

**Can absence of evidence be treated as evidence of absence?
An analysis of Stenger's argument against the existence of God**

Chong Ho Yu

Azusa Pacific University, USA

chonghoyu@gmail.com

<http://www.creative-wisdom.com/pub/pub.html>

The Chinese version of this article was published in the *China Graduate School of Theology Journal* (2012, Issue 52, pp.133-152). The Chinese version is slightly longer than the English version due to a request for revision from the reviewer.

Abstract

Victor Stenger, American philosopher and physicist, is one of the prominent authors of the New Atheism Movement. His central argument against the existence of God is: absence of evidence can be treated as evidence of absence. In other words, when we have no evidence or other reason for believing in some entity, then we can be sure that this entity does not exist. In Stenger's view, if God is the creator and the master of the universe, then he must leave some footprints that are detectable by scientific tests. Therefore, the existence of God is a scientific hypothesis that can be tested by standard methods of science. Stenger asserts that this hypothesis fails the tests. The objective of this article is to analyze his thesis in the perspectives of the history of science and statistics. The discussion will be divided into six domains: 1. Contraposition argument, 2. Problem of testing, 3. Bayesian probability, 4. meta-analysis, 5. data mining, and 6. hypothesis testing. The author concludes that Stenger does not succeed in validating the notion that "absence of evidence can be treated as evidence of absence" and thus consequently he also failed to disprove the existence of God via this argument.

American philosopher and physicist Victor Stenger, one of the prominent authors and spokesmen of the New Atheism Movement, has written numerous books and delivered countless speeches to advocate the course of atheism throughout three decades. His ideas can be epitomized to this central argument: absence of evidence could become evidence of absence (In the following discussion this notion will be abbreviated as AE=EA). The objective of this article is to help theologians and people who are interested in apologetics examine Stenger's thesis in the perspective of history of science and statistics. Because some readers may not be familiar with history of science and statistics, the author tries not to make this article overly technical.

Simply put, Stenger's argument can be summarized as follows: When we have no evidence or a sound reason for believing in some entity, then we are confident that this alleged entity does not exist at all. We have no evidence for Bigfoot, the Abominable Snowman, and the Loch Ness Monster, and therefore we do not believe they exist. By the same token, the Christian God should be examined in the same scientific fashion. Stenger proposes that the existence of Abrahamic God is a scientific hypothesis that can be tested by standard methods of science. If God is the creator and the master of the universe

and he actively interferes with human affairs, then he must leave some footprints that are detectable by scientific tests. But no such evidence is found, thus Stenger concludes that the hypothesis fails the tests (Stenger, 2007, Kindle Locations 115-117).

Contra-positive argument

On some occasions, absence of evidence could be counted as evidence of absence. One of those situations is the contra-positive argument. Specifically, we can utilize the transposition rule of inference found in the classical logic to make the following conclusion: C implies E; Not-E ($\sim E$) must also imply Not-C ($\sim C$). If the cause always leads to the effect, then absence of the expected effect is evidence of absence of the cause. For example, if it is raining outside, then the streets must be wet. If the streets are not wet, then it is not raining outside. The required condition is: the cause always leads to the same effect.

However, the debate about God's existence is not so simplistic. We are not absolutely certain that if God exists what effect would be observed. One of the major arguments against intelligent design and Paley's watch metaphor is that we do not know how the designer, if there is any, created this universe. Sober (2000, 46-57) argued that if we know that an intelligent designer would make the universe in certain ways, then we could argue that since the world appears in this expected order, it is likely that this world was created by an intelligent designer. However, creationism starts from the observed phenomenon of the existing world, which appears to be well-structured. Given the existing world, creationists argue that the world originates from an intelligent designer. Sober insists that it is a form of logical fallacy. In other words, we cannot infer from E to C because we don't know what C looks like; how God would design the world.

However, new atheists infer from $\sim E$ to $\sim C$ as if they know how C operates and what type of E should be observed. For example, Stenger stated (2007) that if God is an intelligent designer, then we must observe a well-structured human body (If C then E). Because our body human body has many signs of poor engineering, then there must be no intelligent designer (If $\sim E$, then $\sim C$). In other words, the absence of well-structured body is considered the evidence of no intelligent designer. Stenger (2007) wrote:

Our bones lose minerals after age thirty, making them susceptible to fracture and osteoporosis. Our rib cage does not fully enclose and protect most internal organs. Our muscles atrophy. Our leg veins become enlarged and twisted, leading to varicose veins. Our joints wear out as their lubricants thin. Our retinas are prone to detachment. The male prostate enlarges, squeezing and obstructing urine flow (Kindle Locations 590-592).

The previous passage assumes that an intelligent designer should make us not to lose bone calcium after age thirty, and should make our rib fully enclose most internal organs...etc. However, how could Stenger know what an intelligent designer should do? Should he make us lose bone calcium after age forty? Age fifty? Age sixty? Age seventy? Even if God makes us not to lose bone calcium until age eighty, I suspect Stenger might still ask: why not age ninety? But how could we know whether this is the

intention of God to make us start aging rapidly after age thirty? The argument made by Sober against creationism is applicable to Stenger's criticism. We cannot infer from E to C because we don't know how God would design humans. In brief, contra-positive argument could support $AE=EA$ if we know how C leads to E and C always result in E. But this is not so simplistic in the debate between theism and atheism.

Detection problem

In the previous discussion, the example is highly simplistic. We can simply use our naked eyes to observe whether the ground is wet or not. But to address a complicated phenomenon that cannot be directly detected by our natural senses, we need powerful instrument. $AE = EA$ is valid if we have sophisticated instrument to accurately detect the absence or presence of the evidence. For example, I suspect that I have a brain tumor and thus I went through a thorough physical examination. The doctor used PET scan, fMRI, X-ray, and other cutting-edge equipment to scan my brain, but did not find any sign of brain cancer. Then the doctor can confidently conclude that the absence of the evidence of brain tumor is the evidence of the absence of brain cancer. But there are many auxiliary assumptions:

1. The search space is finite and can be exhausted.
2. The equipment used for the detection is the right type of equipment and is well-calibrated for accuracy.

If either one of the preceding conditions fails, then it is a fallacy to claim that $AE = EA$.

Stenger (2009) used a very simplistic example to downplay the complexity of the preceding detection problem. He argues:

You often hear theists and even some reputable scientists say, "Absence of evidence is not evidence of absence." I dispute this. Absence of evidence can be evidence of absence beyond a reasonable doubt when the evidence should be there and is not found. For example, no one has seen elephants in Rocky Mountain National Park. But surely, if elephants did roam the park we should have evidence for them: footprints, droppings, smashed grass. While a remote possibility exists that they have remained hidden all this time, we can conclude beyond a reasonable doubt that Rocky Mountain National Park is not inhabited by any elephants. In this manner, the absence of evidence for the Judeo-Christian-Islamic God where there should be clear evidence allows us to conclude, again beyond a reasonable doubt, that such a God does not exist (p.241).

Relatively speaking, the area of Rocky Mountain National Park is small enough for scientists to utilize satellites and standard equipment to detect the signs of elephants. But what would happen if we are looking for outer-space alien civilizations in the universe? This project is called *Search for extraterrestrial intelligence* (SETI), which was launched by the University of California, Berkeley in 1979 and was subsequently funded by the U.S. government in 1992. Even though scientists have not found any sign of the existence of outer-space aliens, no one could conclusively declare that the absence of

evidence of aliens is the evidence of no aliens. First, the scope (the entire universe) is too huge. Consequently, we cannot make a firm conclusion unless we examine every corner of the universe. Second, we don't even know whether we have been using the right equipment or not. Many SETI efforts were devoted to detect radio signals or other electromagnetic signals. But how could we know technologically sophisticated aliens emit radio signals like us? How could we know they are not using laser or other types of technology instead? If the aliens use advanced technologies that are way beyond our imagination, we might never detect their existence at all (Davies, 2010).

In the history of science, there are cases that absence of evidence is mistakenly treated as evidence of absence because the improper detection method and instrument were used. When William Atherstone announced that he found a 21-carat diamond in South Africa in 1867, no one believed him because since the fourth century India had been the only source of diamonds. Diamonds in the raw form are buried at great depths inside the earth. When a volcano erupts, diamonds are thrown out of the top of the volcano along with molten rock, and therefore the best place to find diamonds is in the center of an extinct volcano. However, there are no volcanoes on the mainland of South Africa. In 1868 England sent one of the best mineralogists, James Gregory, to South Africa for further investigation. After examining many rock samples, Professor Gregory concluded that there were no diamonds in the whole of South Africa. You may think that this mistake is laughable because today indeed there are many diamond mines in South Africa, but you must realize that Professor Gregory had used the best scientific apparatus accessible to him at his time (Nigel, 1980; Morton, 1877).

SETI scientists assume that aliens use radio signals. Gregory assumed that diamonds must be present near volcanoes. The above examples teach us this valuable lesson: preconceived misconception might lead us to use the wrong detection method. Seeking for God is far more complicated than looking for aliens and diamonds. How could we exhaust all search space? How can we ensure that we have the right detection method? Interestingly enough, when Stenger formulated his argument against the Anthropic principle, he seemed to realize the limitation of our detection. According to the Anthropic principle, the universe seems to be so fine-tuned that it allows carbon-based life form to inhabit, and thus it implies that the universe has an intelligent designer. Stenger (2007) argues:

We expect any life found in our universe to be carbon-based, or at least based on heavy element chemistry. But that need not be true in every conceivable universe. Even if all the forms of life discovered in our universe turn out to be of the same basic structure, it does not follow that life is impossible under any other arrangement of physical laws and constants (Kindle Locations 1437-1440).

In a more recent book, again Stenger (2011) argued that theistic supporters of fine-tuning assume that any form of life must necessarily resemble our own, such as needing oxygen for water, which is a universal solvent. While it is true for our form of life, the carbon-based life form is not a large sample at all. He asserted that other life forms may someday be found elsewhere in the universe: "I am sure it's out there.....I will wager that extraterrestrial life, sufficiently distant to not be connected to Earth life, will not be based on left-handed DNA. So why make that a requirement for life in this and

every universe? I expect life to occur in any sufficient system in any sufficiently long time, and our kind of biology is not a constraint" (Kindle Locations 2207-2221).

To summarize, Stenger seems to deduce that absence of evidence of non-carbon-based life form is not the evidence of non-existence of non-carbon-based life form, because: 1. The search space is very huge. They might exist in the far side of the universe. 2. We cannot assume that life forms in other galaxies are like us. In short, by Stenger's standard, AE is not EA in the case of looking for non-carbon-based life form! Could the same principle be applied to searching for evidence of the existence of God?

Bayesian probability

Based on the alleged absence of evidence of God's existence, Stenger contended that the probability of God's existence is extremely small. This assertion is a response to British physicist Stephen Unwin (2003). Using Bayesian probability, Unwin computed that the probability of God's existence is .67.

Bayesian probability was invented by 18th century mathematician and theologian Thomas Bayes. To evaluate the probability of a hypothesis, first the Bayesian statistician specifies some prior probability, and then it is updated in the light of new relevant data. Bayesian is a form of subjective probability in the sense that it views probability as a degree of belief, which is in sharp contrast to the frequency approach to probability. In the frequency school, probability is empirical and objective. For example, if you want to know the probability of obtaining a head from flipping a coin, you can simply flip the coins many times and then observe the outcomes. However, many events in the world are not repeatable or allow us to conduct experiments. The debate about the existence of God is an example. To tackle the problem, Unwin started with his estimation of the prior probability of God's existence. He set the prior probability to .5. The rationale is: if we do not have any information about the subject matter, then we can assume that each side has 50% chance. For example, if we do not know anything about Obama and Romney, then we can assume that the probability of either one of them would win the 2012 Presidential election is .5. After setting the prior as .5, Unwin inserted evidence for and against God's existence into the Bayesian equation by evaluating six pieces of evidence and assigning a score called "divine indicator" (D) to each piece of evidence. At each step he calculated the posterior probability after each D is taken into account:

- 1) The evidence for goodness, such as altruism: $D = 10 \Rightarrow p = 0.91$.
- 2) The evidence for moral evil done by humans: $D = 0.5 \Rightarrow p = 0.83$.
- 3) The evidence for natural evil (natural disasters): $D = 0.1 \Rightarrow P = 0.33$.
- 4) The evidence for "intra-natural" miracles (successful prayers, etc.): $D = 2 \Rightarrow P = 0.5$.
- 5) The evidence for "extra-natural" miracles (direct intervention by God in nature): $D = 1 \Rightarrow P = 0.5$.

6) The evidence for religious experience (feeling of awe, etc.): $D = 2 \Rightarrow P = 0.67$.

Stenger (2011) pointed out that Tufts University physicist Larry Ford examined Unwin's calculation and made his own estimate using the same formula. Ford's result is 10^{-17} (Kindle Locations 2948-2950). Stenger sided with Ford by saying that Ford's result is much more plausible. First, Stenger used his famous argument $AE = EA$ to object the prior probability of God's existence as .5: "the lack of any evidence or other reason to believe some entity such as Bigfoot or the Loch Ness Monster exists implies that it is highly unlikely that it does. So the prior probability of God should be more like one in a million or less. $P = 10^{-6}$ (Kindle Locations 2970-2972).

Again, based on the logic of $AE = EA$, Stenger further argued that if God exists, he should be producing miracles. He observed that there are no miracles, and therefore D based on the absence of evidence for miracles is < 1 . He criticized that Unwin made the typical theistic fallacy: goodness can only come from God, and thus Unwin mistakenly assigns an unreasonably high divine indicator, $D = 10$. In Stenger's and Ford's view D should be adjusted to 0.1 only. Further, Ford noted that the existence of both moral and natural evils in the world is evidence against God's existence. Stenger said that Ford's values of $D = 0.01$ and $D = 0.001$ for moral and natural evils, respectively, are far more reasonable than Unwin's estimates. In short, all estimations made by Ford and Stenger are much lower than Unwin. And therefore their probability of God's existence is extremely small (Kindle Locations 2974-2983).

As mentioned before, Bayesian is a type of subjective probability measuring the degree of belief. It is not surprising to see that different people have different estimates. However, it is surprising to see that as a scientist Stenger jumped into a hasty conclusion without using more rigorous statistical methods. When the numbers are involved with subjective ratings, usually statistics compute the inter-rater reliability, such as the Kappa coefficient or intra-class correlation coefficient. In this debate, obviously, the inter-rater reliability is very low. The solutions to this situation are well-developed among Bayesians. In this case, there should be a panel consisting of several judges, such as atheists, agnostics, and people from different religious backgrounds. Assume that there are five judges in the panel and all of them give different numbers. The possible solutions are:

1. Trimmed average: The panel could use the trimmed mean by dropping two of the extremes. Take Olympic Games as an example. Some sports do not have clear winners, such as gymnastics, and as a counter-balance usually Olympics arrange seven judges. Needless to say, every judge might have some bias. An Asian judge might give Asian athletics a higher score; a European judge might favor European athletics. To ensure the outcome is fair, the highest and the lowest scores are dropped. Then the remaining five scores are averaged.

2. Sensitivity analysis: it is the standard way of approaching the problem of diverging estimates. The procedure is: Run the analysis five different times. Each time use a different prior probability distribution. Then compare the different posterior probability distributions. Formally speaking, sensitive analysis aims to quantify the impact of parametric variation on model output. The purpose of doing sensitive analysis is to answer the following questions: How confident is the researcher in the results?

How much the results change if the data are different? If the results are vastly different, it is clear that probability estimate is sensitive to the choice of the input data, then the researcher must investigate where the differences are (Felli & Hazen, 1999; Levy, personal communication; Skene, Shaw, & Lee, 1986).

3. Model averaging: In this method, the researcher creates a prior distribution that is a mixture of the different people beliefs (Levy, personal communication). For example, Unwin has an estimate of prior probability as .5. Other four people might have estimates of .4, .3, .2, and .1. The statistician could construct a distribution as a mixture by weighing each equally. The probability for the parameter is given by $(1/5)(.1) + (1/5)(.2) + (1/5)(.3) + (1/5)(.4) + (1/5)(.5)$ (Levy, 2011). Hoeting, Madigan, Raftery, and Volinsky (1999) warn that many data analysts construct a single model but ignore model uncertainty. However, this is exactly what Ford and Stenger did. Hoeting *et al.* state that model averaging provides a coherent mechanism for accounting for model uncertainty. In addition, as mentioned before, the frequency approach of probability views probability as objective and empirical. Thus, frequentist hypothesis testing offers no method for resolving conflicting findings across alternative models, but Bayesians can (Montgomery & Nyhan, 2010).

In summary, Stenger's use of Bayesian probability is questionable because he did not apply at least one of the above methods to resolve conflicting estimates of probability from different people. Once again, the notion of AE = EA is not substantiated.

Meta-analysis

Bayesian probability is one of many incorrect statistical applications made by Stenger. He also made many other incorrect assertions about statistical methods to support AE = EA. For instance, Stenger objected using meta-analysis to analyze the evidence for supernatural or God. He said:

This procedure is highly questionable. I am unaware of any extraordinary discovery in all of science that was made using meta-analysis. If several, independent experiments do not find significant evidence for a phenomenon, we surely cannot expect a purely mathematical manipulation of the combined data to suddenly produce a major discovery. No doubt parapsychologists and their supporters will dispute my conclusions. But they cannot deny the fact that after one hundred and fifty years of attempting to verify a phenomenon, they have failed to provide any evidence that the phenomenon exists that has caught the attention of the bulk of the scientific community. We safely conclude that, after all this effort, the phenomenon very likely does not exist (2007, Kindle Locations 824-830).

Different studies on the same problem might produce the same result or different results. To get an overall vote count, a meta-analysis combines the results of diverse studies by measuring the effect size (Glass, 1976; Glass, McGraw, & Smith, 1981; Hunter & Schmidt, 1990). Simply put, Stenger said that if several studies show $\sim X$, we should not expect that pulling these studies together would show X. In response to the Stenger's statement, British statistician David Bartholomew (2011) criticized that the words "mathematical manipulation" and "suddenly" are very vague. First, manipulation in

meta-analysis is no more mathematical than other statistical procedures, such as hypothesis testing. Second, there is nothing “sudden” about their emergence. This author sides with Bartholomew.

Stenger does not seem to understand the value of meta-analysis. Indeed, meta-analysis is a standard way of rectifying many challenging situations in statistics. Not only can Meta-analysis be used for synthesizing results of past research, but also for new research studies. For example, Baker and Dwyer (2000) conducted eight studies regarding visualization as an instructional variable (sample size=2000). If all subjects are used for one analysis, the statistical power is too big, and as a result, any trivial effect might be misidentified as a significant effect due to the large sample size. Instead, the effect size should be computed in each study individually. Then the findings of eight studies are pooled to draw inferences as a collective body of research.

Besides the risk of overly high statistical power, using all data in one test may make the researcher overlook the Simpson's Paradox. It is a phenomenon that the conclusion drawn from the aggregate data is opposite to the conclusion drawn from the contingency table based upon the same data. For example, a university conducted a study to examine whether there is a sex bias in admission. The admission data of the MBA program and the law school were analyzed. By looking at the MBA data only, it seems that females are admitted at a slightly higher rate than males in the MBA program. The same pattern can be found in the law school data. Interestingly enough, when the two data sets are pooled, females seem to be admitted at a lower rate than males! In England once a 20-year follow-up study was conducted to examine the survival rate and death rate of smokers and non-smokers. The result implied a significant positive effect of smoking because only 24% of smokers died compared to 31% of non-smokers. However, when the data were broken down by age group in a contingency table, it was found that there were more senior people in the non-smoker group. In short, the conclusion based on partitioned data is different than that of aggregated data (Appleton & French, 1996). To address the Simpson Paradox, Olkin (2000) recommends researchers to employ meta-analysis rather than pooling. In pooling, data sets are first combined and then the groups are compared. In meta-analysis, groups in different data sets are compared first and then the comparisons are combined.

To summarize, it is very common that the conclusion from separate analyses and the conclusion of a single analysis using all data could be different. Meta-analysis has been recognized as a standard way to deal with this situation. In meta-analysis, the information of partitioned datasets is given consideration first. It is not to “manipulate” several studies to produce the result that could not be found in a single study. Stenger’s denial of value of meta-analysis cannot establish $AE = EA$.

Data mining

The reason that Stenger rejected meta-analysis and other research methods was because in his view hypothesis testing is the only rigorous scientific method. In addition to rejecting meta-analysis, he also criticized data mining. According to Stenger, data mining is a cover-up of the absence of evidence. He wrote (2007):

The hypotheses being tested must be established clearly and explicitly before data taking begins, and not changed midway through the process or after looking at the data. In particular, "data mining" in which hypotheses are later changed to agree with some interesting but unanticipated results showing up in the data is unacceptable (Kindle Locations 158-160)They (theists) are often naturally reluctant to accept the negative results that more typically characterize much of research. Investigators may then revert to data mining, continuing to look until they convince themselves they have found what they were looking for. (Kindle Locations 164-166).

Bluntly speaking, Stenger has an incorrect concept of data mining. It is certain that most statisticians would disagree that data mining is to continue looking until finding what we expect to look for. Actually, data mining is a cluster of techniques that has been employed in the field Business Intelligence (BI) for many years (Han & Kamber, 2006). According to Larose (2005), data mining is the process of automatically extracting useful information and relationships from immense quantities of data. Data mining does not start from a strong pre-conception, a specific question, or a narrow hypothesis, rather it aims to detect patterns that are already present in the data and these patterns should be considered relevant to the data miner. Thus, data mining is viewed as an extension of Exploratory Data Analysis (EDA) (Luan, 2002; Yu, 2010). Stenger is correct that in data mining a new hypothesis might be generated, but this is not a problem at all. Rather, this is the characteristic of exploratory data analysis (Behrens & Yu, 2003)! Indeed, data mining have been widely applied in academic research projects that can meet rigorous standards in peer-review journals (Yu, Jannasch-Pennell, DiGangi, Kim, & Andrews, 2007; Yu, C. H., DiGangi, Jannasch-Pennell, & Kaprolet, 2008, 2010).

Hypothesis testing

While rejecting meta-analysis and data mining, Stenger failed to build a strong case of setting hypothesis testing as the highest standard because he has misconception of this methodology. According to Stenger (2007), if the probability or the p value is .05, it implies that "if the experiment were repeated many times in exactly the same fashion, on average one in twenty would produce the same effect, or a greater one, as an artifact of the normal statistical fluctuations that occur in any measurement dealing with finite data. But think of what that means. In every twenty claims that are reported in medical journals, on the average one such report is false-a statistical artifact!" (Kindle location 805-806). Bartholomew (2011) bluntly said that "Stenger makes the elementary mistakes" (p. 125). Bartholomew corrected Stenger by saying "it would be true to say that in every case where there was actually no difference, an experiment will be falsely reported as showing a difference one time in every twenty" (p.125). In other words, the alpha level of .05 is about the Type I error rate. And this basic information is taught in introductory statistics.

Even if Stenger's conceptualization of hypothesis testing is correct, his attempt of using hypothesis testing to support $AE = EA$ cannot succeed. The logic of hypothesis testing is: Given that the null hypothesis is true, what is the probability of observing the data in the long run? It is not addressing this question: Given the data, what is the probability that the hypothesis is true in the long run?

Hypothesis testing is concerned with whether we can confirm that the data and the hypothesis fit each other. In philosophy of science's term, we are testing empirical adequacy (Van Fraassen, 1980). If the data cannot fit the model, it does not necessarily mean that the model is not true. Following this line of reasoning, in the article entitled "Absence of evidence is not evidence of absence" Altman and Bland (1995) gave this warning to medical researchers: "Randomized controlled clinical trials that do not show a significant difference between the treatments being compared are often called 'negative.' This term wrongly implies that the study has shown that there is no difference" (p.485).

Conclusion

Based on the preceding analysis, it is clear that contraposition argument cannot establish $AE = EA$ in the debate of the existence of God. Contraposition argument can be applied in simplistic cases. However, we do not know how God created the universe, how he designed human body, and what footprints in the world would be left by God. The detection problem indicates that the search space of looking for God is too huge. Saying that science has proved that God doesn't exist seems like boldly announcing there is no outer-space alien or non-carbon-based life form after using our available equipment to search our known universe. Further, Stenger did not use Bayesian probability properly in computing the probability of God's existence. Because he omits trimmed means, sensitive analysis, and model averaging, his conclusion $AE = EA$ based on Bayesian is invalid. Stenger heavily relies on hypothesis testing, but exclude meta-analysis and data mining as acceptable research methods. However, his concepts of all of these statistics methods are incorrect. He did not understand that it is common that the finding from separate analysis could be different from that of aggregate analysis, and meta-analysis is one of the standard approaches in data analysis. Further, he did not understand that data mining is an extension of exploratory data analysis, and developing a data-driven hypothesis is acceptable and even desirable. More importantly, he misconstrued the meaning of hypothesis testing even though he viewed it as the most rigorous scientific research method. Take all of the above into account, he did not build the case $AE = EA$; he did not prove there is no God.

References

- Altman, Douglas & Bland, Martin. Statistics notes: Absence of evidence is not evidence of absence. *BioMedical Journal*, 311 (1995): 485.
- Appleton, D. R., & French, J. M. Ignoring a covariate: An example of Simpson's paradox. *American Statistician*, 50(1996): 340–341.
- Baker, R., & Dwyer, F. 2000 February. *A meta-analytic assessment of the effects of visualized instruction*. Paper presented at the 2000 AECT National Convention. Long Beach, CA.
- Bartholomew, David. Victor Stenger's scientific critique of Christian Belief. *Science and Christian Belief* 22(2010): 117-131.
- Behrens, J. T., & Yu, C. H. (2003). Exploratory data analysis. In J. A. Schinka & W. F. Velicer, (Eds.), *Handbook of psychology Volume 2: Research methods in psychology* (pp. 33-64). New Jersey: John Wiley & Sons, Inc.
- Davies, Paul. 2010. *The Eerie Silence: Renewing our search for alien intelligence*. Chicago, IL: Houghton Mifflin Harcourt.
- Felli, James, & Hazen, Gordon. A Bayesian approach to sensitivity analysis. *Health Economics*, 8(1999): 263-268.
- Glass, G. V. Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(1976): 3-8.
- Glass, G. V., McGraw, B., & Smith, M. L. 1981. *Meta-analysis in social research*. Beverly Hills, CA: Sage Publications.
- Han, J., & Kamber, M. 2006. *Data mining: Concepts and techniques* (2nd ed.). Boston, MA: Elsevier.
- Hunter, J. E., & Schmidt, F. L. 1990. *Methods of meta-analysis: Correcting error and bias in research findings*. Newbury Park, CA: Sage Publications.
- Larose, Daniel. 2005. *Discovering knowledge in data: An introduction to data mining*. NJ: Wiley-Interscience.
- Luan, J. 2002. Data mining and its applications in higher education. In A. Serban & J. Luan (Eds.), *Knowledge management: Building a competitive advantage in higher education* (pp. 17-36). PA: Josey-Bass.
- Montgomery, J. M., & Nyhan, B. Bayesian model averaging: Theoretical developments and practical applications. *Political Analysis*, 18(2010): 245-270.
- Morton, William. South African diamond fields, and a journey to the mines. *Journal of the American Geographical Society of New York*, 9(1877): 66-83.

- Nigel, Blundell. 1980. *The world's greatest mistakes*. London: Octopus Books.
- Olkin, I. 2000 November. *Reconcilable differences: Gleaning insight from independent scientific studies*. ASU Phi Beta Kappa Lecturer Program, Tempe, Arizona.
- Skene, A. M, Shaw, J. E. H., & Lee T. D.. Bayesian modeling and sensitivity analysis. *Journal of the Royal Statistical Society. Series D*, 35 (1986): 281-288.
- Sober, E., 2000. *Philosophy of biology* (2nd ed.) Boulder, CO: West View Press.
- Stenger. Victor J. 2007. *God: The failed hypothesis. How science shows that God does not exist* (Kindle Edition). Amherst, NY: Prometheus Books.
- Stenger. Victor J. 2009. *Quantum Gods: Creation, Chaos, and the Search for Cosmic Consciousness*. Amherst, NY: Prometheus Books.
- Stenger, Victor J. 2011. *The fallacy of fine-tuning: Why the universe is not designed for us*. Amherst, NY: Prometheus Books.
- Unwin, Stephen. 2003. *The probability of God: A simple calculation that proves the ultimate truth*. New York, NY: Three Rivers Press.
- Van Fraassen, Bas C. 1980. *The scientific image*. New York : Oxford University Pres.
- Yu, C. H., Jannasch-Pennell, A., DiGangi, S., Kim, C., & Andrews, S. (2007). A data visualization and data mining approach to response and non-response analysis in survey research. *Practical Assessment, Research and Evaluation*, 12(19). Retrieved from <http://pareonline.net/getvn.asp?v=12&n=19>
- Yu, C. H., DiGangi, S., Jannasch-Pennell, A., & Kaprolet, C. (2008). Profiling students who take online courses using data mining methods. *Online Journal of Distance Learning Administration*, 11(2) Retrieved from <http://www.westga.edu/~distance/ojdla/summer112/yu112.html>
- Yu, C. H. (2010). Exploratory data analysis in the context of data mining and resampling. *International Journal of Psychological Research*, 3(1), 9-22. Retrieved from <http://mvint.usbmed.edu.co:8002/ojs/index.php/web/article/download/455/460>
- Yu, C. H., DiGangi, S., Jannasch-Pennell, A., & Kaprolet, C. (2010). A data mining approach for identifying predictors of student retention from sophomore to junior year. *Journal of Data Science*, 8, 307-325. Retrieved from <http://www.jds-online.com/>