained observation. Copernicus' model was simpler, and therefore optimal. This simplicity had advantages in the transmission and use from person to person, and it required fewer assumptions (e.g., combinations of epicycles and eccentrics). Furthermore, the Ptolemy model is an example of a classic problem with model complexity. As model complexity adds degrees of freedom, it can be increasingly tuned to match the received data. The model becomes nothing more than an instrumental representation of past observation. There is a general belief that if several models more or less equally explain observation, the simpler model comes closer to capturing the physical essence.

One way to resolve this dilemma rationally is to take a controller that is optimally designed for a complex dynamic model and reduce its order while maintaining a resemblance to the original system in some meaningful sense. Each degree of freedom of a mechanical model usually has two states associated with it, position and velocity. The state trajectory of the system can be visualized as a multi-dimensional plot of the state variables of the system. A simple single-degree-of-freedom system is a mass that is allowed to move in one direction only, like a train on a straight track. The state of this mass can easily be plotted as position vs. velocity on a piece of two-dimensional paper. As forces act on the mass, the state trajectory becomes a two-dimensional line. The state of complex multi-dimensional modes cannot be plotted so simply. Although dynamic dimensions are not equivalent to spatial dimensions, they are often represented in the two and three spatial dimensions available to us. Our ability to visualize higher dynamic dimensions is clearly inhibited by our three-dimensional spatial experience. If the train becomes a car and is allowed to move in two dimensions on a planar surface, the state of the system requires four dimensions to express it. This system state cannot be represented in our three-dimensional world in a way that our senses can interpret, like a graph on paper, but it is no less valid a representation of a physical system. A suspended vehicle going straight along a bumpy road has 14 states —position and velocity of each of four wheel hubs and position and velocity of pitch, roll, and heave of the car body. Control theorists have been forced to conceive of the physical world in dimensions far greater than our five senses have trained us to think as we control our muscles in three-dimensional space.

So even the most complex imaginable model will be incomplete and very likely inaccurate. These complicated multi-dimensional models can be projected into a smaller dimension with robustness and some sense of optimality preserved. This projection of a system into a smaller dimension is every bit analogous to Plato's famous projection of three-dimensional space as shadows on a two-dimensional wall. Information is lost as the dimensions of the system are reduced, but the Kantian brain processing the images of the shadows does not know what it does not know.

As instrumentation has extended our senses to probe our habitable spatial dimensions, we have had to revise our understanding of the physical world. It is possible that there are dimensions completely hidden from our senses. David Bohm has shown that if such hidden dimensions are assumed, quantum effects can become deterministic. String theory requires more than three spatial dimensions. Advanced control theorists are quite comfortable projecting a complex model of reality into a reduced dimension state space and using the result to perform a useful function in reality.

Beginning with the Greeks through Descartes to Einstein and his famous thought experiments, men have believed that they could develop a mental concept of nature. Such thinking is known as rationalism and prefers reason to observation. Rationalists are inclined to discount observations that do not fit their model. Such rationalists can be dogmatic and, in the extreme, delusional. Opposing rationalists are empiricists, who believe that we know nothing apart from what we observe. Extreme empiricism, typified by David Hume, can infer nothing about the future from the past, denying causality. Tension between rationalists and empiricists has long been evident in Western culture.

A Theory of Scientific Knowledge

As the scientific method was formalized, Francis Bacon classically considered the hypothesis to be a mental model resulting from an inductive process of summarizing observations. Thus the hypothesis can be thought of as a parsimonious mental model of observation. In Bacon's traditional view of scientific knowledge, theory (hypothesis) is a sequential, rational forward processing of particulars to form generalities. Relatively recently, Karl Popper and Thomas Kuhn claim that the reverse is actually true. A tentative theory is developed, guiding subsequent observation that either reinforces its validity or refutes it. This is essentially a sequential process as well, with hypothesis (theory) leading to observation, an empirical deductive processing of generalities to particulars.

Empiricists and rationalists are both right. Theory results from observation and in turn promotes further observation and theoretical refinement. This establishes a *chicken-and-egg* loop, of the form similar to the closed loops of control theory. This particular closed loop was formed when science became the significant input to technology. Technology can be considered the manipulation of nature to suit man. Throughout most of history, technology has been divorced from science. Practical men were artists, craftsmen, and guildsmen using rules of thumb passed along by word of mouth. Prior to the rise of science, technology was developed through a combination of individual innovation and trial and error, with the trials largely motivated by extrapolation of experience, mimicking nature, religion, and mysticism. Technological innovation can produce positive results unanticipated and, in fact, unexplained by scientific theory. This somewhat embarrassing situation for scientists can drive theoretical creativity. Thomas Kuhn analyzed the simultaneity of *discovery* of the principle of the conservation of energy among 12 independent researchers. The vast majority of these were technologists working on steam engines, and yet they discovered one of the most elegant consistencies of nature.

As the predictive qualities of science improved, scientific theory became the prime technological input, largely replacing trial and error. Technology became a practical, positive test of scientific theories. The significance of this event cannot be underestimated. It formed a feedback loop.

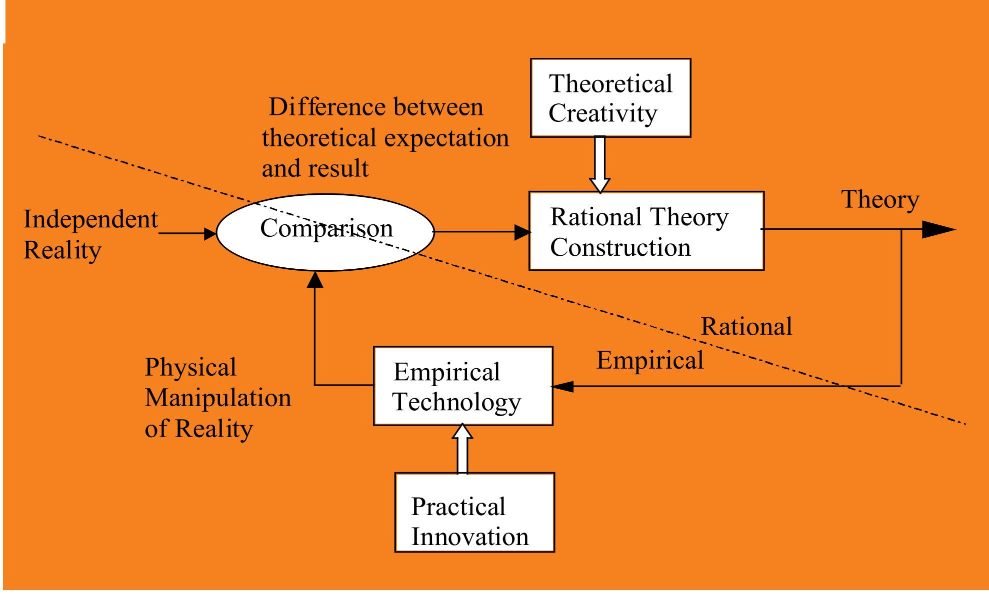

Figure 2: The scientific closed-loop

If science was a process of constructing theory from reality, technology became the inverse process of manipulating reality based on theory. When theory is used to manipulate reality, the expected result can be compared with the actual result. This allows the full attention of science to be placed on developing theories that are successful in predicting reality and thus driving technology, as shown in Figure 2.

No longer is scientific theory driven by cumulative observations of independent reality as suggested by Bacon. To the extent that these observations are successfully incorporated into scientific theory, they cease to be important.When experiments produce the theoretically expected results, there is no error signal in the closed loop. No new information is transmitted to the mental model, as the existing mental model is producing the appropriate results in practice. The theoretical process of science is driven not by observation, but by the error between observation and prediction. Thus the closed-loop scientific process agrees with much of Popper's emphasis on falsification. Only when there is an error can there be an improvement in the mental model. Popper insightfully states that the easier it is potentially to falsify a hypothesis, the stronger the hypothesis is if it stands. A hypothesis attributing a specific effect to divine intervention is very hard to prove wrong. However, a hypothesis regarding the outcome of a specific experiment is easily falsified if the results are unexpected. Using Popper's reasoning, the strongest theories are those most successful in predicting reality. The weakest theories explain, but do not predict.

Theory Tracks Reality

As Popper rightly suggests a falsified hypothesis should be celebrated because something new is learned. In addition to the positive successful application of theory, technological activity is here defined to include any contrived test of theory, such as an experiment. This is a falsifying or negative technology, and plays a key part of Popper's refutation. A hypothesis is refuted when there is a large error between the observation and the result predicted by theory that cannot be reduced by slight modifications and extensions of the theory. In this case, it is quite likely that a radical, structural change in the theory is required to reduce the error.

The flyball governor of a steam engine is an example of a closed-loop process that might be familiar to the reader.


NATURALIS
PRINCIPIA
MATHEMATICA.

Without a governor, a steam engine will reciprocate with increasing speed until failure. The action of the flyball governor serves to reduce the input to the steam engine as its speed increases. The reference signal is the desired speed, the forward process is the engine turning steam into motion, and the feedback is the governor turning motion into flyball position, which can be compared with the reference position (equivalent to the desired speed). Thus, if the engine runs slower than the desired speed, more steam is allowed, and, if the engine runs faster than desired, the input is reduced.

In such a closed loop, it is impossible to say if the engine drives the governor (which it does) or the governor drives the engine (which it does). Similarly, in our model of the scientific process, it is equally valid to say that theory drives practice (as Popper) or that practice drives theory (as Bacon). Just as the closed loop of the steam engine ensures that actual engine speed corresponds to the desired engine speed, the closed-loop scientific process ensures that our scientific theories correspond to reality.

This closed-loop characterization of the scientific process is well known to applied mathematicians and engineers familiar with control theory. A well designed closed-loop system works to minimize the error or the deviation between the reference and feedback. In this case, the reference signal is reality, and the feedback is the output of technology. When technology transforms theory into a desired result useful in reality, the error is small, and in many practical ways the feedback signal equals the input. The empirical technology feedback block can be thought of as a transformation from the mental model of reality to manipulation of the material world. The mental model determines how reality needs to be acted upon based on its sensory inputs, and this intention is transformed by the empirical technology transfer function. If there is no error between reality and feedback, the mental model is static, and the inverse of the empirical technology transfer function is the mapping of reality to the mental model. This transformation is the key to understanding Kant.


David Hume

A closed-loop system will more completely eliminate error by accumulating (integrating) it, which is equivalent to memory. Errors accumulate over time and increasingly drive the theoretical process to work to eliminate them. When accumulated error is zero, the mental model explains all past observations. It is difficult to agree with Popper that this model remains *provisional* when it cannot be possibly improved with all that is known in the history of observation. When the accumulated error is zero, the inverse of the technology transfer function is the optimal mapping of reality to our brain. The Kantian brain does not impose an arbitrary order on reality. Rather, as the loop is closed, there is a specific mapping from reality to the mental model.

The function of a closed-loop system is to force an output to track a reference. In this case, scientific theory tracks reality. Hume was correct in that Bacon's linear inductive process of theory from observation does not necessarily result in truth. When scientific theories are driven by accumulated errors between the consequences of their theories and reality, theories will resemble reality in an optimal sense. All that is required are assumptions that reality is consistent and sufficiently well behaved in the mathematical sense to allow interpolation between observations to say, contrary to Hume, that there are things we can absolutely know about reality and causality. The closed-loop process of science rescues epistemology from Hume's empirical dead end. At the same time, the closed-loop process refutes the Kantian notion that man imposes order on nature. Man does not impose order on nature; rather nature rewards man for better theoretical models of her deeper order. Thus the closed-loop theory of knowledge eliminates the excesses found in purely rational or empirical knowledge.

Changing Theories

Frequently, new knowledge is the result of a technological innovation extending or amplifying the actions of our muscles and sensitivity of our senses. Initially the model is perturbed to eliminate the error. Usually this is successful, but rarely structural changes are required, and this is referred to by Kuhn as a *paradigm shift*. Self-organizing closed-loop systems can be constructed that change their structure, in effect adapting to maximize their performance. In mathematical language, such systems are known to converge to local optimums, but in general cannot be guaranteed to converge to absolute optimums. In laymen's terms, such adaptive systems will usually converge to very good solutions—in fact the best in a mathematically defined neighborhood. When starting from different initial conditions. the adaptation could converge to a different solution, which would be the best in a different neighborhood. A paradigm change can be interpreted as a movement from one neighborhood of solutions to another. This suggests another function of the scientific hypothesis or model, agreeable to Kuhn. It directs experimentation and is able to absorb new information as it arises. A particularly good scientific theory will be incorporated into the structure of a succeeding theory as a limiting or special case, as with Newton and Einstein.

Recall that no information was obtained when the old model correlated with reality. Therefore, to the extent that the old model represents accumulated verification by its stability over time in the face of potential falsification, the new paradigm should resemble the old in its domain of effectiveness. When a paradigm shift occurs, the new paradigm must provide a similar response to the old paradigm in the input region where the old paradigm was accumulating inputs, or else accumulated verification will be lost.

Incorporating Sociology and Values

Scientific learning occurs at the personal level, but there are significant interactions between the individual and society. The role of memory implies the importance of recorded material to transfer accumulated errors between generations. Since accumulated error drives theoretical innovation, the more participants in technology, whether in the negative sense of experimentation (in an attempt to falsify) or the positive sense of technology (in an attempt to utilize), implies a value in scale, or numbers of participants, and, therefore, the value of community. Communication is important both between and within generations. Simple theories are easier to communicate and use than more complex theories. Our early concepts of theory are often learned from books, which represent the accumulated experience of the ages. Therefore, it is not surprising that the scientific process accelerated with the invention of the printing press and can be expected to accelerate again with the advent of the internet.

The proposed theory fundamentally agrees with Kuhn's conception of a sociology of science determining the paradigm around which it will organize. Kuhn believes that theories are false, and the decision to accept one theory over another is made with considerations other than absolute truth. Additional considerations in paradigm selection, such as parsimony, serve to make the paradigm more useful as a driver of practical technology. When Copernicus replaced Ptolemy, his system was not absolutely accurate, nor was it even relatively more accurate. It was roughly as accurate as Ptolemy, but far simpler, and, therefore, more parsimonious and useful. As an approximation, our mental representation of reality is most valuable when it is parsimonious. At the end of the day, the proposition must explain reality at least as well as its competing paradigms and must be more effectively transmitted within the community of scientists in order to survive in the marketplace of theories.

Conclusion

Science is not unique in its endeavor to transform observation into theory. All of the social sciences in some sense attempt this explanatory function. The difference between science and these fields is that their theory is not sufficiently predictive to drive technology, and conversely they do not as efficiently incorporate new information into their models. If other theoretical disciplines aspire to advance as quickly as science, they must develop a closed-loop linkage between theory and practice. The rationalist feedforward process of inductively generating a parsimonious mental model of reality is required to generate new hypothetical extensions and to be efficiently communicated to the scientific community. By itself, the rationalist feedforward process can support an explanative function common in social science while lacking predictive ability and therefore not be subject to falsification. The empirical feedback process of experimental technology ensures that theory will resemble reality. Both rationalism and empiricism are required for the scientific closed-loop process to work.

A desire exists to consider science as objective and quantifiable and reserve such subjective notions as values to a different realm. This is an overly simplistic view of performing science. As one of several competing paradigms is accepted, it is because of such things as elegance and simplicity, or in other words—values. At a personal level, the experience that we accumulate is largely determined by values. So we must conclude that, despite our aspirations otherwise, our encounter with the material world is shaped by our values and filtered by our senses. Only by using theoretical knowledge to anticipate a desired interaction with the physical world—and accumulating the empirical result—does man understand nature.

Bibliography

- Bacon, Francis, *Novum Organum.* Reprint. New York: Wiley Book Co., 1944.
- Bronowski, J. and Bruce Mazlish, *The Western Intellectual Tradition: From Leonardo to Hegel.* Reprint. New York: Barnes and Noble Books, 1960.
- Kant, I, *Critique of Pure Reason.* London: Penguin, 2007.
- Kuhn, Thomas S., *The Copernican Revolution,* Harvard University Press, 1957.
- Kuhn, Thomas S., *The Essential Tension: Selected Studies in Scientific Tradition andChange.* Chicago: University of Chicago Press, 1977.
- Kuhn, Thomas S., *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press, 1962.
- Popper, Karl, *The Poverty of Historicism.* Reprint. London: Routledge, 2002.

Daniel E. Williams, Ph.D., P.E., *Florida Zeta '95*, received B.S. (1985) and M.S. (1987) degrees in engineering from the University of Illinois, an M.B.A. (1990) and Ph.D. (1995) in mechanical engineering from the Florida Institute of Technology, and an M.A. (2008) in history from Purdue University.

Chief engineer of advanced engineering for TRW Commercial Steering Systems in Lafayette, IN, he has worked for the firm since 1987 in engineering and planning positions in passenger-car suspension and commercial-vehicle steering business units. He has managed the development and launch of chassis-control systems in North America and Europe.

Dan (*dan.williams@trw.com)* lives with his wife Christina and four children in Lebanon, IN.