

# Three Techniques for Weighing Evidence to Reach a Conclusion

*Britten Coyne Partners*

In a radically uncertain world, the ability to systematically weigh evidence to reach a justifiable conclusion is undoubtedly a critical skill. Unfortunately, it is one that too many schools fail to teach. Hence this short note, which will cover some basic aspects of evidence, and quickly review three approaches to weighing it.

Evidence has been defined as “any factual datum which in some manner assists in drawing conclusions, either favorable or unfavorable, retarding a hypothesis.”

Broadly, there are at least four types of evidence:

- Corroborating: Two or more sources report same information, or one source reports the information and the other attests to the first’s credibility;
- Convergent: Two or more sources provide information about different events, all of which support the same hypothesis;
- Contradictory evidence is two or more pieces of information that are mutually exclusive, and cannot both or all be true;
- Conflicting evidence supports different hypotheses, but the pieces of information are not mutually exclusive.

Regardless of its type, all evidence has three fundamental properties:

- Relevance: “Relevant evidence is evidence having any tendency to make [a hypothesis] more or less probable than it would be without the evidence” (from the US Federal Rules of Evidence);
- Believability: Is a function of the credibility and competence of the source of the evidence;

- Probative Force or Weight: Is concerned with the incremental impact of a piece of evidence on the probabilities associated with one or more of the hypotheses under consideration.

There are three systematic approaches to weighing evidence in order to reach a conclusion.

In the 17<sup>th</sup> century, Sir Francis Bacon developed a method for weighing evidence. Bacon believed the weight of evidence for or against a hypothesis depends on both how much relevant and credible evidence you have, and on how complete your evidence is with respect to matters that you believe are relevant to evaluating the hypothesis.

Bacon recognized that we can be “out on an evidential limb” if we draw conclusions about the probability a hypothesis is true based on our existing evidence without also taking into account the number relevant questions that are still not answered by the evidence in our possession. We typically fill in these gaps with assumptions, about which we have varying degrees of uncertainty.

In the 18<sup>th</sup> century, Reverend Thomas Bayes invented a quantitative method for using new information to update a prior degree of belief in the truth of a hypothesis.

“Bayes Theorem” says that given new evidence (E), the updated (posterior) belief that a hypothesis is true ( $p(H|E)$ ) is a function of the conditional probability of observing the evidence given the hypothesis ( $p(E|H)$ ), times the prior probability that the hypothesis is true ( $p(H)$ ), divided by the probability of observing the new evidence ( $p(E)$ ).

In qualitative terms, we start with a prior belief in the probability a hypothesis is true or false. When we receive a new piece of evidence, we use it to update our prior probability to a new, posterior probability.

The “Likelihood Ratio” is a critical concept in this process of Bayesian updating. It is the probability of observing a piece of evidence if a hypothesis is true divided by the probability of observing the evidence

if the hypothesis is false. The greater the Likelihood ratio for a piece of new evidence (i.e., the greater the information value of the new evidence), the larger should be the difference between our prior and posterior probabilities that a given hypothesis is true.

In the 20<sup>th</sup> century, Arthur Dempster and Glenn Shafer developed a new theory of evidence.

Assume a set of competing hypotheses. For each of these hypotheses, a new piece of evidence is assigned to one of three categories: (1) It supports the hypothesis; (2) It disconfirms the hypothesis (i.e., it supports "Not-H"); or (3) it neither supports nor disconfirms the hypothesis.

The accumulated and categorized evidence can then be used to calculate a lower bound on the belief that each hypothesis is true (based on the number of pieces of evidence that support them, and the quality of that evidence), as well as an upper bound (equal to one less the probability that the hypothesis is false, again, based on the evidence that disconfirms the hypothesis, and its quality). This upper bound is also known as the plausibility of each hypothesis.

The difference between the upper (plausibility) and lower (belief) probabilities for each hypothesis is the degree of uncertainty associated with it. Hypotheses are then ranked based on their degrees of uncertainty.

While there are quantitative methods for applying all of these theories, they can also be applied qualitatively, to quickly and systematically produce an initial conclusion about which of a given set of hypotheses is most likely to be true.