On a Quantitative Standard for Assessing
Genuineness of Alleged UFO Contactees

J. W. Deardorff, PhD

Introduction

In a review of the book, *Majic Eyes Only* by Ryan Wood,[1] Michael Salla has pointed out the value of Wood's quantitative scheme for rating the authenticity of crash-retrieval cases.[2] Wood established seven categories of evidence: witnesses, sources, zingers, content, chronology, no-anachronisms, and forensics. He assigned different weightings to the different categories, and within each category assigned points ranging from 0 (least credible) to 5 (most credible). A summation of points for any one crash-retrieval case then provided a measure of its credibility relative to other cases, of which he analyzed 74. For consistency, he kept the weighting scale for the different categories constant for all cases. His scoring yielded 5 cases out of the 74 with the highest grade (100-125 points), 14 with a medium-high grade (75-100 points), 54 with medium level of authenticity (50-75 points), 1 with a medium-low level (25-50 points), and none in the lowest category.

Something similar might be done for assessing the genuineness of alleged UFO contactees. A quantitative method is needed, for example, by the Exopolitics Institute since its programs include investigating cases of UFO contactees backed up by credible evidence of voluntary physical contact and communications with ETs, with a goal of analyzing their credibility.[3] Which contactees can be

considered credible and which not? Mainstream ufology long ago elected to consider all alleged contactees as unworthy of investigation.[4,5] However, that position has become increasingly less tenable as more and more abductees proved credible and yet sometimes shared much in common with contactees. Both have received similar content in the telepathic communications from their aliens, especially with respect to future catastrophes mankind appears to be heading towards through warfare and over-utilization of Earth's resources. Some investigators have found that abductees as well as contactees sometimes receive benefits of healing and increased psychic awareness from their alien contactors.[6] This has led to the use of the term "UFO experiencers" to describe both categories.[7]

Consideration of Wood's system for alleged contactees

Ryan Wood presenting at the First Annual Crash Retrieval Conference - 2003

Wood's "witnesses" category (0 to 15 points) refers to first-hand witnesses and to factors entering into their credibility, such as their professions, state of mind, their number and relationships to others. This is directly applicable to our goal of contactee genuineness assessment (**CGA**). The category would refer to all those who have some first-hand observation that associates the alleged contactee with the alleged UFO aliens or ETs. If this category is extended to include the alleged contactee himself (or herself; interestingly, a large preponderance of alleged contactees are male) as one of these witnesses, we must keep in mind that his credibility as a contactee has to be determined primarily not by himself but by all else associated with the alleged contacts. However, his background, his character as described by others, and his willingness to tell his story or answer questions, and submit to a lie-detector or voice-stress analysis test, are all relevant. Here we shall name this category "*witness credibility*."

Wood's "sources" category (0 to 25 points) refers to how the case surfaced. This applies to the present application also – did it become known through natural means? Thus secondary witnesses, as well as first-hand ones and the alleged contactee himself, can be involved in this category. They may report a contactee-related event to a newspaper or a UFO data-collection center or ufologist

friend, out of which an article or book about the event or case may emerge. Possibly, information allegedly leaked by an undercover government agent may be involved. The number of such sources, their internal consistency and credibility then all enter in. Here we notice that for purposes of CGA, considerable overlap may occur between this *source* category and the *witness credibility* category, since UFO witnesses are the primary sources of UFO-related events associated with contactees. Unless redefined for CGA, the two categories are likely to be entangled.

In Wood's study, "Zingers" (0 to 20 points) refer to obscure, rare or subtle evidence or actions that strongly suggest genuineness rather than hoax. For CGA, the analog could include the question: Does the contactee or first-hand witness describe events that are unique rather than close repeats of what may be found in earlier UFO literature? This category does seem to carry over well to CGA, although again some overlap may occur. For example, a *zinger* may be observed by both the alleged contactee and one or more witnesses, and perhaps fall within the *witness credibility* or *source* category, so that care would be needed not to place the same event in more than one category.

For Wood, "content" (0 to 10 points) refers to all the informational data of the case – key data -- and its degree of verification. Since for CGA much of the key data would have fallen under the previous three categories, and much of the rest will fall under a "forensics" category still to be discussed, here one could only label this category as *miscellaneous.*

With "chronology" (0 to 10 points), Wood asks the question, "Is the timeline of the events and data consistent, believable and accurate?" This category can also apply to CGA, since a hoaxer with invented data can easily slip up and fail to have carefully researched the chronology of relevant historical facts into which to insert his invented incidents. However, for CGA this category may more often apply to the verification of alleged prophecies or scientific information the contactee may have received from his ET contactors.[8,9] Chronology comes into play here in that the time a prophecy was received by the contactee and made known to others, relative to the time the prophecy was fulfilled, is of paramount importance for its validation.

By "no anachronisms" (0 to 20 points) Wood refers to cases being more credible if they suffer only minimally from anachronisms. An anachronism occurs when a reported piece of information about

a case could only have happened at a date later than that supplied by the case itself, as determined from historical facts or records. A well confirmed and unexplained anachronism can point to the work of a hoaxer or disinformation agent. For CGA, staying alert for such data is therefore also very relevant, although it might fall under the previous *chronology* category.

With CGA, Wood's category of "forensics" (0 to 25 points) would take on the more general role of "physical evidence." Various events associated with a contact with an ET can leave evidence behind, whether these are due to the alien's UFO having been photographed, having left ground traces behind, or to the UFO alien having left artifacts behind for the contactee. Since UFO photographs and video images are of such great importance to ufology, this category would play a large role within CGA. However, care needs to be taken here also that an event is not placed both in a *physical evidence* category and in a *witness credibility* or *source* category.

Example of a multiple-category event for CGA. In the Billy Meier case, with which I am most familiar, in an instance in May of 2000 when he was outside his residence next to the garage with another -- Freddy Kropf, Meier suddenly looked up, searched the sky and told Kropf to look up and see the beamship of Florena, which Kropf soon noticed. It was a disk-shaped craft.[10] They then called to others around the residence (numbering 15) who saw it plus a second craft that had joined
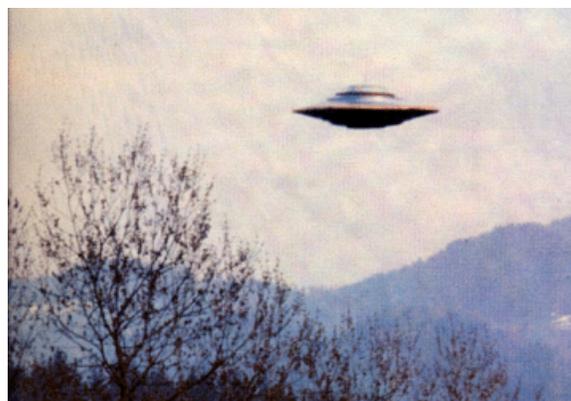
From one of the photos taken by Meier on March 8, 1975 near Schmidrüti, Switzerland

the first. Meanwhile Kropf retrieved his camera and took some photos of the two craft, which were sufficiently distant, however, that the images showed only as spots of light. Thus in this instance we have witness testimony to Meier apparently having received a telepathic communication with one of his contacting ETs located in one of their UFO craft overhead, as well as a bit of photographic evidence. The former could be categorized under *witness credibility,* or under *source*, since Lanzendorfer reported it (he was not a witness himself but knows Meier and most of the witnesses to this event quite well). Or, it could be listed under a category that would include indications of telepathy, which would fall under the heading of *miscellaneous* in the absence of an explicit

"telepathy" category. The taking of the photos would come under the *physical evidence* category. Hence one is led to ask if some general, consistent and flexible procedure exists for handling such a problem that would be free of category overlap and also avoid the use of an arbitrary point scale.

A probabilistic approach

Fortunately, just such a method was presented by Sturrock.[11] The ultimate application he had in mind was the UFO question: What is the overall assessed probability that UFOs are real (not man-made or natural objects)? Although the derivation of the approach involves Bayesian statistics, which can be daunting to non-mathematicians, the resulting equation that one can apply is quite simple. Sturrock presented the approach in detail, but did not attempt to answer the ultimate question in that paper!

Thomas Bayes set out his theory of probability in 1764

The procedure starts with collecting all the known information that bears upon a particular yes-no type question, such as: "Was this a sighting of a 'genuine' UFO or not?" or in our case, "Is this alleged contactee genuine or not"? The data are listed (in any order) as mutually exclusive events or pieces of information, rather than attempting to place them into distinct categories unless that is deemed convenient.

The next step, preferably accomplished by a small committee in a comprehensive effort, is to assign a probability, *p*, to each such piece of information, giving the committee's best estimate of the degree to which the particular datum by itself indicates the alleged contactee either is or is not a genuine contactee of a UFO/ET or UFO/alien.

A probability of $p = 0.9$, say, would express the opinion that the particular datum under scrutiny indicates strongly that the alleged contactee is genuine, while an assigned value of $p = 0.1$ would indicate the opposite just as strongly. Although some piece of data might by itself seem to be

absolute proof of genuineness or hoax, there are always some other possibilities, no matter how slim, that prevent assignment of the upper limit, 1.0 or lower limit, 0.0 for any individual value of *p*.

A probability of $p = 0.5$ would indicate complete uncertainty -- such a datum would lean as much in one direction as the other. Thus, if it is confirmed that the alleged contactee had witnessed a UFO, and that is all that is known about it, $p = 0.5$ could be assigned, since such an incident, by itself, doesn't suggest the witness either is or is not a contactee.

Assigning probabilities: An example

In the Meier-case incident described above, the data could be separated into three pieces, consisting of:

> (a) Meier yelling out to Kropf that he just received a telepathic message from Florena (an ET), and Kropf supporting this by noticing that Meier had suddenly looked up, briefly searched the sky and then spotted a (UFO) craft, which Kropf soon confirmed; and Lanzendorfer later reported in an article;
>
> (b) 15 others at this time visually witnessing the craft, noticing its disk shape, then noticing a second craft, with Lanzendorfer reporting this and the names of the witnesses; and
>
> (c) Kropf then having taken some photos of the craft, also mentioned by Lanzendorfer in his report.

For datum (a), one might assign $p = 0.9$ at the least, upon taking the background credibility of Meier, Kropf and Lanzendorfer into consideration. In perhaps not estimating a higher value for *p*, one allows the possibility that what were seen were high-flying balloons, with Meier and Kropf collaborating in a hoax to be sprung on an occasion when such balloons were overhead, to make it seem as if Meier had received a telepathic message. Or perhaps these were UFOs that had nothing to do with Meier's alleged contactee experiences and it was only a coincidence that Meier spotted them when he did. Another (extra slim) possibility is that Lanzendorfer invented the whole incident, without any of the named persons objecting to such a lie. These and perhaps similar considerations prevent any assignment of $p = 1$.

Concerning datum (b), we may treat it as mutually exclusive of (a) and (c) because these witnesses weren't at first in the immediate vicinity of Meier and Kropf to be able to notice the apparent telepathy that had taken place. Instead, their witness to the craft is relevant for CGA because the larger craft, at least, was reported as having a disk shape, and this is connected to Meier as contactee through most of his multiply photographed UFOs, one of which was witnessed by several persons, having this general shape, and because the sighting occurred high in the sky but directly over Meier's residence at a time when he was present. Certainly, if they had looked for the craft but not spotted them, after having been informed of the event, which lasted a half hour, this would count against the genuineness of the event. But since they did confirm the craft's presence and the disk shape, this one incident has to weigh in favor of Meier's genuineness as contactee. After consideration of the rather small fraction of overall UFO sightings that report the disk shape, and considering witness credibility, one might assign $p = 0.8$ to datum (b).

In assigning a probability to (b), one doesn't explicitly take the number of witnesses who saw the same thing into account. However, the say-so of several such witnesses is almost always of more value than that from only one, so that one assigns a higher probability for genuineness (a value of p closer to unity) for multiple witnesses than for a single witness, if there is no reason to doubt their credibility. Their individual witnessings do not each receive separate p values, since in observing the same thing their data were not mutually exclusive.

Salem Massachusetts 1952, Coast Guard

For datum (c), the photos confirmed that two UFOs were present then, but they only showed up as tiny orbs. Hence one might assign $p = 0.5$ here, since it might be assumed they were balloons, not UFOs/beamships.

If there is uncertainty for two or three pieces of information whether or not they are mutually exclusive, it is best to lump them together and

treat them as a single piece of data. In so doing, one may find it logical to assign a *p* value to the lumped data that indicates a higher degree of certainty one way or the other (lie further from 0.5) than what would be assigned to any of the individual *p* values of the data before lumping them.

It can be seen that much discussion could result within a committee in coming to agreement on which data constitute mutually exclusive evidence, and what value of *p* to assign to each mutually exclusive piece of evidence. However, a similar problem would hold for Wood's study, if a committee had been involved.

This same procedure is to be followed for each mutually exclusive piece of data bearing on the question: Is the alleged contactee genuine or not? This might involve only 5 or 10 incidents or items, or, as in the Meier case, many hundreds. One keeps track of the probabilities assigned, as p_1, p_2, p_3, and so on. Upon going over all the rest of the evidence, how does one accumulate the resulting *p*-values involved to obtain an *overall* probability, P, that the contactee is genuine?

Accumulating the individual probability estimates

This is where the formula resulting from Sturrock's analysis enters in. In simplest form, it is:

P = A / (A + B), where   (Eqn. 1)

$A = p_1$ x p_2 x p_3 x p_4…. and $B = (1-p_1)$ x $(1-p_2)$ x $(1-p_3)$ x $(1-p_4)$….

The slash indicates division, the "x" multiplication. The dots indicate that all further estimated probabilities are to be included also within the chains of multiplication. Three derivations of Eqn. 1, one from heuristic combinatorial arguments and the other two using Bayesian statistics, are presented by Deardorff.[12]

If Eqn. 1 is applied to the above little example of items (a), (b) and (c), one finds P = 0.97. Item (c) does not affect this outcome, as it shouldn't for an estimate of *p* = 0.5 indicating complete uncertainty either for or against. Item (b) with *p* = 0.8 enhances the probability of genuineness beyond that of the single item, (a) with *p* = 0.9. This may seem peculiar at first until one realizes

that it accords with common sense: the more mutually independent pieces of information that favor a particular outcome, the more probable that outcome is correct. Or, "the more evidence, the better." Hence the *p* values do not get averaged together. If, however, item (c) had been 0.2 rather than 0.5, P would have been 0.9; the 20% chance of genuineness would have cancelled the 80% chance, leaving just 90%.

In the derivation of Eqn. 1, a factor known as the "prior probability" enters in, if there is an initial bias not taken into account in the rest of the analysis. I believe it preferable to take the additional evidence associated with any such bias explicitly into account within the calculation of A and B of Eqn. 1, so that the "prior probability" can be set to 0.5. I.e., if everything known about the alleged contactee is included within the explicit input data, including whatever may have provoked an initial bias for or against the alleged contactee, there should be no hidden bias left, and one starts with a clean slate of no bias. In the simple expression of Eqn. 1, it has been assumed that this is the procedure that will be followed. Hence, everything known about a contactee case should be taken into account for this Bayesian approach to be most valid. In practice, everything is not known, but a reasonable result may be expected if all that is known is utilized, provided the number of *p* values is not too small..

Due to the commutative nature of Eqn. 1, the *p* values may be grouped, if desired, into whatever categories seem most suitable for that case, with Eqn. 1 used for the data within each category; then those resulting probabilities for genuineness can be accumulated into the final value, P, again using Eqn. 1. This final value will not depend upon how or whether the input data was grouped or categorized, but only upon the total number of estimated *p* values and the probability assigned to each one. Thus, one may utilize various different categories, as suggested earlier, such as *witness credibility, source, zingers, chronology, no anachronisms* and *miscellaneous*, or not use them, provided that within each and all categories each piece of evidence is mutually exclusive from all the rest of the data. None of the data is totally independent, since all of it relates to the same yes-no question.

A key advantage of the present approach is that it is derivable mathematically or through proper statistical analysis, and can be universally understood wherever the meaning is understood of what a

probability, lying between zero and unity, represents. It is to be kept in mind, however, that the final answer, P, depends entirely upon the individual probability (*p*) estimates, which are subjective.

A nice feature of the method is that the various pieces of evidence used as input can be taken in whatever order one desires, without affecting the value of P. And as noted, one may utilize categories if desired, or dispense with them entirely and just place the input data in chronological order.

Another convenient feature of this CGA approach is that pieces of evidence suggesting hoax (*p* < 0.5) are to be input, in any desired order, right along with the rest of the data, and can count as strongly against genuineness as other evidence counts in favor. With the point system as used by Wood, however, a definitive piece of evidence of hoax would only give zero points within some category, and may not allow sufficient offset of positive points from other categories. Similarly, if in a given case the evidence of a particular type or within a particular category is overwhelmingly powerful, its value will not be unduly limited by an arbitrary cap to that category.

**Utilizing the final probability estimate, P**

In a trial by jury, a final guilty or innocent pronouncement is based upon a subjective criterion, such as "innocent unless proven guilty beyond a reasonable doubt." For our application, the final step is also subjective, as was also the case in Wood's study. We must agree upon some standard probabilistic level of genuineness as acceptable for our purposes. For example, P = 0.95 might be an acceptable level of genuineness; or it might be considered that P = 0.999 would be a better criterion. It depends upon how damaging it is considered to be to the purpose at hand if a false contactee is treated seriously.

An alleged contactee with a level of P judged to lie between 0.5 and the higher criterion level might be considered tentatively untrustworthy but worthy of further evaluation if new evidence came along. For P below 0.5 the possibility of hoax seems much too strong to allow acceptance of the alleged contactee as genuine.

In summary, the main advantage of this probabilistic approach is that a meaningful assessment of overall genuineness is provided by the outcome, as opposed to an approach that utilizes points on an arbitrary scale within an arbitrary set of categories.

***

About the author: James W. Deardorff is the author of three books – *Celestial Teachings*, *The Problems of New Testament Gospel Origins*, and *Jesus in India*, and creator of a website that researches the *Talmud of Jmmanuel* (http://www.tjresearch.info). His professional career was in the atmospheric sciences, first at the National Center for Atmospheric Research in Boulder, CO, then in the Dept. of Atmospheric Sciences at Oregon State University. His curriculum vitae is available at:

http://www.tjresearch.info/resume.htm.

***
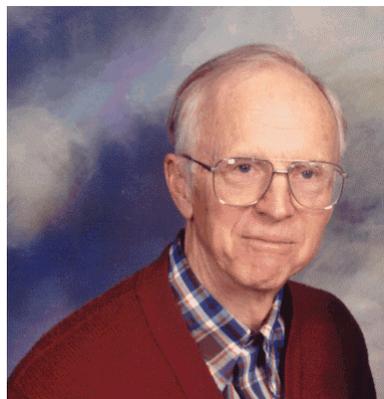
References

[1] Wood, Ryan (2005). *Majic Eyes Only: Earth's Encounters with Extraterrestrial Technology* (Broomfield, CO: Wood Enterprises).

[2] Salla, Michael (2006). "Review of Ryan S. Wood, Majic Eyes Only," *Exopolitics Journal*, vol. 1, no. 2 (January 2006), 160-63. Available at: http://www.exopoliticsinstitute.org/Journal-home.htm .

[3] See http://www.exopoliticsinstitute.org/Programs.htm .

[4] Hynek, A. J. (1972). *The UFO Experience: A Scientific Enquiry* (New York: Ballantine Books), 31-34.

[5] Bryan, C.D.B. (1995). *Close Encounters of the Fourth Kind* (New York: Alfred A. Knopf), 8-9.

[6] Mack, John (1994). *Abduction: Human Encounters with Aliens* (New York: Charles Scribner's Sons), 398-99.

[7] Sprinkle, Leo (1999). *Soul Samples* (Columbus, NC: Granite Publishing).

[8] See: http://www.theyfly.com/prophecies/prophecies.htm .

[9] See: http://www.tjresearch.info/denial.htm, under item (i).

[10] Lanzendorfer, Hans Georg (2001). "Anatomie liederlicher Recherchen," in FIGU Bull. #32, 5; available at: http://www.lanzendorfer.ch/Artikel_Daten/bulletin_27d.htm .

[11] Sturrock, Peter (1994). "Applied scientific inference," *J. Scientific Exploration* (Vol. 8, No. 4), 491-508.

[12] See: http://www.tjresearch.info/cumulate.htm .