Journal of Humanistic Mathematics

Volume 6 | Issue 2 July 2016

Simple Tools with Nontrivial Implications for Assessment of Hypothesis-Evidence Relationships: The Interrogator's Fallacy

Justus R. Riek
Hofstra University

Recommended Citation
Riek, J. R. "Simple Tools with Nontrivial Implications for Assessment of Hypothesis-Evidence Relationships: The Interrogator's Fallacy," *Journal of Humanistic Mathematics*, Volume 6 Issue 2 (July 2016), pages 179-199. DOI: 10.5642/jhummath.201602.13 .
Available at: http://scholarship.claremont.edu/jhm/vol6/iss2/13

JHM is an open access bi-annual journal sponsored by the Claremont Center for the Mathematical Sciences and published by the Claremont Colleges Library | ISSN 2159-8118 | http://scholarship.claremont.edu/jhm/

Simple Tools with Nontrivial Implications for Assessment of Hypothesis-Evidence Relationships: The Interrogator's Fallacy

Cover Page Footnote
The author wishes to acknowledge the assistance of Mark V. McEvoy (Hofstra University Philosophy Department) for helpful discussions during the early, formative stages of this paper.

Simple Tools with Nontrivial Implications for Assessment of Hypothesis-Evidence Relationships: The Interrogator's Fallacy

Justus R. Riek

Department of Computer Science, Hofstra University, New York, USA
jriek@optonline.net

Abstract

This paper takes a mathematical analysis technique derived from the Interrogator's Fallacy (in a legal context), expands upon it to identify a set of three interrelated probabilistic tools with wide applicability, and demonstrates their ability to assess hypothesis-evidence relationships associated with important problems.

1. Introduction

During the past 50 years, it has been well documented that humans suffer from a variety of cognitive biases. Kahneman and Tversky are generally credited with early ground-breaking research in this area [11], resulting in the award of a Nobel Prize in 2002. Although some of their conclusions (e.g., concerning base-rate neglect) have not gone unchallenged [8], it is today generally acknowledged that the human mind is susceptible to systematic cognitive errors. In response, the study of methods to improve information assessment and reasoning strategies has endeavored to mitigate these errors. An area that has received particular attention involves the issue of *overconfidence* in one's beliefs. Among the strategies which have evolved to address the problem of overconfidence is one referred to as *consider-the-opposite* [1, page 146]. This paper investigates the situation in which an individual is confronted with a solitary hypothesis (h) and wishes to evaluate the effect that a piece of evidence (e) may have on the probability of h being viewed as true, namely $p(h|e)$.

This paper's principal focus is on a set of simple tools with wide-ranging application which can be used to evaluate such conditional statements. The following sections will provide: (1) a brief discussion of the overconfidence issue and the benefit of considering a conflicting hypothesis, (2) an overview of the Interrogator's Fallacy from which the probabilistic tools are derived, (3) a generalization of the tools' range of applicability, and (4) practical examples to demonstrate their usefulness. The reader will see, in the practical application of these tools, that they can be collectively viewed as a concrete, formula-based approach within the family of *consider-the-opposite* strategies. Their advantage is that they are simple and explicit. Rather than a descriptive, subjective process (e.g., *consider-the-opposite*, Inference to the Best Explanation), these tools depend upon the use of a simple formula: a mathematical relation involving two conditional probability statements in which either (1) the relation holds or (2) it does not. In the material that follows, the tools will be shown to constitute an explicit, formula-based contribution to improved information assessment and decision making.

This material involves two distinct but related background issues: (1) purely formal results derived from mathematical probability as axiomatized by Kolomorgorov [13] and (2) the interpretation of these results. Generally speaking, there is agreement concerning the validity of the mathematical derivation of probability formulas and relations. However, with regard to interpretation of probabilities, opinions differ, and this paper will concentrate on two: (1) the classical interpretation of probability as a ratio between favorable-cases and all-possible-cases as standardized by Laplace [14] and (2) probabilities as subjective degrees of belief, an approach stimulated by the early work of Borel [2], Ramsey [17] and de Finetti [7]. It is well known that a duality exists between (1) the laws of probability based on a set-theoretic approach and (2) the calculus of belief as applied to sentences within a language [10, pages 14–15].

2. The Importance of the Negation of a Hypothesis

People adhering to a hypothesis (or dearly held belief) seem often to be guilty of overconfidence in the correctness of their beliefs. This is a topic that continues to receive much attention, comprising, for example, all of Part III (titled "Overconfidence") in Kahneman's recent book [12, pages 199–265]. It has been suggested by researchers that at least one cause of this behavior is the failure to seriously consider and confront alternative hypotheses, especially alternative and conflicting hypotheses. For example, Robyn Dawes, in his highly acclaimed book, provides an extended discussion of three historically significant examples [4, Chapter 11 titled "Giving Up," pages 230-253] in which people maintained (i.e., failed to "give up") unwarranted beliefs held over long periods of time while failing to attend to substantial, well publicized, disconfirming evidence. The three examples involve: (1) a failure of military intelligence, (2) use of the Rorschach inkblot test and (3) graphology (handwriting analysis). In the process of evaluating what went wrong in each of the situations, Dawes notes that an important and common characteristic within this group of examples is the lack of an alternative hypothesis. He concludes that what is needed to combat such lapses in the rational evaluation of beliefs is something to contrast with the dearly held belief: a conflicting hypothesis. Fortunately, if one is working with a hypothesis, h, then an immediately available *conflicting* hypothesis is simply its negation, $\neg h$. The tools discussed below explicitly force one to confront and evaluate $\neg h$.

3. The Interrogator's Fallacy

This section initiates the discussion of what will ultimately be seen to be a set of three related tools that may be productively applied to evaluate a hypothesis-evidence relationship. Each of the tools forces a comparative evaluation involving the negation of the original hypothesis. The discussion starts with the introduction of a paper by Matthews called "The interrogator's fallacy" [15]. The analysis technique used by Matthews, in a legal setting, will be used as a model to develop the above mentioned set of tools. The reader will see that we are more interested in the analysis technique used in [15] than we are in the fallacy itself.

The following discussion deals with the use of a confession in conjunction with the evaluation of the guilt or innocence of an individual accused of a crime. In times past, it is to be noted that many cultures and civilizations have viewed the confession as a powerful piece of evidence, even to the point of requiring no corroborating evidence [15, page 3]. Over time this attitude has softened. For example, extracting a confession under duress is now inadmissible. But it is nevertheless true that the legal system of today remains generally convinced of the value of confessional evidence [15, page 3]. Confessions continue to play a role in a significant number of legal cases. So now following the Matthewsf paper and its Bayesian approach, consider the situation in which a crime has been committed, evidence has been gathered by the police prior to any interrogation, and a person is accused of the crime. Let $p(G)$ be the prior probability of the guilt of the accused based on the evidence gathered, where G is the event that the person is guilty. Suppose that, subsequently, the accused person confesses to the crime. Let $p(G|C)$ be the posterior probability of guilt, where C is the event that the person has confessed. Now, if the confession is to increase the probability of the person's guilt, then it must be true that

$$p(G|C) > p(G). \tag{3.1}$$

However, using straightforward manipulations of simple conditional probabilities, Matthews proves the following: If one wishes to assert that $p(G|C) > p(G)$, then it necessarily follows that

$$p(C|\neg G) < p(C|G). \tag{3.2}$$

In other words, it is a *necessary* consequence of Equation (3.1) that it must be more probable that a guilty person will confess than that a not-guilty, $\neg G$ (innocent), person will confess, all under the same conditions. The mere existence of a confession does not automatically guarantee increased probability of guilt.

The Interrogator's Fallacy sheds light on why, and under what circumstances, a confession may constitute evidence of questionable value. Before continuing with our generalization of the above inequalities, let me stop for a moment to explain why. Briefly, psychological research has identified certain personality traits (e.g., *interrogative suggestibility* and *compliance*) which affect how individuals behave during a police interrogation. Simply put, individuals who are highly suggestive and compliant may exhibit, among other things, a lack of assertiveness and a tendency to show obedience to authority [15, page 4]. Matthews provides a more detailed explanation of situations (especially involving cases of terrorism) in which the inequality of (3.2) above is actually *reversed*. Matthews notes that the final effect is that the confession by an innocent party may (surprisingly) constitute a *disconfirming* piece of evidence which reduces the probably of guilt, and he names the phenomenon the Interrogator's Fallacy.

Note well: Use of the tools described in this paper may depend on the ability of a person to perform *relative* probability assessments. Ask people to estimate the numerical probabilities (on a ratio scale) of both an automobile accident, p(auto accident) and an airplane accident, p(airplane accident), and they will invariably give answers which are not only incorrect, but also highly variable. However, ask people to judge the relative magnitudes of the two probabilities (on an ordinal scale) and most will answer that $p(\text{auto accident}) > p(\text{airplane accident})$. They would be correct. In fact, many people will go even further and state that the probability of an auto accident is much greater than an airplane accident. Generally, people's answers to relative-magnitude questions will be more accurate than numerical probability estimates.

In this paper, the importance of the Interrogator's Fallacy is related only little to its legal implications. The justice system of today generally recognizes the possibility that a confession may be unreliable and that additional evidence should be required in order to prosecute. Happily, the work by Matthews provides scientific support for this recognition since it offers a formal, mathematical justification for the skepticism. However, of equal importance is the recognition that what Matthews did, analytically in a legal setting, can be used in other application domains to provide insight by *restructuring* problems. In addition, the Matthews result generalizes directly to a set of three simple formulas (the "tool set") with widespread application. The next section is concerned with generalizing the applicable range of Matthews' result.

4. Expanding the Useful Range of Matthews' Result

It turns out that the *necessary* condition (3.2) given above by Matthews is also *sufficient*. This means that (3.1) is logically equivalent to (3.2). As such, the legal result by Matthews represents an instantiation of the first item in the following set of three historically standard definitions of Bayesian hypothesis confirmation pertaining to (1) confirming evidence, (2) disconfirming evidence and (3) neutral (independent) evidence [10, page 92]. For ease of reference, the three definitions are stated below (LHS) together with their derived equivalencies (RHS).

- Confirming evidence:

$$p(h|e) > p(h) \text{ iff } p(e|\neg h) < p(e|h) \tag{4.1}$$

- Disconfirming evidence:

$$p(h|e) < p(h) \text{ iff } p(e|\neg h) > p(e|h) \tag{4.2}$$

- Neutral evidence:

$$p(h|e) = p(h) \text{ iff } p(e|\neg h) = p(e|h) \tag{4.3}$$

The following provides a brief proof of (4.1) as an indication of how the other two equivalent formulations are derived:

(Necessary) First note that it is possible to represent $p(h|e)$ in the following manner [10, page 21]:

$$p(h|e) = p(h)/[p(h) + [p(e|\neg h)/p(e|h)]p(\neg h)].$$

To insure that the conditional probabilities are well-defined, assume that each of $p(e)$, $p(h)$ and $p(\neg h)$ is non-zero. The inequality $p(h|e) > p(h)$ in (4.1) is then equal to the following inequality

$$p(h)/[p(h) + [p(e|\neg h)/p(e|h)]p(\neg h)] > p(h).$$

Solving for the quantity $[p(e|\neg h)/p(e|h)]$ gives

$$1 > [p(e|\neg h)/p(e|h)],$$

which implies

$$p(e|h) > p(e|\neg h).$$

(Sufficient) Starting with the necessary condition, it is a simple matter to reverse the above steps.

I wish to emphasize three important points:

1. The domain of application for these three equivalency formulations is wide indeed. In theory, they can be applied to test any hypothesis-evidence relationship whose equivalent formulation is well defined. However, there is even more to say regarding this claimed domain of applicability. In some cases, the tools can be applied to a problem not

obviously presented in the form of a solitary hypothesis-evidence relationship (see Section 5.3). It may be possible to take the problem and recast it in the form of a hypothesis. Next it may be possible to identify a piece of evidence and then use the formulas to evaluate the *manufactured* hypothesis-evidence relationship. The point to be emphasized is that the tools may be even more widely applicable than initially appreciated.

2. As mentioned above, the Matthews result reminds us of a time-honored conceptual approach for reasoning about a problem: it *reframes* the problem. That is, it takes an initial problem such as (3.1) above, and transforms it into an apparently different problem, such as (3.2) above, while introducing the additional component $\neg h$. The original problem is transformed into a new and different view of what turns out to be logically equivalent to the original. As is well known, viewing things from a different perspective frequently produces insight. For example, in Matthews' legal example, many (most?) people would be prepared to accept the truth of (3.1) without reservation, until they are exposed to the necessary consequence (3.2). This may lead them to reason more deeply (thoughtfully) and recognize the possibility that the inequality in (3.2) may not hold in all cases. The three formulas given above may be viewed as concrete, Bayesian, reason-guiding tools.
3. Finally, as will be demonstrated below, there is value to be gained by forcing an explicit consideration of $\neg h$ into the discussion. Simply making the effort to logically negate a statement-of-hypothesis may create important insight (see Section 5.1). Furthermore, attending to $\neg h$ is not only useful but actually mandatory. If one accepts the definitions of Bayesian hypothesis confirmation, then the equivalencies to the right of iff would seem to require an analysis involving $\neg h$. Perhaps surprisingly, it will be shown below that researchers and decision makers often ignore $\neg h$, as well as the need to gather the information (evidence) required for its valuation. It will become clear that ignoring $\neg h$ can lead to very bad consequences.

5. Examples and Applications

In the discussion that follows, I refer to the above three formulas as the "tool set." By that I mean not only the three definitions of Bayesian confirmation together with their corresponding equivalent formulations, but also the closely related admonition (i.e., *operational philosophy*) to *consider-the-opposite*: $\neg h$. The importance of the tool set will be demonstrated by its ability to guide the analysis/solution of a variety of difficult and important problems. The solution process may not always be the most direct and will sometimes border on the inelegant. However, the tools do seem to constitute a comprehensive, single-source capable of guiding the reasoning process in solving problems that have confused and even baffled people over the years. The tools may be simple, but they work.

5.1. The Wason four card problem

The four card problem is easy to state and yet difficult for most people to solve [19, pages 273–281]. Conceptually, it may be viewed as a psychology experiment involving a participant seated at a table on which four cards have been placed. The four cards are showing A, B, 2 and 3 respectively. The participant is told that each of the cards has a letter on one side and a number on the other side. The experiment coordinator then makes the following assertion: “All cards with a vowel on one side have an even number on the other side of the card.” The participant is asked which cards should be turned over in order to test the validity of the assertion.

Over the past 45 years, this experiment (and variations on it) has been repeated thousands of times, both (1) formally, in rigorously conducted experiments to investigate the dynamics of why people answer the way they do, and (2) informally with groups of students in psychology, logic, and critical thinking classes. The results are consistent over time. In the version of the experiment described above, the great majority of participants (approximately 80%–90%) do not arrive at the correct solution. For example, Dawes, himself a mathematical psychologist, informally administered the problem to five colleagues who were also highly regarded (unnamed) PhD mathematical psychologists. Only one of the five solved the problem correctly [3, page 126].

While a great deal of research and commentary has been generated in a search for an answer to the question of why so many people do not arrive at the correct solution, I have a different goal in this section. I choose not to become embroiled in the discussion of “why?”, but will instead concentrate on one particular method (among several possible) to solve the problem correctly utilizing the *operational philosophy* of the tool set; we will consider–the–opposite by explicitly negating the statement of hypothesis.

The assertion by the experiment coordinator can be recast as an if-then hypothesis statement h: "If a card has a vowel on the letter-side (p), then the card has an even number on the number-side (q)" which is merely $p \to q$, and which is equivalent to $\neg p \vee q$. We next form the negation of the hypothesis to see what it may reveal: $\neg h = p \wedge \neg q$. This negated hypothesis is easily translated back into English as: "vowel on letter-side of card and odd number on number-side of card." Given the choices available among the four cards, the answer suddenly becomes clear without any further effort; turn over the cards marked with a vowel (A) and an odd number (3). Merely translating h into $\neg h$ provides the insight necessary to solve the problem. The solution is simple, but it eludes most of the people who attempt the problem.

5.2. The NASA space shuttle Challenger *disaster*

This example involves the United States space agency (NASA) and the space shuttle *Challenger* disaster of 1986. Following a high-level summary discussion by Dawes, recall NASA's worry involving the possibility of malfunctions associated with O-rings during cold temperature operations [6, pages 8–9]. In preparation for the launch, concern over the issue of low air temperatures led to a request for engineers to graph the temperatures at launch time for previous flights in which O-ring problems had occurred. Two figures are shown below; Figure 5.1 and Figure 5.2 [6, page 9].

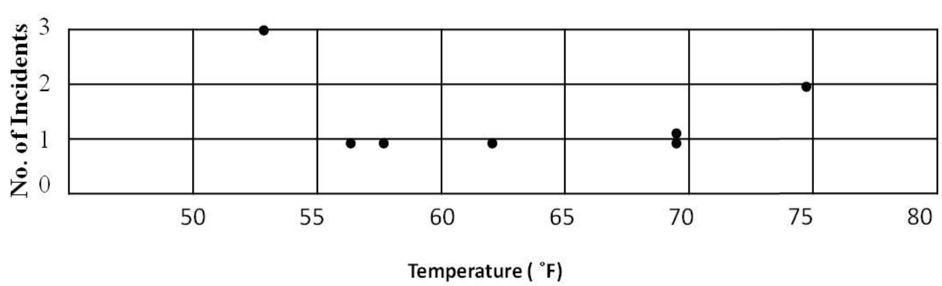

Figure 5.1: O-ring performance against temperature for problem launches

Figure 5.1 plots temperatures vs. frequency of incidents (black dots) for launches in which problems had occurred. This is the graph that was requested.

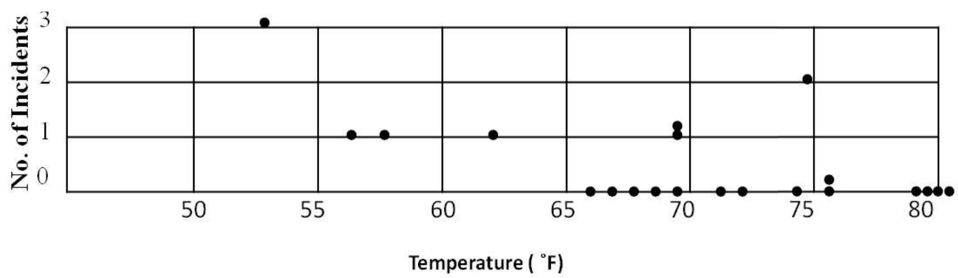

Figure 5.2: O-ring performance against temperature for problem and no-problem launches

Figure 5.2 also plots temperatures vs. incident frequency, but includes the additional launches in which *no problems occurred*. This is the graph that was not requested. Figure 5.1 shows only problem launches, while Figure 5.2 shows all launches. When viewing data in Figure 5.1 for the 10 of the previous 24 launches in which incidents were recorded, there does not appear to be any particular relationship. A person looking at this plot might easily believe that frequency of incidents does not depend on temperature. However, when the 14 of the previous 24 launches in which there were no incidents are included (Figure 5.2), there is a powerful incentive to believe that a relationship between incidents and temperature does truly exist. In spite of the small sample size, incident-free launches seem much more likely to take place above 65 °F.

Imagine what might have happened if an analysis of the situation had taken place using one of the tools in Section 4. Someone looking only at Figure 5.1 (which is the only plot that had been requested and created) and speculating that incident frequency is independent of temperature might proceed as follows. Let the set of launches be viewed as composed of two subsets: (1) launches with a non-zero incident frequency and (2) launches with a zero incident frequency. If one sets I to be the event that the launch has an incident, and T to be the event that the launch temperature ranges from 50 °F to 80 °F in (4.3) above, then under the belief that the frequency of incidents does not appear to be dependent on temperature, one might believe that

$$p(I|T) = p(I). \tag{5.1}$$

If this equality is believed to be true, then its equivalent formulation using (4.3) above must also be true:

$$p(T|\neg I) = p(T|I). \tag{5.2}$$

Given that the graph in Figure 5.2 had not already been created, the need to verify the equality in (5.2) would have led to the gathering/plotting of the necessary data (i.e., noting the temperatures for launches with zero incident frequency). Plotting the data for Figure 5.2 would have cast serious doubt on the truth of (5.2). The distribution of temperatures in the absence of incidents $\neg I$ is substantially different from the distribution of temperatures in the presence of incidents I. If (5.2) does not hold, it follows that (5.1) does not hold. Incidents of O-ring problems at launch time are not independent of temperature. The use of a tool that sensitizes one to adopt a *consider-the-opposite* attitude and, further, encourages one to challenge an analysis that does not reflect insight into "what is the situation when incidents do not occur?" might have had a significant impact on the decision to launch. A disaster might have been averted. Note that the air temperature at the time of launch was 36 °F.

Dawes describes an analysis which fails to compare problem launches with problem-free launches as irrational [6, pages 8–11]. He then goes on to label a conclusion based on such a flawed analysis as pseudodiagnostic. Other examples of pseudodiagnosticity identified by Dawes involve:

1. Investigations of problems in the United States at nuclear power plants in which great effort is made to find the cause of the *particular* problem in question (h), without making the comparison to other situations and behaviors in which no problem occurred ($\neg h$).
2. Failure in the U.S. to investigate safe airplane landings ($\neg h$) and compare them to a landing when an airplane crashes (h). For example, the great emphasis placed on a detailed analysis of the crashed aircraft's cockpit flight recorder ("black box") does not provide any such comparison.
3. The tragic episode in the U.S. of Kip Kingel who killed his parents with a gun and went on a high school shooting spree in Springfield, Oregon in 1998. In the aftermath, analysis concentrated on Kip himself and his family life (h), while ignoring any in-depth comparison of Kip to high school, adolescent peers who did not engage in violent behavior ($\neg h$).

There is more. The work of Russo and Schoemaker documents situations in which "people often ignore important feedback about their failures because feedback about successful decisions seems somehow more relevant" [18, pages 204–205]. Consider for example:

1. The product manager at a major U.S. consumer goods company who admitted that "Of the new products we decide to field test, only the successes are evaluated. We don't have time to determine why the other product ideas failed" [18, page 204]. Therefore, failures were not compared to successes.
2. A nonprofit U.S. social service agency, which, when trying to characterize what constituted a successful grant application, realized that the applications concerning unsuccessful grants (90% of the proposals written) had been destroyed. Only files for the winning applications had been retained. No comparison was possible.
3. A large U.S. bank, which, when it decided to study why so many of its loans to small/medium companies had "gone bad," restricted its analysis to the bad loans and failed to contrast them with the good loans.

The conclusion is clear. It is not uncommon for investigators of important problems to fail to explicitly address $\neg h$ in a meaningful way. This failure will lead to a flawed or incomplete analysis of the situation. Furthermore, at the end of the analytical process, it is likely that any "answer" obtained is actually an answer to the wrong question. Tools utilizing easily applied formulas, which are reason-guiding and encourage investigators to consider the negations of their hypotheses, will help to improve the quality of decisions.

5.3. A questionable claim regarding child sexual abusers

The purpose of this example is to demonstrate how the tools may be used to examine a claim made in the U.S. by some licensed PhD clinical psychologists regarding the treatment of child sexual abusers (pedophiles). The questionable claim is this: "If there is one thing we know about child abusers, it's that they never stop on their own without therapy" [5, page 126]. This quote is reported by Dawes who was himself once preparing to become a clinical psychologist but subsequently withdrew in his second year of clinical training to pursue research as a mathematical psychologist. Similar quotes are found in [4, page 4] and [9, pages 3–4]. The tool set can be used to demonstrate that, with high probability, clinical psychologists who make the above assertion possess merely a belief; a belief which is neither *justified* nor *true*.

To render the claim susceptible to analysis using the tool set, I rephrase it in the following form: "If a person has previously committed child sexual abuse, then the person will not stop unless the person undergoes psychological therapy." This may be more compactly stated as: "If the abuser does not receive therapy, then the abuser does not stop." But this statement is logically equivalent to: "If the abuser does stop, then the abuser has received therapy." So, if the abuser has stopped because of therapy, it is certainly the case that the probability of stopping is increased by the application of psychological therapy. This final reformulation implies the following probabilistic statement:

$$p(\text{stop}|\text{therapy}) > p(\text{stop}). \tag{5.3}$$

Therefore, the original claim implies that therapy increases the probability of stopping. We now test (5.3) using (4.1) from the test set. The reframed, equivalent formulation is:

$$p(\text{therapy}|\neg\text{stop}) < p(\text{therapy}|\text{stop}). \tag{5.4}$$

The strategy here is to identify abusers drawn from the populations associated with LHS and RHS and use them to create a contradiction to the inequality in (5.4). This contradiction is then used to work backward, up the chain of inferences, to show that the original claim is most probably false.

Consider the set of abusers in the LHS of (5.4) that do not stop abusing (¬stop). With regard to these abusers, the probability of entering therapy to mitigate abusive behavior, given that they do not stop, is probably large. The longer abusive behavior goes on (i.e., does not stop), the greater is the chance of these individuals being found out (e.g., caught in the act, accused/exposed by a growing number of abuse victims), arrested and finally forced into therapy as part of a criminal sentence or court-mandated condition of parole. The probability of entering into therapy is large.

Consider the set of abusers in the RHS of (5.4) that do actually stop abusive behavior (stop) without therapy. But how can an abuser stop without therapy? Is it possible? By what mechanism could it be possible? Fortunately, this reformulation of the problem at least forces one to begin to think about the issue. One of the possibilities that comes to mind is that, perhaps, some of the abusers have simply summoned the courage and personal discipline to quit on their own: they may have self-quit. I claim that this set of former abusers is non-empty with non-zero probability. If any reader is skeptical of the ability of a child abuser to stop without therapy (e.g., self-quit), consider the following observation. Some people do actually self-quit very serious addictions (e.g., smoking, over eating/obesity, drugs, alcoholism, etc.). Specific statistics are not so readily available, but we all know that it certainly does happen. Given this insight, it is reasonable to believe that some child sex abusers might also self-quit. Thus, the probability of some abusers stopping (self-quitting) is not zero which means that the conditional probability on the RHS of (5.4) is well-defined. We next examine the magnitude of the RHS of (5.4). Note that the probability of entering abuse-mitigation therapy given that these persons have stopped without therapy (e.g., self-quit) is probably small. After all, the persons in question have actually stopped the abusive behavior. Entering child-sexual-abuse-therapy voluntarily would involve admitting a history of child sexual abuse, an admission that most people would be hesitant to make, either because of personal embarrassment or for fear of being reported to the legal authorities. The probability of voluntarily entering into therapy is small.

If the reader finds the above discussion unconvincing, then consider the following minimalist example. First, select one particular individual from the population associated with LHS who is a habitual abuser (labeled HA) over a long period of time and who does not stop (¬stop) the behavior. Next, select one particular individual from RHS, who abused a child one time (labeled OT), was greatly disgusted/upset by the incident, did truly stop and never abused again (stop). Now compare the relative probabilities of each receiving therapy. Clearly, it is likely that the habitual abuser (HA) will enter therapy and more likely that the one-time abuser (OT) will not enter therapy.

The contradiction now follows. Since it is almost certain that LHS is relatively larger than RHS (most certainly true for the single pair of individuals OT and HA), the inequality in (5.4) almost surely does not hold. Given this analysis, it may now be concluded: the claim in the introductory paragraph is very likely false. It is very likely true that these clinical psychologists cannot know what they claim to know.

6. The Monty Hall Problem

The problem is patterned after a popular TV game show in the United States in which a contestant selects (but does not open) one of three closed doors, behind one of which is concealed a valuable prize. One door conceals a new automobile, while the other two doors each conceal a goat. Subsequently, the game show's host Monty Hall opens one of the two doors not selected to reveal a goat and then offers the contestant a chance to switch (1) from the door initially selected (but not yet opened) (2) to the other unopened door.

To elaborate on the detailed rules of the game, note the following. Monty knows (but does not divulge) the uniformly at random placed location of the car. If the contestant should initially happen to select the door with the car behind it, then Monty has two doors from which to uniformly at random choose when picking one to open. If the contestant should initially happen to pick one of the doors with a goat behind it, then the car is behind one of the other two doors and Monty is constrained to open the only remaining door which has a goat behind it. Finally, Monty does not open the contestant's initially selected door. It remains closed for the duration of the discussion. The question posed in the game is this: is it beneficial to accept Monty's offer to switch doors?

The Monty Hall Problem gained nationwide attention in the U.S. when it was discussed in 1990, in a popular, nationally syndicated newspaper column written by Marilyn vos Savant. For more details on this fascinating controversy, see [16, pages 42–45]. In her column, vos Savant stated correctly that the high-probability choice is to accept Monty Hall's offer and switch doors. The American reading public erupted in protest! The vast majority (92%) of 10,000 letter-writing-readers (including approximately 1,000 letters from PhD scientists and mathematicians, many writing on their university department stationary) strongly disagreed. In point of fact, most of the readers were wrong, and vos Savant was correct. Interestingly, the prominent mathematician Paul Erdös also rejected vos Savant's solution [16, page 45]. He later changed his mind only after witnessing a computer simulation of the problem in which the solution converged to vos Savant's answer.

The purpose of this example is not to provide yet-one-more contribution to the large volume of writing that has been previously directed at the Monty Hall Problem. It is, rather, to demonstrate how the tool set provides guidance in the analysis/synthesis of a problem that confounded a great many people (including many PhD scientists and a world-class mathematician). The tools can be used to (1) ask whether or not the opening of a door has changed the probability of the car being behind the door offered by Monty in trade for the one originally selected, and, if so, (2) in which direction (up or down).

I will not attempt to step through all possible combinations of contestant door choices, identity of the automobile-hiding-door, and choices of doors to be opened by Monty Hall, etc. The logic of the demonstration for one particular sequence of events translates easily to the other combinations. Therefore, without loss of generality, suppose that the doors are labeled A, B and C, that (1) the contestant initially chooses door A, and that (2) subsequently Monty opens door B or door C (call this $MontyB$ and $MontyC$, respectively) to reveal a goat. Let A, B and C denote respectively the events that the car is behind door A, door B or door C. According to the ground rules, the initial probabilities for the car's location are $p(A) = p(B) = p(C) = 1/3$. Under the conditions stated, there are four important conditional probabilities relating to the location of the car and the door which Monty opens for the contestant. They are:

$$p(MontyB|A) = p(MontyC|A) = 1/2, \tag{6.1}$$

since if the car is behind door A, Monty chooses one of the other two doors uniformly at random. Similarly, it is easily seen that

$$p(MontyB|B) = 0, \tag{6.2}$$

and

$$p(MontyB|C) = 1. \tag{6.3}$$

Again, assuming that the contestant has made an initial selection of door A, I will now formulate and answer three questions using (4.1) and (4.3) above.

Question 1. Does $p(C)$ remain constant following the contestant's initial selection of door A and Monty's opening of, say, door B? Using (4.3) above, this is question is written as $p(C|MontyB) = p(C)$? It can be rewritten equivalently as the following question,

$$p(MontyB|\neg C) = p(MontyB|C)?$$

The question is now seen to be, does the right hand side equal the left hand side? Note $p(MontyB|C) = 1$: if the car is behind door C and the player picks door A then Monty has no choice but to open door B. The value of $p(MontyB|\neg C)$ is 1/4. Given the car is not behind door C it is equally likely to be behind doors A and B. If it is behind door B (with probability 1/2), then Monty does not open door B. If it is behind door A (with probability 1/2) then Monty opens door B with probability 1/2, giving a left hand side of $(1/2)(1/2) = 1/4$. Since the right hand side does not equal the left hand side, equality does not hold and $p(C|MontyB)$ does not remain constant after Monty opens door B. Its value changes.

Question 2. Does $p(C|MontyB)$ increase in value? Using (4.1) above, this can be written as $p(C|MontyB) > p(C)$? The equivalently formulated question becomes: is $p(MontyB|\neg C) < p(MontyB|C)$? Using a calculation analogous to Question 1, one computes that $p(MontyB|\neg C) = 1/4$ while $p(MontyB|C) = 1$, so the inequality holds. Following the opening of door B by Monty, $p(C|MontyB)$ does actually increase in value.

Question 3. Does $p(A|MontyB)$ remain constant? That is, does the probability that the car is behind the originally selected door A remain constant given that Monty opens door B? Analogously, using (4.3) above, one finds that $p(MontyB|\neg A) = p(MontyB|A) = 1/2$ and equality holds. The probability of the car being behind door A has not changed.

To summarize, originally $p(C) = 1/3$, but opening a door increased its value to some unknown quantity. Originally, $p(A) = 1/3$ and opening a door has not changed its value. Since $p(C|MontyB)$ is the only remaining unknown, it can be concluded that the updated probabilities are $p(A|MontyB) = 1/3$ and $p(C|MontyB) = 2/3$. The high probability choice is to switch. This is the solution originally given by Marilyn vos Savant and the demonstration is complete.

Note that an interesting consequence of the above question/answer process is the realization that the tool set is capable of functioning as a probability calculation engine. Within a "small" finite system in which some numerical probability values are known, the equivalent formulations on the right sides of (4.1), (4.2), and (4.3) may be helpful in calculating some unknown probabilities.

The process demonstrated here can also be adapted to two other famous problems which are closely related to the Monty Hall Problem. They are the (1) Three Prisoners Problem and (2) Bertrand's Box Paradox.

7. Final Comments on the Utility of the Tool Set

It may be claimed by some readers that the usefulness of the tool set championed in this paper is diminished by the prior existence of other tools, techniques, and processes. For example, the process of Inference to the Best Explanation (IBE) certainly urges one to consider alternative explanations for the observed evidence. Similarly, consideration of $\neg h$ in the form of a null hypothesis in statistical hypothesis testing involving, say, Fisher's famous lady-tasting-tea thought experiment, has been routine since the mid 20^{th} century. Finally, one might implement a full, formal Bayesian analysis of a hypothesis-evidence relationship (using the odds-ratio form) requiring the estimation of prior numerical probabilities and likelihood ratios in order to arrive at estimates of a posterior probability.

These and other approaches exist, but I think that they are comparatively more elaborate or less straightforward and difficult to use when stood next to the simple, reason-guiding tool set of Section 4. The beauty of the tools lies in their ease-of-use. Rather than merely urging one to "consider alternative explanations", they offer an explicit formula which says "plug in the information and try to assess the relative magnitudes of the resulting conditional probabilities." Remember, in the absence of concrete, numerical probability information, use of the tools relies on comparative magnitude assessment (using an ordinal scale) rather than the computation of explicit probability values (on a ratio scale). In fact, the tool set requires no estimation of prior or numerical probabilities of any kind. It merely requires that the user input parameter information (e.g., "the person is guilty," "the person has confessed," "the abuser does not stop," etc.) and then let a formula produce conditional probability statements. The tools do not merely suggest, in a general and non-specific way, that one should "consider alternative explanations", but instead mandate that one should form $\neg h$ and see what effect it has on the comparative magnitudes of the resulting conditional probabilities. The tools reframe a problem. They introduce a new and different view of the problem which frequently produces insight.

The technical preparation required to use the tools is not large. Of course, one needs to understand the elements of probability theory and how to formulate and read conditional probability statements. However, I believe that, for example, typical university students can probably be "up and running" with test problems and conjectures using the tool set while other students are still struggling with the complexities of the word "statistical" involved in statistical hypothesis testing. Both the (1) ease-of-use and (2) power of the tools are perfectly mirrored in Matthews' original legal example. Most any person can probably understand (and, moreover, agree with) the unexpected legal conclusion, that a suspect from the not-guilty population might be highly suggestive and compliant, and therefore might be more likely to confess than a guilty suspect. Moreover, in the hands of subject matter experts in other fields of inquiry, use of the tools may prove even more powerful. Such experts would possess an understanding of additional underlying issues and their subtle relationships. This broader and deeper understanding might result in surprising insight when the domain specialists are led to contrast and compare the conditional probabilities created by use of the tools.

If readers are still unconvinced, then consider using the tool set as a screening device. Use it for preliminary evaluation of hypothesis-evidence relationships and then move on to more complex techniques depending on the outcome.

8. Conclusion

Motivated by the reading of the paper by Matthews, this paper has investigated the situation in which an individual evaluates a solitary hypothesis-evidence relationship. Further motivated by the reason-guiding analytical process demonstrated in Matthews' work, it has concentrated on a set of three simple analytical tools which are straightforward consequences of the historically standard definitions of Bayesian hypothesis confirmation. Simple does not mean trivial. Each member of the tool set involves a binary relation between two conditional probability statements, one of which always includes $\neg h$. The tool set is applicable to any candidate hypothesis-evidence relationship in which the equivalently formulated conditional probabilities are not ill-defined. This collection of candidate problems is large.

The tools are not difficult to apply. At this point in the paper, five major examples have been presented to illustrate their useful (practical) application.

They are the analysis and solution of the Monty Hall Problem, criticism of a questionable claim involving child sexual abusers, analysis of the NASA disaster, solving of Wason's card problem, and the Interrogator's Fallacy itself. The demonstrated ability of the tools to conform to the natural process of analysis and synthesis for these problems is sufficient to hypothesize their practical utility for solving a wide range of other problems. Even in situations where the tools may not lead to an explicit solution of a specific problem, the effort to apply them may reframe an issue and produce insight. In either case, they will have proven themselves to be of value. Their application would appear to constitute a formula-based, solution-guiding contribution to information assessment and problem solving which can help people to reason more effectively.

References

[1] Michael A. Bishop, M. and J. D. Trout, *Epistemology and the Psychology of Human Judgment*, Oxford University Press, New York NY, 2005.

[2] Emile Borel, (1924) "Á propos d'un traité de probabilities," *Revue Philosophique*, Volume **98**: pages 321–336. Tr. in H. E. Kyburg and H. E. Smokler, eds., *Studies in Subjective Probability*, John Wiley, New York, NY, 1964.

[3] Robyn M. Dawes, (1975) "The mind, the model and the task," Volume **1**, pages 119-129, in F. Restle, et al., eds., Cognitive Theory, Lawrence Erlbaum, Hillsdale NJ.

[4] Robyn M. Dawes, *Rational Choice in an Uncertain World*, Harcourt, New York NY, 1988.

[5] Robyn M. Dawes, *House of Cards: Psychology and Psychotherapy Built on Myth*, The Free Press, New York NY, 1996.

[6] Robyn M. Dawes, *Everyday Irrationality*, Westview Press, Oxford, 2001.

[7] Bruno de Finetti, (1937) "La prévision: ses lois logiques, ses sources subjectives," *Annales de l'Institut Henri Poincaré*, Volume **7**, pages 1–68. Tr. in H. E. Kyburg and H. E. Smokler, eds., *Studies in Subjective Probability*, John Wiley, New York NY, 1964.

[8] Gerd Gigerenzer and Ulrich Hoffrage, (1995) "How to improve Bayesian reasoning without instruction: Frequency formats," *Psychological Review*, Volume **102**, 4, pages 684–704.

[9] Reid Hastie and Robyn M. Dawes, *Rational Choice in an Uncertain World*, 2nd ed., SAGE Publications, Thousand Oaks CA, 2010.

[10] Colin Howson and Peter Urbach, *Scientific Reasoning: The Bayesian Approach, 3rd ed.*, Open Court Publishing Company, Chicago, 2006.

[11] Daniel Kahneman, Paul Slovic and Amos Tversky (eds.), *Judgments Under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge, 1982.

[12] Daniel Kahneman, *Thinking, Fast and Slow, Farrar, Straus and Giroux*, New York NY, 2011.

[13] Andre N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitrechnung*, J. Springer, Berlin, 1933. *Foundations of the Theory of Probability* (trans. N. Morrison), Chelsea Publishing, New York NY, 1950.

[14] Pierre-Simon Laplace, (1814) Essai Philosophique sur les Probabilités. Paris, 1814. Trans., A Philosophical Essay on Probabilities, John Wiley, New York NY, 1917.

[15] Robert A. J. Matthews, (1995) "The interrogator's fallacy," *Bulletin of the Institute of Mathematics and its Applications*, Volume **31**, pages 3–5.

[16] Leonard Mlodinow, *The Drunkard's Walk*, Pantheon Books, New York NY, 2008.

[17] Frank Plumpton Ramsey, "Truth and probability," *The Foundations of Mathematics and Other Logical Essays*, Harcourt Brace, New York NY, 1931.

[18] J. Edward Russo and Paul J. H. Schoemaker, *Winning Decisions*, Currency-Doubleday, New York NY, 2002.

[19] Peter C. Wason, (1968) "Reasoning about a rule," *Quarterly Journal of Experimental Psychology*, Volume **20**, pages 273–281.