



On the nature of thought processes and their relationship to the accumulation of knowledge, Part XIII: The nature of evidence

Cris Anderson¹

¹ Pathology, Southern Illinois University School of Medicine, Carbondale, Illinois, USA

Citation: Anderson C. On the nature of thought processes and their relationship to the accumulation of knowledge, Part XIII: The nature of evidence *Dermatol Pract Concept*. 2014;4(1):3. <http://dx.doi.org/10.5826/dpc.0401a03>

Copyright: ©2014 Anderson. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Corresponding author: Cris Anderson, M.D., Email: crisj@me.com

Foreword

I decided to write this series of essays, “On the nature of thought processes and their relationship to the accumulation of knowledge” because it seems to me logical that if we do not understand the mechanisms that occur within our brains to generate thoughts (those thoughts underlying each and everything we do) we cannot begin to understand how or why events either occur according to plan or seem to go awry. The series looks at thought processes and the “knowledge” generated by them from a variety of perspectives. The goal is to better understand just how “true” items of knowledge are, or can be, and how that understanding might help us generate and communicate “knowledge” more accurately.

This current essay about the Nature of Evidence was originally published in *Dermatopathology: Practical & Conceptual* in October 2009, but because of a series of unlucky events the essay has been withdrawn shortly after its publication without being ever replaced for the readership. After now 4 years, the essay will be finally again available for the readers of *Dermatology Practical & Conceptual*.

Further essays in the series are “in the works” and this essay is an important topic in the series overall, so I am pleased that it is to be republished.

Introduction

How do we decide what to believe? This question is most important for us because we each believe many things. Each of us assumes that what we believe is true, otherwise we

would not believe it. And we are aware that we do not all believe all of the same things. In fact, we are convinced that many other people believe things which by definition they believe to be true, that are just plain wrong, or not true, because we do not believe those same things.

Of course we each have reasons for our beliefs. If someone asks us, “Why do you believe that?” we can respond, “Because. . . .” And we all believe that our reasons support adequately our decision to believe that item. So if all of my beliefs are true, and all of your beliefs are true, and all of our reasons collectively support our beliefs, why do we disagree so often?

Let us move on to the topic of this essay, the nature of evidence. The Random House Dictionary, second edition, defines “evidence” as “that which tends to prove or disprove something; ground for belief; proof” and “something that makes plain or clear.” Serving as it does as a justification for belief, evidence is something we simply cannot live without. The definition implies that evidence must be true and unassailable, serving as a “ground for belief”, “belief” being something we assume to be true, and “proof”, which implies that there can be no argument. Also, the definition implies that evidence serves to clarify an issue, serving to improve our understanding of it. When we hear the “evidence” we are supposed to say “Aha! Now I understand” and we all move on to the next item of interest.

So the question arises, are “reasons for belief” the same as “evidence”? If we do consider them to be the same, should we consider them to be the same? In this essay we will look at the concept of evidence from a variety of perspectives and

try to gain some insight into the powerful hold the very word “evidence” holds over us. We will examine reasons why evidence cannot be expected to be true and unassailable in each and every case. We will examine the question of whether or not the word “evidence” has become more of an instrument of persuasion than out-and-out “proof” and how this affects us in the day-to-day practice of medicine, which is supposed to be Evidence-Based.

Evidence—the ideal

Preston Cloud, in *Oasis in Space: Earth History from the Beginning*, states, “A subject becomes a science when it is brought within the bounds of natural law and a body of ordering principles that make its hypotheses testable—where evidence takes priority over authority and revelation.” This statement implies that evidence is more than merely something that “justifies belief.” It implies that the justification of belief must be a sort of “proof.” That is to say, a body of evidence—that body which can be evaluated independently by one or more persons other than the person who puts forth that body—should lead a reasonable person to conclude the belief on his or her own.

No doubt a person who accepts revelation as a reason for belief feels justified in believing what has come to him by means of revelation. Encarta World English Dictionary, first edition, defines “belief” as “acceptance by the mind that something is true or real, often underpinned by an emotional or spiritual sense of certainty.” That dictionary also defines “justify” as “to serve as an acceptable reason or excuse for something.” Thus, using the term “evidence” as “that which justifies belief” does not necessarily imply “proof.” Encarta defines “proof” as “evidence or an argument that serves to establish conclusively a fact or the truth of something.” However, I would argue that most of us *infer* “proof” when we hear the label “evidence” applied to some information given to us.

In a similar vein, returning to Cloud’s statement about science, “authority”, if we accept it as a reason to believe some item of information, it serves to justify our acceptance of that item and becomes “evidence.”

In fact, “evidence” is not unique to science at all. What really makes science “science” is the “body of ordering principles that make its hypotheses testable” and the actual testing of those hypotheses. The “evidence” that we scientists use does require “proof.” We try to use “proof” formally in the sense of deductive logic or mathematical proof, in which each ensuing step is agreed upon by all as following its predecessor.

Evidence in practice

We tend to use the term “evidence” rather freely, applying the term to a reason that we accept as justifying our belief,

but without “thoroughly vetting” the term each time we use it to see if a rational person would accept necessarily our argument in the case at hand. We may do this without malice aforethought, being somewhat confused at the time as to our own argument; but occasionally a person might purposefully apply the term “evidence” in an attempt to prevent the audience from working through the argument themselves to see if the argument is actually a valid one. Unfortunately, often there is no easy way to tell whether the arguer is being logical, is confused, or is trying to hoodwink the audience.

Many problems exist in this complex world in which we live for which legitimate disagreements occur. For some of the reasons to be discussed later in this essay, a formal, undeniable proof cannot be put forth. This does not mean that there are no rules available with which such a problem can be discussed in a meaningful and enlightening way.

In fact, informal logic, or argumentation, is a discipline which recognizes that “proof” more often than not represents a consensus arrived at by the process of justifying and rebutting claims, refining through the process an understanding of the topic of the argument. David Zarefsky, in *Argumentation: The Study of Effective Reasoning*, 2nd edition, an audio course produced by The Teaching Company, explains the five key assumptions that underlie the process of argumentation. First, argumentation takes place with an audience in mind and that audience judges ultimately the success or failure of the argument. Second, argumentation takes place under the condition of uncertainty. Third, argumentation involves justification of claims made by the arguers. Fourth, argumentation is a cooperative exercise; although it may seem adversarial, the arguers share a common goal of arriving at the best decision under the prevailing circumstances. Fifth, argumentation involves risks—an arguer may be shown to be wrong and lose the argument and the loser may lose face if s/he is perceived to have performed badly during the argument.

“Proof”, or justification, rests on being able to put forth and/or to recognize a “valid” argument. That is to say the argument follows logically and is thus consistent and without contradiction. A problem arises when we humans cannot, or do not make the attempt to, discern the difference between proof and persuasion. Persuasion, in this sense, means that the words sound nice, but the argument is not valid in some way. Recall from the earlier essay in this series *Reasoning* the words of Rick Garlikov, in which he compares reasoning with a game of chess. Garlikov writes, “If little children are playing with chess pieces and a chess board, but are making arbitrary moves in what they think is emulation of adults they have seen playing chess, it is not that they are playing chess badly. It is that they are not playing chess at all, regardless of what they think they are doing or what they call it.” If the rules are not followed, there is no argument in the true sense of the word. There is merely the melodious sound of words.

If I concede a point to you when you put forth an invalid argument, you have persuaded me, yes; but the proof or justification is not there.

Interlude about learning

It is most important for us to understand how knowledge is gained. It is for the most part a painful process based entirely in trial-and-error. More involved discussions about concepts related to learning are to be found in earlier essays in this series, especially Interpretation, Causation, Reasoning, and Patterns. I will mention briefly a few salient points.

What one is able to learn and conclude is dependent on how the problem is defined—that is to say, what boundaries are drawn that define the system being studied. Different boundaries lead to different questions being asked, different types of data collected, and different interpretations of data.

All learning of material previously discovered is based in ostensive definition, whereby someone is shown something by someone more knowledgeable about the object and the learner must try to figure out exactly what it is s/he is supposed to be looking at (more complete discussion of this in Interpretation).

Abstract concepts are thought about and discussed in terms of metaphor. Metaphors are based in everyday experience and are related to features of human embodiment, for example “happiness” is discussed in terms of “up”, the “future” in terms of “forward”, “life” is a “journey”, and the like (also discussed in Interpretation). A metaphor chosen will necessarily highlight one aspect of the abstract concept, while simultaneously downplaying another aspect of that concept. This gets back to the problem of how boundaries are drawn around the abstract concept, or from what perspective the concept is examined.

Patterns are sought, which are converted mentally to rules that can then be applied to future situations. If an expected outcome ensues, we assume the rule must be correct. If we are surprised, if we see a contradiction unexpectedly or the like, we revise the rule and see if we can make a better prediction for the next event (trial-and-error at work). The rules we “discover” serve as boundaries for the system we are examining.

We start by imagining some event and then trying, “by hook or by crook”, to get to where we imagine. It is very difficult to get somewhere without some inkling of where one might be going. Serendipitous events occur, but only for the “prepared” mind.

True to pattern, when we find something, a method or an approach perhaps, we try it again and again until we finally are forced to admit that some sorts of problems require a different approach (the rules applied routinely result in surprises or contradictions). This is discussed in more depth in the essay on Reasoning.

Paradox and contradiction, if recognized, lead to the opportunity to redraw boundaries around a problem and try to “get it right.” An implicit assumption that we all hold is that, if we identify the pattern correctly and describe it with a precise and correct rule, we will avoid surprise and contradiction. We are slaves to expectation!

Knowledge and understanding

Basically, while we think we “know” and “understand” things, we only feel that way. In the earlier essay in this series on Truth it was noted that when we think we know something, our “truth bell” rings in our brains. The “truth bell” is in the ventral striatum and associated with the nucleus accumbens and is associated with pleasure and motivation. In the earlier essay in this series on Emotion, we learned that Damasio posits that even the most **objective** thought we think we are having is run through various centers in the brain associated with feelings, including the somatosensory cortex. We evolved this way because being able to assess easily our feelings better enabled us to determine whether we were in danger imminently. When the “truth bell” rings for each of us, there is no assurance that the thought ringing the “truth bell” is actually true in the sense that it conforms with Universal Law.

“Evidence” is a similar concept, since it is that which justifies belief. “Belief” is a personal feeling for each of us that what we believe is true. Something that we accept as justifying for us a belief also rings the “truth bell” and, thus, is not necessarily true in the conforms-with-Universal-Law sense.

“Evidence” as an attribute applied during argument

We all understand the world from a slightly different perspective. Each of us thinks we are correct about many of the things we understand. It is human nature to share knowledge and understanding. Often, however, the person with whom we are sharing our unique perspective requires convincing. To convince someone that our belief is justified, we offer them “evidence.” The evidence we offer them is the evidence that justifies our own belief. We assume that if the evidence is good enough for us, it should be good enough for anyone else and the evidence we are sharing should serve to justify belief in the mind of the other person.

Although at first blush we think of the word “evidence” as being associated with objectivity and science, the label “evidence” is laden intensely with feeling. If you reject my evidence, if you say my evidence is insufficient to justify your belief in what I believe, I infer that you are implying that my own belief is not justified and that I should not accept the

item I have labeled as “evidence” as sufficient to justify my own belief.

The grip of “authority”

From our earliest experiences we rely on other people to tell us “the truth.” Our teachers, our parents initially, instruct us by first ostensive and then verbal definition. There is much to learn and learning from the experience of others saves us time and is often safer than direct experience. “Don’t touch that hot stove!”, “Don’t run into the street without looking both ways!”, we are told. If someone tells us this or that plant is safe to eat, we do not risk poisoning ourselves. To look to authority is natural to humans. Following authority helps us to become more easily acculturated into our very complex society.

As our lives become more complex and as we are bombarded from all sides by information coming with increasing frequency, our tendency naturally is to believe what we hear, since we humans look automatically to authority. It is only when we hear statements that we recognize as contradicting one another that we are prompted to examine situations more carefully and to make more reasoned decisions about what to accept as true. Contradiction requires us to ask, “Who really is an authority in this subject?” “Who should we believe?” “How can we find out the truth?” In the end, we look for and evaluate information put forth as “evidence” and we evaluate carefully those items before we accept them internally as evidence, which justifies our belief. For most of us this means adhering to the “scientific method”, which is defined by Encarta dictionary as “the system of advancing knowledge by formulating a question, collecting data about it through observation and experiment, and testing a hypothetical answer.” The problem here is that we must call up into Working Memory, which can only hold simultaneously about seven items, propositions that we can recognize as contradictory. We all have so many potential thoughts and memories stored in our brains, we often do not call up information that is contradictory and we continue to believe contradictory propositions; that is we all live with cognitive dissonance.

Finding the “best fit” definition of a system

In the essay in this series on Reasoning, we noted that Lawrence Slobodkin, in *Simplicity and Complexity in Games of the Intellect*, opined that complex problems require a look by both philosophers and mathematicians in order to gain a better understanding of those problems. “Evidence” has been considered by both and we will see what each of them has to say.

What sorts of things can be accepted as evidence—philosophers

As with most abstract concepts, one’s conclusion depends ultimately on the boundaries one constructs around one’s system of thought concerning the concept.

Thomas Kelly, in *Evidence* published in the online Stanford Encyclopedia of Philosophy, notes that most people think of evidence as “something that might be placed in a plastic bag”, such as a bloody knife, or an historical document, or the result obtained by biochemical analysis of a blood specimen. However, philosophers through the ages have applied variably the term evidence to mental items present in one’s consciousness (Bertrand Russell), stimulation of one’s sensory receptors (Willard Quine), or even the totality of propositions that one knows (Timothy Williamson). Bayesians consider evidence to be those beliefs for which one is psychologically certain.

Should an item in my consciousness, say an idea I have for writing a story of science fiction (and, therefore, for which very little physical evidence is available) be considered as evidence? Perhaps it could serve as evidence that I have an imagination. Should the fact that I say I feel cold serve as evidence to support my action to turn up the thermostat? Or should I wait until someone sees me shivering so that more than one person agrees that I feel cold? Should the totality of propositions I “know” serve as evidence, even if all those propositions are not satisfiable with one another (some propositions contradict one another, if only I could bring them into my consciousness at the same time, I would see that they are contradictory and could act to resolve my cognitive dissonance)?

Kelly advises us that philosophers have offered “quite divergent theories of what sorts of things are eligible to serve as evidence.” He points out that, depending on one’s frame of reference, “the concept of evidence has often been called upon to fill a number of distinct roles”, and that, while sometimes the roles complement each other, sometimes tension is created.

One interesting concept is that, in the minds of those who are skeptical about our knowledge of the external world, one’s evidence does not favor ordinary commonsense interpretations of one’s environment over unusual alternatives. For example, skeptics may point out that we do not sense what we think we are sensing; those skeptics posit that we could be hallucinating in an undetectable way. Our evidence does not mean we see that tree, even if we pluck a leaf and feel the features of that leaf under our fingers. We could be hallucinating the entire episode. This has been dubbed the “brains in a vat” paradox. Skeptics maintain that there is no way to be certain about anything we think we “know.”

Another interesting concept is that an item of evidence, E, can never be the ultimate evidence in support of a belief

because there is always the possibility that additional evidence, E', may become known that defeats the prior evidence. Evidence that is susceptible to being undermined is called "defeasible" evidence. Kelly points out that there is controversy among philosophers whether "indefeasible" evidence could even exist.

Important to this discussion about the availability of additional evidence is whether one is considering a subset of evidence or a theoretical "total body of evidence." I remember attending courses in dermatopathology given by A. Bernard Ackerman, M.D. Often, when Bernie was discussing his diagnosis of a particular case, a member of the audience would ask (this is a paraphrased quotation), "what if you just had this area—what would your diagnosis be then?" Bernie would point out, "but you don't just have a small area, you have the entire case. It makes no sense to consider a small part of the specimen in isolation; you must consider all the information available."

It is true that our brain power has limitations, but Bernie insisted that we should not draw our boundaries around a problem more narrowly than necessary. We will return to the problem of narrow boundaries later.

Additionally, evidence supports a belief in relation to the number of hypotheses available. If there is only one hypothesis, then evidence supports only that hypothesis. But if there are multiple plausible competing hypotheses, evidence may support more than one of those hypotheses and not serve to differentiate between competing hypotheses.

Bayesians theorize that evidence can support an hypothesis only if one considers the probability that an hypothesis is correct. We physicians are all familiar with Bayes theorem as it relates to diagnosis. If the prevalence of disease A is 30% and the prevalence of disease B is 3%, the presence of a symptom common to both diseases is more likely to be evidence of disease A than disease B. What we need is a symptom or sign or test result that will serve to differentiate between the two diseases.

Another consideration philosophically is whether additional confirmatory evidence provides a stronger argument than disconfirmatory evidence. Karl Popper developed the example using the premise "All swans are white." He averred that evidence of each additional white swan provided little additional proof of the claim. He pointed out that a single black swan would disprove the hypothesis. He pointed out that the goal of science should be to "look for black swans." Scientific experiments should attempt to disprove an hypothesis. If the experiment failed to disprove the hypothesis, that hypothesis could still serve as a working hypothesis and as a (temporary) basis for further progress of knowledge about some aspect of our universe.

Victor DiFate, in his entry Evidence on the Internet Encyclopedia of Philosophy, describes Carl Hempel's work on

"The Raven's Paradox." Hempel begins with the hypothesis "All ravens are black" and then draws a logically equivalent statement, "All non-black things are non-ravens." Hempel then points out that one could provide evidence that all ravens are black by merely looking around the room and saying, for example, "That book is green. It is a non-black thing and a non-raven. Therefore it is evidence in support of my hypothesis." He further points out that, using "evidence" to support a statement that is equivalent from a logical standpoint to the original premise does not do much to prove the hypothesis in a meaningful way and that a similar argument could be used to "prove" the hypothesis that "all ravens are white." A way around this, points out DiFate, might be to "test severely" a hypothesis. It may be true that a green book provides evidence that all ravens are black, but the evidence is weak in the extreme. One should make a good faith effort to ensure that, if a non-black raven exists, one is likely to find it, just as Karl Popper has recommended.

Philosophers have spent much time and brain power on the problem of evidence. The main point I wish to make here is that how one frames the problem determines the sorts of conclusions one may draw ultimately. One cannot draw a system with infinite boundaries and, therefore, paradoxes and disagreements will arise that might suggest a better, or at least an alternate, way to draw boundaries in an effort to minimize the presence and effects of paradox.

Problem definition from a mathematical perspective

Lawrence Slobodkin, in *Simplicity & Complexity in Games of the Intellect*, explains the role mathematicians play in our world. Says Slobodkin, "The work of mathematicians is to imagine new worlds and what their regularities might be. In their choice of difficult intellectual problems they can exclude questions that require physical apparatus for collecting information and knowledge or social history in their answers. In this sense they are free of the limitations placed on empirical scientists . . . The purpose of their activity is clarity and understanding only. While they are free to make new worlds, they carefully avoid dealing with the inconvenient parts of this one."

Mathematicians draw up a set of rules, which are often represented symbolically by formulae, they then select a set of initial conditions (by defining some of the variables in the equations), and then set the whole process in motion by attempting to solve the equations to see what occurs. This is very like the concept of games discussed in the essay in this series on Reasoning. Slobodkin states of games, "Basically, all good games involve a complete, circumscribed, and simplified model of a world, consisting of a clear playing

field with understandable and unchanging rules whose rich consequences are then developed by the players.”

When mathematicians approach a problem, they set certain assumptions (which are some of the rules) and make certain predictions about what they expect will happen. If their predictions come to fruition, they accept that their assumptions are valid. If something unexpected occurs, they must consider whether their assumptions were correct or whether there could be some other reason for the unexpected occurrence. If a contradiction arises when the rules are executed flawlessly, there is a paradox.

Paradox as a clue to deficient knowledge

Thomas Bolander, in *Self-Reference* in the online Stanford Encyclopedia of Philosophy, writes of paradox “A paradox is a seemingly sound piece of reasoning based on apparently true assumptions that leads to a contradiction . . . The significance of a paradox is its indication of a flaw or deficiency in our understanding of the central concepts involved in it.” Bolander describes semantic and set-theoretic paradoxes and explains, “In [the] case of the semantic paradoxes, it seems that it is our understanding of fundamental semantic concepts such as truth (in the liar paradox and Grelling’s paradox) and definability (in Berry’s paradox) that are deficient. In the case of set-theoretic paradoxes, it is our understanding of the concept of the set. If we fully understood those concepts, we should be able to deal with them without being led to contradictions.”

For the purposes of this essay I am using the definitions, from Encarta, of “truth” as “correspondence to fact or reality” and “definability” as ability to “give the precise meaning of a word or expression or to state or describe something clearly.” As for “reality”, I invoke the words of Steven Pinker (as discussed in the essay in this series on Patterns), in *The Stuff of Thought: Language as a Window Into Human Nature*, “But reality can’t be riddled with paradoxes and inconsistencies; reality just is.”

You can see that, if we humans are deficient in our understanding of the concepts of “truth” and “definability,” it is only natural that we will have difficulty understanding the concept of “evidence.”

Parenthetically, the paradoxes mentioned by Bolander are as follows. But first, a few more definitions. Logic, as defined in *Encyclopedia Britannica*, 15th edition, is the study of propositions and of their use in argumentation. Deborah Bennett, in *Logic Made Easy: How to Know When Language Deceives You*, states, “A proposition is any statement that has the property of truth or falsity.” A proposition is composed of a subject and a predicate. As per The Random House Dictionary of the English Language, second edition, the subject

is the syntactic unit that is performing the action or being in the state expressed by the predicate and the predicate is the syntactic unit that expresses the action or state attributed to the subject. From the standpoint of logic, the predicate is that which is affirmed or denied by the subject .

The liar sentence is the proposition (and thus having the property of truth or falsity) “This sentence is not true.” “Sentence” is the subject and “true” is the predicate, which is denied by the subject in this example. When one tries to determine the truth of this proposition, one either starts out assuming the sentence is true or that the sentence is false. If one assumes it is true, then it cannot be true because the sentence itself (self-reference) states that it is not true. If one assumes the sentence is false, then it must be true since the “false” of “not true” is “true.”

Grelling’s paradox involves the use of a predicate defined in a self-referent way. States Bolander, “Say a predicate is heterological if it is not true of itself, that is if it does not itself have the property it expresses. [“hetero-” being a prefix meaning “different or other” and “logical” meaning “based on facts”] Thus the predicate “German” is heterological, since it is not itself a German word, but the predicate “deutsch” is not heterological. The question that leads to the paradox now: ‘is “heterological” heterological?’” Bolander continues his explanation, “It is easy to see that we obtain a contradiction independently of whether we answer ‘yes’ or ‘no’ to this question (the argument runs more or less like in the liar’s paradox). Grelling’s paradox is self-referential, since the definition of the predicate heterological refers to all predicates, including heterological itself. Definitions such as these, which depend on a set of entities, at least one of which is the entity being defined, are called impredicative.

Berry’s paradox is also based on an impredicative description. Bolander explains, “Some phrases of the English language are descriptions of natural numbers, for example, “the sum of five and seven” is a description of the number 12. Berry’s paradox arises when trying to determine the denotation of the following description: ‘the least number that cannot be referred to by a description containing less than 100 symbols’. The contradiction is that this description containing 93 symbols denotes a number which, by definition, cannot be denoted by any description containing less than 100 symbols. The description is of course impredicative, since it implicitly refers to all descriptions, including itself.”

Russell’s paradox is a set-theoretic paradox. States Bolander, “Russell’s paradox arises from considering the Russell set R of all sets that are not members of themselves . . . The contradiction is derived by asking whether R is a member of itself . . . If R is an element of itself, then by the definition of the Russell set, it is not a member of itself. If R is not an element of itself, then by definition, R is an element of itself.

Language can interfere with logic and evidence

This section and the next few sections will address problems with language as we use language to attempt to provide logical reasons to support our claims and to use the term “evidence” in its ideal sense.

The vast majority of us are well-intentioned. When we speak to each other and argue a point, we try very hard to follow the rules as described by Garlikov and Zarefsky and others. It is just that certain limitations in the system get in the way of our goals as we imagine those goals. Some of these limitations were addressed briefly in the section in this essay on Interlude About Learning. Another limitation is the way we are able to use language.

Language serves as an instrument of thought and thought serves as a means to understanding. Thus language, which is composed of words and syntax, is the currency of understanding. Encarta defines understanding as the ability to explain to oneself a concept and seems to imply some degree of precision since, in theory, we would explain to ourselves a concept in the same way each time we explain it, otherwise, we would not truly “understand” the concept if we kept changing our minds.

When we explain to ourselves, we may well understand the “message” in exactly the same way each time we ponder it, but what happens when we try to explain the concept or message to another person? We assume that another person understands language the same way we do, but study after study has shown that people do not understand seemingly simple bits of language in the same way. We all know from practical experience that we are often surprised by what ensues when we have relayed a message to someone.

Thought and logic

In earlier essays in this series, especially in Reasoning, we learned that the rules of logic are not the rules of our “natural” thinking processes. We looked at the work of Gerd Gigerenzer, who in Adaptive Thinking and Gut Feelings, explains the concept of “fast and frugal” heuristics and the role of Environment of Evolutionary Adaptation in the evolution of what are truly our “natural” thinking processes.

Our “natural” thinking processes often lead to errors, although usually not life-threatening errors, which is why we have survived to pursue our studies. Perhaps because we find our errors discombobulating, if not downright embarrassing, we maintain steadfastly that we can “do better.” To that end, certain patterns of thought and language have been observed throughout the ages and studies have been made to understand rules that lead to consistent and noncontradictory thought. We will examine both the rules and common misunderstandings of the rules later in this essay.

Rules of logic as rules of communication

Rules exist so that we can know what to expect. Humans, despite occasional protestations to the contrary, hate surprises. One very important set of rules that we have relate to communication. Without communication we could not exist. And without rules, there can be no effective communication.

Bennett, in *Logic Made Easy*, states, “There are certain principles of ordinary conversation that we expect ourselves and others to follow. These principles underlie all reasoning that occurs in the normal course of the day and we expect that if a person is honest and reasonable, these principles will be followed. The guiding principle of rational behavior is consistency. If you are consistently consistent, I trust that you are not trying to pull the wool over my eyes or slip one by me . . . [these principles of conversation are] consistency and noncontradiction [which were] recognized very early on to be at the core of mathematical proof.”

So in a way we expect the same rigor of ordinary conversation that we expect from mathematical proof. In fact, as we shall see, although logic is a function of language (discussed in the essay in this series on Reasoning) the rules of logic, and even messages themselves, can be written symbolically and the formulae and equations can even be manipulated by computer algorithms. Parenthetically, this should not surprise us much since mathematics in its purest form is the language of relationships (patterns) and not merely a way to manipulate numbers.

Common symbols used are *S* for subject and *P* for predicate. For conditional propositions (if/then statements), *p* is the antecedent and *q* is the consequent. Aristotle noted that form determines the validity of an argument, regardless of the truth or falsity of the propositions. This allows the simplicity of using symbols, which in some systems are manipulated according to mathematical rules, to determine validity since one will not be distracted by the subject matter. In fact, mathematical proofs are evaluated for validity by computer programs developed for that purpose since to determine validity “by hand” is tedious, time-consuming, and prone to human error. Truth or falsity of propositions must be determined separately since only a true and valid argument is a **sound** argument (as discussed in the essay in this series on Reasoning).

Logic as a means of evaluating evidence

Many things in life are not straightforward. It may be obvious that “B” follows from “A”, but it may be less obvious that “G” or “M” or “Z” is a consequence ultimately of “A.” In order to get from “A” to a conclusion, “G,” we require some sort of

“proof.” Encarta defines “proof” as “evidence or an argument that serves to establish a fact or truth of something.”

Keith Devlin, in *Mathematics: The Science of Patterns*, states, “In propositional logic, a proof, or valid deduction, consists of a series of propositions such that each proposition in the series is either deduced from previous ones by means of modus ponens, or else is one of the assumptions that underlie the proof.” Modus ponens, as discussed in the essay in this series on Reasoning, follows the form, “if p , then q .” Bennett explains, “The basic steps in any deductive proof, either mathematical or metaphysical, are the same. We begin with true (or agreed upon) statements, called premises, and concede at each step that the next statement or construction follows legitimately from the previous statements. When we arrive at the final statement, called our conclusion, we know it must be necessarily true due to our logical chain of reasoning.”

I want to point out early in this discussion that ultimately, despite all these rules we are going to examine, our decision as to valid or not valid, true or not true, is based in trial-and-error and common sense, which result from our experiences. We use rules, or at least attempt to use them, because we think rules will lead us necessarily to a correct conclusion, but if at the end of our intellectual labors we can think of a counterexample that shows our conclusion to be false in some instance, we know our thread of reasoning contains an error or that we have run into a paradox that exposes our lack of understanding.

Bennett discusses the work of Douglas Hofstadter, who, in the words of Bennett, “said that the study of logic began as an attempt to mechanize the thought processes of reasoning. Hofstadter pointed out that even the ancient Greeks knew ‘that reasoning is a patterned process, and is at least partially governed by storable laws.’ Indeed, the Greeks believed that deductive thought had patterns and quite possibly laws that could be articulated . . . [troubled by the Sophists, who used ‘deliberate confusion and verbal tricks in the course of a debate to win an argument’] Aristotle . . . attempted to systematically lay out rules that all might agree dealt exclusively with correct usage of certain statements called propositions.” Bennett describes Aristotle’s basic work in logic. Aristotle strived to develop methods “to reason from generally accepted opinions about any problem set before us and shall ourselves, when sustaining an argument, avoid saying anything contradictory.” Aristotle’s two basic axioms were “the law of the excluded middle” and the “law of noncontradiction.”

Law of the excluded middle

Explains Bennett, “The law of the excluded middle requires that a thing must either possess a given attribute or must not possess it. A thing must be one way or another; there is no middle. In other words, the middle ground is excluded. A

shape is either a circle or not a circle . . . A statement is either true or not true.” Bennett points out that this law cannot be applied reasonably to all situations, since “fuzzy logic” applies to circumstances when application of the law of the excluded middle is inappropriate. A common tactic in debate is to pretend that the law of the excluded middle applies, when it does not apply. An opponent is encouraged, thereby, to accept a position he does not hold. Examples offered by Bennett, “Either you are with me or you are against me. Either you favor assisted suicide or you favor people suffering a lingering death.” Recall the discussion on fuzzy logic put forth by Bart Kosko and described in the essay in this series on Reasoning. How many apples are entirely red or entirely green? How many people like their jobs 100% of the time? Inappropriate use of the law of the excluded middle is called the “black-and-white fallacy.”

Law of noncontradiction

Bennett describes the law of noncontradiction, “. . . a thing cannot both be and not be at the same time. A figure cannot be both a square and not a square. Two lines in a plane cannot both intersect and not intersect.”

Using syllogisms

Logicians understand that the basis of argumentation is the syllogism. Bennett states that Aristotle defined the syllogism as “. . . discourse in which, certain things being stated, something other than what is stated [a conclusion] follows of necessity from their being so.” Adds Bennett, “In other words, a syllogism accepts only those conclusions that are inescapable from the stated premises.”

Syllogisms, as series of premises related to each other in some way, are composed of words, and certain of those words appear consistently in syllogisms, defining as they do the relationship of the subject (S) to predicate (P) of the premises offered. These simple words are used often by us and we each assume that “the other guy” understands them in exactly the same manner that we understand them. But we do not understand words such as “all”, “a”, “any”, “some”, “not”, “if”, and “then” in the same way.

States Bennett, “In reasoning and language comprehension, there are several factors to consider. Sentences take on meaning based on the denotative (dictionary) meaning, the linguistic structure (syntax and semantics), and the connotation. Connotation includes the factual and experiential knowledge that we bring to the material at hand . . . If p , then q can be expressed as: p never without q ; p only if q ; q if p ; p is a sufficient condition for q ; p implies q ; q is a necessary condition for p ; q is implied by p ; or q whenever p . Though they are identical statements in logic, there is no reason to believe that individuals interpret these sentence forms in the same way.”

Quantifiers may be universal or particular. For example, “all,” “every,” and “none” are universal and specific, applying to the entire class under discussion, but the quantifiers “some,” “few,” and “many” are particular and vague, applying to only some members of the class under discussion.

In classifying valid conclusions, the symbols **A** and **I** are assigned to propositions that are affirmations (derived from the word **A**ffirmo), **A** for universal and **I** for particular, while **E** and **O** are assigned to propositions that are negations (derived from the word **nEgO**), **E** for universal and **O** for particular.

Aristotle discovered that for three line syllogisms there were sixty-four possible ways that the four types of quantifiers (**A**, **I**, **E**, and **O**) could be combined, but only four of those combinations resulted in valid conclusions. (It seems that Aristotle had discovered the problems that are associated with the “combinatorial explosion” as described by Gigerenzer and discussed more completely in the earlier essay in this series on Reasoning.) Each combination is referred to as a “mood.” The valid moods are **AAA**, **EAE**, **AII**, and **EIO**. Examples of valid syllogisms are **AAA**—All birds are animals; All canaries are birds; Therefore, all canaries are animals; **EAE**—No beans are animals; All chickpeas are beans; Therefore, no chickpeas are animals; **AII**—All biographers are authors; Some curators are biographers; Therefore, some curators are authors; **EIO**—No bases are acids; Some chemicals are bases; Therefore, some chemicals are not acids.

Syllogisms were also classified as to “figure,” which referred to the arrangement of terms in the syllogism. The conclusion is symbolized as, for example, “All **S** (subject) are **P** (predicate).” Between the two premises of the syllogism, one contains the subject and the other the predicate of the conclusion. Both premises contain a common term, called **M**, the middle term. Bennett gives the example using the syllogism, “All poodles are dogs (first premise); All dogs are animals (second premise); Therefore, all poodles are animals (conclusion).” In the conclusion, “poodles” is **S** and “animals” is **P**. The first premise is **S-M** and the second premise is **M-P**, where “dogs” is the middle term, **M**. Considering the possible arrangements, **S-M** or **M-S**, **M-P** or **P-M**, Aristotle called the “first figure” **S-M**, and **M-P**, the “second figure” **S-M** and **P-M**, and the “third figure” **M-S** and **M-P** for the premises. Only the “first figure” is valid, recognized Aristotle (“All dogs are poodles.” or **M-S**, for instance, is not true). Aristotle resolved these syllogistic moods by applying his Law of Non-contradiction and settled ultimately on the rules we now use.

Reductio ad absurdum as proof

Bennett describes the basic steps in any deductive proof. One begins with true or agreed upon premises and concedes at each step that the ensuing statement follows legitimately from the previous statement(s). The final statement is the

conclusion. A common type of proof is one of *reductio ad absurdum*, by which one arguer begins by accepting the opponent’s premise as true and argues logically by refutation to a contradiction by exposing inconsistencies in the opponent’s original argument, causing the opponent to give up his original premise as false. In all systems, agreement on the rules and definitions used in the system of interest is essential to maintaining consistency, so that reasoning in that system can be valid. For example, there is a difference between contradictories and contraries. Confusing one for the other can lead to invalid arguments and unsound conclusions.

Contradictories versus Contraries

Contradictories are paired statements such that both statements cannot be true and both cannot be false. For contradictories, one statement is universal (affirmation or denial) and its contradictory is particular (denial or affirmation, respectively). For example, “every person has enough to eat” is a universal statement, applying as it does to an entire set. The contradictory must be a particular (and not a universal) denial—“some people do not have enough to eat.” A particular applies to a subset, not to the entire set (humans in this example). In another example, “No individuals are altruistic” serves as the universal denial, while “some people are altruistic” serves as the particular affirmation in this contradictory pair. Aristotle pointed out that every affirmative statement has its own opposite negative and vice versa. For contradictories, one statement of the pair will always be true and the other false. Contradictories are represented by **A** with **O** and **E** with **I** in **A**ffirmo/**nEgO** terminology.

Contraries consist of opposite pairs in which both the affirmation and denial are universal or both are particulars. For contraries, both cannot be true, but it is also possible that neither is true. For example, “All people are rich” and “No people are rich” are contraries. Contraries are represented by **A** with **E** and **I** with **O** in the **A**ffirmo/**nEgO** terminology.

Bennett credits John Stuart Mill with noting that people frequently confuse contradictories with contraries and that this confusion also occurs in one’s private thoughts. He opined that if people made these pairs of statements aloud they may detect their errors.

Bennett gives as example the common complaint “Nobody around here helps out.” The contradictory is that “Some of us help out.” and the contrary is “We all help out.” How often is the appropriate affirmation used in the heat of complaint?

Using quantifiers in logic

Quantifiers, mentioned in an earlier section in this essay, may be expressed explicitly or implicitly, and this variation in usage can lead to alternate interpretations. Consider the uni-

versal quantifier “all.” Bennett gives the example, “Members in good standing may vote.” “All” is implied. The article “a” may be used as a universal quantifier, as in the example, “A library is a place to borrow books.” implying “all” libraries. Confusion may arise in deduction logically by using different universal quantifiers. Bennett cites a study in 1989 by David O’Brien in which people of various ages (second graders, fourth graders, eighth graders and adults) were tested for their understanding of the universality of “all,” “any,” and “a” used in propositional statements. States Bennett, “Without exception, in every age group the tendency to err was greatest when the indefinite article a was used, “If a thing . . .” For older children and adults, errors decreased when any was used, “If any thing . . .,” and errors virtually vanished with the universality was made explicit, “all things . . .” With the youngest children, though the errors did not vanish, they were reduced significantly when the universality was made clear with the word all.”

Problems with converse statements

Beginning with the basic form of a propositional statement, “quantifier subject (S) predicate (P)”, the statements “All S are P” and “All P are S” are converse statements. People may think they mean the same thing, but that is not a correct interpretation. Bennett states that conversion is a common error during argument. Both statements may be true, but they are not equivalent statements. Bennett gives examples “All mothers are parents” versus “All parents are mothers.” The first statement is true, but the second is not true. “All dogs love their owners” versus “All (dog) owners love their dogs.” Possibly neither statement is true. Bennett describes a study in which children of varying ages were asked about a series of drawings of squares and circles. The children were shown a picture in which, from left to right, were drawn a gray circle, a white square, a gray circle, a white square, a gray square, a gray circle, a gray circle, and a gray square. The children were then asked questions such as “Are all the squares white?” (considered by the examiners an easy question), “Are all the circles gray?”, and “Are all the white ones squares?” (considered by the examiners a more difficult question). States Bennett, “The youngest subjects converted the quantification 50% of the time, thinking ‘All the squares are white’ meant the same as ‘All the white ones are squares’ . This may be explained in part by the less developed language ability of the youngest children (ages 5-6), but their mistakes may also be explained by their inability to focus their attention on the relevant information [such as just squares or just white things] . . . By ages 8 and 9, children were able to correctly answer the easier questions 100% of the time and produced

the incorrect conversion on the more difficult questions only 10 to 20% of the time.”

Another factor involved in reasoning is that of familiarity, or lack thereof, with the subject being reasoned about. States Bennett, “The rules of inference dictating how one statement can follow from another and lead to logical conclusions are the same regardless of the content of the argument. Logical reasoning is supposed to take place without regard to either the sense or truth of the statement or the material being reasoned about. Yet, often reasoning is more difficult if the material under consideration is obscure or alien.”

Recall from the essay in this series on Reasoning Gigerenzer’s comparison of performance of test subjects on the Wason Selection Task. One group was asked to reason about the cards “D,” “E,” “3,” and “4.” while the other group was asked to reason about the cards “subway,” “Arlington,” “cab”, and “Boston.” Gigerenzer’s hypothesis was that, if logic is “natural” thought, most test subjects should perform correctly whether the problem was abstract (as in the D, E, 3, 4 scenario) or whether the problem concerned social interactions (as in the subway, Arlington, cab, Boston scenario). The tests were quite similar from a logical viewpoint. When the four cards on the table read “D,” “E,” “3,” and “4.” the conditional statement to be evaluated logically was “If there is a ‘D’ on one side of the card, then there is a ‘3’ on the other side.” The test subjects were instructed to turn over any cards necessary to test the validity of the conditional statement. Logic would dictate that, faced with a statement “If P, then Q,” one must rule out, “P and not Q.” The appropriate cards to turn over are “D” and “4.” In other scenario, when the four cards on the table were “subway,” “Arlington,” “cab,” and “Boston”, the conditional statement to be evaluated was “If a person goes to Boston, then he takes the subway.” The cards to turn over are P (Boston) and not Q (cab). Gigerenzer found that 10% of test subjects answered correctly in the abstract scenario and 30-40% of test subjects answered correctly in the social scenario, a much better performance; however, one might be tempted to conclude from this study that, even under the best of circumstances, less than half of us can think logically.

Bennett points out that familiarity is not always a help. She describes a study performed in 1928 by M. C. Wilkins. The premise put forth to Wilkins’ students was “All freshman take History I.” States Bennett, “. . . only 8% of her subjects accepted the conversion, ‘All students taking History I are freshmen.’ However, 20% of them accepted the equally erroneous conclusion, ‘Some students taking History I are not freshmen.’ With strictly symbolic material (All S are P), the errors ‘All P are S’ and ‘Some P are not S’ were made by 25% and 14% of the subjects, respectively. One might guess that in the first instance students retrieved common knowledge about their world—given the fact that all freshmen take History I does not mean that only freshmen take it. In fact,

they may have themselves observed nonfreshmen taking History I. So their conclusion was correct and they were able to construct a counterexample to prevent making the erroneous conversion. However, as they continued thinking along those lines, knowledge about their own world encouraged them to draw a (possibly true) conclusion that was not based on correct logical inference. ‘Some students taking History I are not freshmen’ may or may not be true, but it does not follow logically from ‘All freshmen take History I.’

Negation

The interplay between language and logic can prove especially troublesome. As Bennett pointed out earlier, statements that are equivalent in meaning by logic are not always interpreted as equivalent by a human reasoner. Negation, or saying that something is not, proves difficult at times for humans to process. Bennett explains that Aristotle recognized early on in his studies that propositions meaning the same thing can be explained as either an affirmation or a negation. For example, “all humans are imperfect” is the affirmation, while “no humans are perfect” is a negation with an equivalent meaning. It is possible to affirm the absence of something or to deny the presence of something; thus, the same set of facts can be presented by affirmation or negation.

How facts are presented, however, affects how people understand them. Bennett describes a study in which test subjects were asked to perform certain written tasks. Test subjects were timed and their accuracy assessed. Basically, the same instruction was written as an affirmative, an implicit negative, and an explicit negative. Papers were given to the test subjects, each with the Arabic numerals 1 through 8 listed consecutively at the top of the page. One instruction said, “Mark the numbers 1, 3, 4, 6, 7.” The next instruction said, “Do not mark the numbers, 2, 5, 8, mark all the rest.” The third instruction said, “Mark all the numbers except 2, 5, 8.” States Bennett, “The subjects performed the task faster and with fewer errors of omission following the affirmative instruction even though the list of numbers was considerably longer.

Subjects performing the task using ‘except’ were clearly faster than those following the ‘not’ instruction, signifying that the implicit negatives were easier to understand than the instructions containing the word ‘not.’” Bennett adds about the understanding of implicit negatives, “Some negatives do not have an implicit negative counterpart, and those negatives are more difficult to evaluate. The statement, ‘The dress is not red’ is harder to process than a statement like ‘Seven is not even’, because the negation ‘not even’ can be easily exchanged for ‘odd’, but ‘not red’ is not easily translated [and very difficult to visualize]. The difficulties involved with trying to visualize something that is not may well interfere with one’s ability to reason with negatives. If I say that I did not come

by car, what do you see in your mind’s eye? It may be that, wherever possible, we translate negatives into affirmatives to more easily process information.”

Double negatives can cause problems with processing information as well. In reasoning by *reductio ad absurdum*, we want to prove a proposition, **P**, but to do so we assume **not-P** and argue to a contradiction. We conclude “**not-not-P**” or “**P**.” Bennett points out that referendum questions at the voting booth often use wording that makes one’s choice seem counterintuitive and confuses voters. Her example is from a ballot managing her retirement funds. “Proposal: To stop investing in companies supporting gun control . . . voting ‘for’ means you are against gun control [are for easily obtainable guns] and voting ‘against’ means you favor gun control [are against easily obtainable guns] . . . If the process of negation involves an extra mental step, a double negative can be mind boggling. ‘The probability of a false negative for the pregnancy test is 1%’ or ‘No non-New Yorkers are required to complete form 203’ or ‘The statistical test indicates that you cannot reject the hypothesis with no difference’ can cause listeners to scratch their heads (or give them a headache).”

In a line from a story from John Mortimer’s *Rumpole of the Bailey*, Horace Rumpole, barrister at law, cross-examines an expert witness on the subject of bloodstains. The witness states, “The finding is not inconsistent with your hypothesis.” An exasperated Rumpole counters, “Does not ‘not inconsistent’, when translated to plain English, mean ‘consistent’?” The witness: “Yes, it does.” Common sense reigns, albeit temporarily, at last.

A brief recap

So far we have seen that, although rules have been discovered that assist us in reasoning logically so that we might evaluate “evidence” according to an ideal, problems with the use of language, resulting as it does in differences in interpretation, often interfere with the ideal of formal logic to which we aspire. Also, speaking from the standpoint of philosophy, we may not even understand fully the concepts of “truth” and “definability” since paradoxes exist involving those terms. There is disagreement about what sorts of things should be allowed to be considered evidence. Nonetheless, each of us on some level think we understand what “evidence” is and that “evidence” is, or at least should be, true and unassailable. This leads to a cognitive dissonance that simmers just under our awareness and causes problems for us as we interact with our fellow man.

Evidence Based Medicine

The cognitive dissonance we all experience relative to evidence spills necessarily over into our professional lives. Today,

we are all expected to practice so-called Evidence Based Medicine. To do otherwise, we are told, is to risk unnecessarily patients' lives and to cause too much money to be spent on healthcare.

So what is meant by evidence in the case of Evidence Based Medicine? Despite a well-meaning start, pretty much "evidence" in medicine has come to mean the outcome of a clinical trial, that outcome having been interpreted by an expert in the field.

At its outset, the premise underlying Evidence Based Medicine was that each physician, having completed his or her studies of basic medical science (anatomy, physiology, pathology, pharmacology and the like), would determine for each patient the best course of action by doing his/her own research of the literature. By doing so, each physician would break the bond to "authority." No longer would each physician "parrot" what s/he thought a mentor would do in a similar situation. S/he would find out the results of the latest studies and act according to the "best evidence."

What ensued, however, was a distortion of the ideal. Maya Goldenberg, in "Iconoclast or creed? Objectivism, pragmatism, and the hierarchy of evidence" in *Perspectives in Biology and Medicine*, states, "Objectivity is an epistemic virtue in science that stands for an aperspectival 'view from nowhere', certainty, and freedom from bias, values, interpretation, and prejudice. Even if objectivity cannot be achieved, it is perceived to be an ideal worth striving for."

Goldenberg explains that the idea for Evidence Based Medicine arose from the "pragmatism" movement in philosophy, a movement founded by Charles Sanders Peirce and William James. The doctrines underlying pragmatism are "(1) the meaning of concepts is to be sought in their practical bearings; (2) the function of thought is to guide action; and (3) truth is preeminently to be tested by the practical consequences of belief." Goldenberg quotes James stating that pragmatism stands for "the open air and possibilities of nature, as against dogma, artificiality and the pretense of finality in truth . . . [pragmatists] turn toward concreteness and adequacy, towards facts, towards action, and towards power. That means the empiricist temper regnant, and the rationalist temper sincerely given up . . . [pragmatism] does not stand for any special results [or doctrines]. It is a method only."

Goldenberg agrees that the allegiance of Evidence Based Medicine to the Randomized Controlled Trial captures the tenor of pragmatism; however, the strict adherence of Evidence Based Medicine to its hierarchy of evidence, which is based in a ranking of study designs based on the supposed rigor of their methodology, goes against the spirit of pragmatism.

Of hierarchies, Kirstin Borgerson, in "Valuing evidence: bias and the evidence hierarchy of evidence-based medicine" in *Perspectives in Biology and Medicine*, states that Evidence Based Medicine advocates "are not just claiming that it is

helpful to be able to distinguish, for instance, good from bad RCTs [Randomized Controlled Trials] or better from worse cohort studies. They have made an assumption about the necessity of ranking these methods against one another so that a critical review of the literature will produce one, hopefully decisive, answer . . . [there are multiple possible hierarchies] . . . no one has seriously (that is, explicitly and methodologically) argued for any particular hierarchy. Hierarchies are more often asserted than argued for."

Performance of Randomized Controlled Trials (RCTs) is a pragmatic methodology, writes Goldenberg, since the purpose of them is to "temporarily suspend prior knowledge of human physiology, disease, and pharmacology, all of which might allow for inferences regarding the effectiveness of a particular drug in treating a given condition, and instead determine whether a treatment works by trying the treatment in a large number of cases under controlled conditions." Goldenberg explains that RCTs are "ostensibly unhindered by the pre-theoretic expectations and commitments that can bias the deductive methods of basic science and the less systematic experimental methods of clinical experience and observational studies . . . Absent the influence of anticipated outcomes, scientists faced with recalcitrant empirical data should be more open to revising even well-established views about treatment efficacy."

Goldenberg points out that randomization and blinding are not always appropriate. Small sample size renders randomization futile, for example, and blinding is not always possible, as in studying the negative effects of smoking. Ethical considerations arise in studying effects of surgery, since sham surgical procedures would be required to truly blind interpretation of results. Furthermore, some social contexts may result in bleeding of effects of the study if, for example, study patients share their medications with friends participating in the study, hoping that everyone will benefit from a new therapeutic agent. In a blinded study, subjects may have an incentive to lie about their actions, if they fear being reprimanded for breaking the blinding effect of the study.

The hierarchy of Evidence Based Medicine asserts that Randomized Controlled Trials are **always** better evidence than alternative studies such as observational studies, even though different health research studies call for different designs so there really is no gold standard methodology. For example, Robin Bluhm, in "From hierarchy to network: a richer view of evidence for evidence-based medicine" in *Perspectives in Biology and Medicine*, states of studies of epidemiology, ". . . epidemiology is not generally an experimental science. The epidemiologist must try to identify causal factors 'in the wild' rather than in the controlled environment of the laboratory." Bluhm points out that progress in medicine requires the close interaction of epidemiologists, who attempt to establish causes of disease, and laboratory scientists, who can design

experiments to compare outcomes between study and control groups. The most common study design for epidemiology, asserts Bluhm, is the cohort study, in which groups, similar in most ways, differ in a finding of interest, and are observed and compared over time to determine if the finding of interest leads to a different incidence of a separate finding at the later time. For example, is smoking associated with an increased incidence of lung cancer, or is a high cholesterol level associated with an increased incidence of myocardial infarction?

Cohort studies are similar to Randomized Controlled Trials in that the populations are followed over time. In the cohort study one population is **exposed** to some factor, while in the RCT one population receives an **intervention**, such as a drug or surgical procedure. Other sorts of studies include case reports and cross sectional surveys. Different types of studies have varying strengths and weaknesses.

Goldenberg discusses the work of Upshur and Tracy, “. . . the entire edifice of evidence hierarchies is . . . based upon expert judgment or consensus. They charge that ‘the structuring of evidence according to hierarchy is by no means natural, intuitive, or even logically justified’ . . . Upshur and Tracy propose that the initial creation of an evidence hierarchy was intended to link the quality of evidence to the soundness of the recommendations based on the evidence . . . on the belief that these methods [RCTs and meta-analyses] are less susceptible than observational designs to bias. The key is the ability of randomization to eliminate selection bias and the unprovable claim that randomization balances all relevant known and unknown factors in a probabilistic sense. The hierarchy attributes lower reliability to expert judgment, and specifically subordinates theory and pathophysiological reasoning to designs with randomization. The reasoning behind the latter subordination is unclear, as pathophysiology often provides more fundamental understanding of causation and is in no way scientifically inferior. Thus, Upshur and Tracy conclude, the hierarchy has been advanced on the basis of expert opinion rather than reasoned argument—a move unbecoming of evidence-based thought and practice.”

Goldenberg asserts that adherents of Evidence Based Medicine have failed to recognize the fallibility of scientific evidence, preferring instead to undertake a “sort of absolutist search for certainty.” She discusses the work of Paul Feyerabend who “described science as being obsessed with its own mythology of objectivity and universality.” Goldenberg continues, “If the hierarchy of evidence was put in place to refute skepticism and ensure certainty, it stands as an example of what Feyerabend abhorred: science making claims to truth well beyond its actual capacity.”

Mark Tonelli, in “Evidence-free medicine: forgoing evidence in clinical decision making” in *Perspectives in Biology and Medicine*, opens his article, “At some time in the not-too-distant past, medicine was apparently practiced in the absence

of evidence. I know this to be the case because a Medline search for the term ‘evidence-based medicine’ prior to 1990 yields virtually no returns . . . By the time I had completed residency and fellowship in 1996, evidence-based medicine (EBM) had gone from ‘paradigm shifting’ concept to widely accepted dogma.”

Tonelli bemoans that establishing hierarchies of evidence has distracted the medical profession from its calling, which is to arrive at the best possible decision, be it diagnosis or treatment, for the individual patient. Says Tonelli, “. . . the myriad attempts to define, categorize, and stratify evidence for use in clinical medicine have only resulted in a epistemic and linguistic morass . . . Instead of arguing about what does or does not constitute evidence in support of a clinical decision, we should more carefully examine the central question regarding the legitimacy of various potential facts and warrants for clinical decision making.”

Tonelli notes that clinical medicine does not lend itself to formal, deductive logic, but utilizes the informal logic of argumentation. States Tonelli, “Clinical judgment requires the use of various kinds of reasoning. Done well, clinical decision making closely resembles the structure of argumentation, with careful consideration of many facts, warrants, backings, and rebuttals, ultimately resulting in a conclusion that is only probably, never demonstrably, correct.” Tonelli refers to the work of Toulmin, *The Uses of Argument*, noting, “. . . data, or basic facts, are often invoked as a foundation to a claim, but that facts alone are inherently insufficient to provide legitimate support to any claim.

In arriving at or defending a particular conclusion, we must go beyond producing facts to providing warrants, more general and hypothetical propositions that are necessary to have the particular fact support a particular claim.”

Tonelli opines that proponents of Evidence Based Medicine consider fact and warrant as a bundle—the fact serves as its own warrant. He gives the example that a certain pitcher may be considered the best based on a low Earned Run Average (ERA)—the fact. But, maintains Tonelli, the arguer of this claim must warrant that the ERA is superior to all other measures of pitching excellence to back his claim. Physicians, asked about their evidence for a certain decision, often merely cite a reference. They do not then warrant their use of that reference-datum by explaining why it is superior to other possible choices **in this particular patient at this time**. Tonelli insists that it is most important to show that the warrants that are invoked to support a claim based upon the facts in the reference are legitimate. He adds that establishing legitimacy of a warrant requires “an understanding of the underlying metaphysical and epistemic underpinnings of a healing discipline.”

Tonelli recognizes five broad classes of legitimate potential warrants: pathophysiologic rationale, results of clinical

research, clinical experience, patient goals and values, and system features. He notes that the relative importance of any warrant will depend on the specifics of the patient. A warrant from any of the five classes may be the most important in a particular case. Since the classes of warrants “differ from one another in kind, not in degree . . . no meaningful hierarchies of potential warrants can exist across the [classes].” Within a class, a basic hierarchy might be appropriate. For example, the clinical experience of a medical student would likely carry less weight than the experience of the attending generalist, whose experience would carry less weight than a specialist. However, even within a class of warrants, hierarchies “must be considered general and not prescriptive, serving only as guidelines, not clinical rules . . . often facts and warrants from each of the five classes are likely to be coherent, all supporting the same clinical decision . . . [but] when warrants conflict, supporting different conclusions regarding the right diagnosis or course of action . . . the clinician’s judgment is put to the test.”

Tonelli opines, “The demand for evidence distracts clinicians, requiring them to search outside of the clinician-patient relationship for answers to questions that are not directly applicable to the patient at hand . . . The answer to a great many important and interesting general medical questions can be discovered by searching, analyzing, and interpreting published clinical research. The answer to the question of what to do about the individual patient, however, does not exist outside of the clinician-patient relationship waiting to be discovered.”

Tonelli observes, “The consummate clinician is one who can identify all the relevant facts and warrants and, when necessary, negotiate between conflicting warrants by weighing each in the context of the particular patient at hand. The excellent clinician must, by necessity, have a well-developed knowledge base that includes understanding of biologic and physiologic concepts and principles, the relevant clinical research in her specialty, substantial personal experience, and, preferably, access to clinicians with even more. In addition to this knowledge base, the clinician must also have the skills and inclination to understand patient preferences, goals, and values, as well as an understanding of the facilitators and barriers to optimal care inherent in the system in which she practices. No wonder EBM’s simple five-step process has such appeal. Training clinicians to practice an evidence-free medicine is significantly more challenging than training them to practice EBM.”

Parenthetically, the “simple five-step process” of Evidence-Based Practice is as follows: 1) Formulate a well-built question, 2) Identify articles and other evidence-based resources that answer the question, 3) Critically appraise the evidence to assess its validity, 4) Apply the evidence, and 5) Re-evaluate the application of evidence and areas for improvement. Rela-

tive to step three, family practitioner Ross Upshur, in “Looking for rules in a world of exceptions,” in *Perspectives in Biology and Medicine*, noted that it had been estimated that “a physician would need to spend 627.5 hours just to read the 7,287 articles relevant to primary care **each month**. And that is just reading, never mind the “critically appraise” part.

So Evidence Based Medicine seems to be another distortion of the profession of medical practice, along with the expectation that no errors are tolerable or should occur while we care for our patients, that we can take care of the total needs of a patient in a seven minute office visit, and similar expectations with which we are all familiar.

Attributed to Albert Einstein is the saying “All things should be made as simple as possible, but not simpler.” Slobodkin has pointed out, in *Simplicity & Complexity in Games of the Intellect*, that it is human nature to simplify some aspects of life and to complexify (for example by ritual) other aspects of life. We all decry “error” in the practice of medicine. But how much of the aggregate of unexpected and less than desired outcomes is really due to error per se, how much is due to unreasonable expectation on the part of patients and ourselves as medical practitioners, and how much is due to a lack of understanding (because we humans simply have not yet made the discoveries) of basic science, clinical medicine, and principles of complexity theory or system theory?

It seems that perhaps Evidence Based Medicine has evolved into an attempt to simplify the practice of medicine too much.

Conclusion

In the end, it seems to me that “evidence” truly is, as defined, “that which justifies belief” and nothing more. We each feel justified in believing what we believe. If our belief comes to us by revelation or authority, we still consider it a belief and we feel justified in believing it. If this is true, then evidence is best considered just another word for a “reason” or “reasons.” Attempts have been made throughout the ages to give “evidence” a “higher calling” by insisting on proof that something labeled as “evidence” is true and warrants any conclusions drawn from that evidence. But, as we have seen in this essay, there are so many factors involved, from language to education to paying attention to the proper things at the moment of reasoning, that the formal rules discovered are likely to never be our “natural” process of thinking. Where most of us are concerned, the pretense of “evidence” to a scientific ideal is merely an illusion. As Garlikov might say, we have not learned the rules of the game to use evidence as ideal.

Evidence is the aggregate of reasons we use to justify our belief in some fact or claim. Those reasons may or may not follow the rules of argumentation and those reasons may not

actually lead an astute reasoner to conclude our claim. In the end, evidence does not guarantee truth or certainty. The label “evidence” applied to a fact does not make that fact any more likely to support a valid conclusion. What is important is the **process** we use to support a claim, the process of argumentation, by which we warrant **why** a piece of evidence is pertinent to the claim.

We hold in our minds the ideal that evidence is true and unassailable, but we can never reach that ideal and for a number of reasons. We are limited by our understanding of basic concepts, such as “truth” and “definability”, that lead unavoidably to paradox. We cannot possibly ensure that we understand the same message in exactly the same way, nor do we understand the most basic concepts of our language in the same way, for example dealing with “if/then” hypotheses, negations, and the like. We all live unavoidably with cognitive dissonance, believing contradictory statements, most often because we cannot drag the contradictions into our working memories simultaneously to recognize that they are contradictory.

Nonetheless, in a manner analogous to discovering the rules of logical thought, a process begun in earnest by Aristotle, following the principles of the scientific method, whereby hypotheses are tested according to a body of organizing principles and the test results are examined and interpreted independently by multiple investigators, seems to provide the best way to advance knowledge, while minimizing opportunities for inconsistency and contradiction. Argumentation and science share the same basic principles in that people with differing views are willing to look at a problem from different perspectives and are willing to risk being proved wrong in the interest of acquiring a common understanding of an issue.

Reasons, properly warranted, comprise the best form of evidence, evidence serving as a means to the end of gaining knowledge. However, human nature being what it is, we sometimes “put the cart before the horse” and think that the end justifies the means. If we imagine a desired end and then create a “story” that we label as “evidence” for the purpose of persuasion only, if we put consistently the goal of winning an argument ahead of better understanding, as Aristotle accused the Sophists of doing, then we are lost.

In a manner analogous to vaccination to minimize the spread of infectious disease, the vast majority of people should take the time to understand more fully the rules of logic and the common problems that interfere with thinking consistently and without contradiction. Then the populace at large would be at least in a position to know a reasoned argument when confronted with such an animal. Michael Shermer, as discussed in the essay in this series on Interpretation, has pointed out, “Anecdotal thinking comes naturally; science requires training.” It is up to each of us to do our part to further the

common cause of ensuring that our collective knowledge base is as consistent and noncontradictory as possible.

Upshur, in “Making the grade: assuring trustworthiness in evidence” in *Perspectives in Biology and Medicine*, quotes Alfred North Whitehead (1929), “The chief danger to philosophy is narrowness in the selection of evidence. This narrowness arises from the idiosyncracies and timidities of particular authors, of particular groups, of particular schools of thought, of particular epochs in the history of civilization. The evidence relied upon is arbitrarily biased by the temperaments of individuals, by the provincialities of groups, and by the limitations of schemes of thought.”

When confronted with inconsistencies and contradiction, we must try to discover the source of inconsistency or contradiction. We may find, by reviewing and applying assiduously the known rules of logic, an error in our proof. We can be aware of our biases and try to think more broadly. We can ask ourselves if we are being a bit slavish to authority. If a true paradox arises, we can try to expand the scheme of our thought and try to look at the problem from a new perspective from which paradox disappears.

Parenthetically, as an example, for Zeno’s paradox involving infinity, Zeno proposed that Achilles could never catch up to or pass the tortoise, who started with a 10-meter head start, in the race because he (Achilles) would only close a certain percentage of the gap in what mathematically is represented by an infinite series. Manipulating the formula mathematically gives a finite equation, showing that Achilles draws level with the tortoise when the tortoise has moved 1 and 1/9 meters. A better understanding of infinite series eliminated the paradox. I am not sure we are yet ready to solve the liar’s paradox, but theoretically, as suggested by Whitehead, we could look at the problem from another less limited perspective to solve the paradox. But, as Bolander explains, that will require a new understanding of “truth.” Until we better understand truth, we must stop thinking of “evidence” as true and unassailable. And we should understand that our collective knowledge base has necessarily items in it that will require revision in the future.

I think it is most important that we learn more about the limitations of humans and to ensure that as many of us as possible have access to that knowledge. Only then will we be able to design systems that are effective, including an effective healthcare system. That humans have imagination is both boon and bane. Imagination has enabled us to survive, to prepare for the future, and to dream of a better world. But the downside of this is that we are tempted to believe that what we imagine is true and/or attainable, regardless of whether or not we truly understand “reality,” or principles of Universal Law. We are in danger, as Goldenberg said of Feyerabend’s words, of making claims to truth well beyond our capacity.

Learning to accept “evidence” for what it really represents instead of what we want it to represent will be a start.

Summary

We tend to think of evidence that it must be of necessity true and unassailable, but that does not conform to reality. Evidence is really just a reason or reasons that we use to justify our beliefs. We try to hold evidence to an ideal, but because of our innate fallibility, which occurs as a direct result of our evolution within the environment with which we have co-evolved, we cannot ensure evidence meets that ideal. The nature of language and the use of that language by us is a major impediment to maintaining consistency and avoiding contradiction in our communication with one another. Additionally, our “natural” thought processes do not follow the rules of logic as discovered by logicians throughout the ages. This being the case, we must endeavor to understand better our human fallibility and learn to work within our capabilities. To do so, we must accept “evidence” as what it is instead of what we would like it to be.

Dr. Anderson is a retired general pathologist, residing in Carbondale, Illinois, and teaching part-time at her alma mater, Southern Illinois University School of Medicine.

Bibliography

Anderson C. On the nature of thought processes and their relationship to the accumulation of knowledge, Part II: The role of emotion, feeling, and consciousness. *Dermatopathology: Practical & Conceptual*. 2004;10(2):15.

Anderson C. On the nature of thought processes and their relationship to the accumulation of knowledge, Part III: The concept of truth. *Dermatopathology: Practical & Conceptual*. 2004;10(3):20.

Anderson C. On the nature of thought processes and their relationship to the accumulation of knowledge, Part VIII: Interpretation. *Dermatopathology: Practical & Conceptual*. 2006;12(4):23.

Anderson C. On the nature of thought processes and their relationship to the accumulation of knowledge, Part X: Causation. *Dermatopathology: Practical & Conceptual*. 2008;14(2):20.

Anderson C. On the nature of thought processes and their relationship to the accumulation of knowledge, Part XI: Reasoning. *Dermatopathology: Practical & Conceptual*. 2008;14(4):22.

Anderson C. On the nature of thought processes and their relationship to the accumulation of knowledge, Part XII: Detecting and using patterns. *Dermatopathology: Practical & Conceptual*. 2009;15(2):16.

Bennett D. *Logic Made Easy: How to Know When Language Deceives You*. New York: W. W. Norton & Co. 2004: pp 19-20, 30, 29, 31, 35-36, 40, 49, 54-57, 66, 73-79, 102,

Bluhm R. From hierarchy to network: a richer view of evidence for evidence-based medicine. *Perspect Biol Med*. 2005;48(4):535-47.

Bolander T. “Self-Reference.” *Stanford Encyclopedia of Philosophy* (summer 2008). Edward N. Zalta (ed). URL=<<http://plato.stanford.edu/entries/self-reference/>>

Borgerson K. Valuing evidence: bias and the evidence hierarchy of evidence-based medicine. *Perspect Biol Med*. 2009;52(2):218-33.

Cloud P. *Oasis in Space: Earth History from the Beginning*. New York: W. W. Norton & Co. 1988. p 47.

Devlin K. *Mathematics: The Science of Patterns*. New York: W. H. Freeman and Company. 1994. p 47, 74.

DiFate V. “Evidence.” *The Internet Encyclopedia of Philosophy* (2007). URL=<<http://iep.utm.edu/e/evidence.htm>>

Goldenberg M. Iconoclast or creed? Objectivism, pragmatism, and the hierarchy of evidence. *Perspect Biol Med*. 2009;52(2):168-87.

Kelly T. “Evidence.” *The Stanford Encyclopedia of Philosophy* (Fall 2008 ed). Edward N. Zalta (ed). URL=<<http://plato.stanford.edu/archives/fall2008/entries/evidence/>>

Pinker S. *The Stuff of Thought: Language as a Window Into Human Nature*. New York: Penguin Books. 2007: 72, 154-5.

Slobodkin L. *Simplicity & Complexity in Games of the Intellect*. Cambridge, Massachusetts: Harvard University Press. 1992. p 74.

Tonelli MR. Evidence-free medicine: forgoing evidence in clinical decision making. *Perspect Biol Med*. 2009;52(2):319-31.

Upshur RE. Looking for Rules in a world of exceptions: reflections on evidence-based practice. *Perspect Biol Med*. 2005;48(4):477-89.

Upshur RE. Making the grade: assuring trustworthiness in evidence. *Perspect Biol Med*. 2009;52(2):264-75.

Zarefsky D. *Argumentation: The Study of Effective Reasoning*. 2nd edition. Audio Course. Chantilly, VA: The Teaching Company. 2005. Lecture 2.